



Universidad Internacional de La Rioja  
Facultad de Ciencias de la Salud

Máster Universitario en Bioinformática

**Revisión exhaustiva de los marcadores  
nutrigenómicos en la celiaquía: Nuevas  
estrategias de identificación mediante  
bioinformática, prevención y diagnóstico.**

Trabajo fin de Estudio presentado por:	Adela Arias Rodríguez
Tipo de trabajo:	Bibliográfico
Director/a:	Víctor Jesús Fernández Ramírez
Fecha:	05-03-25

## Resumen

La enfermedad celíaca (EC) es una afección autoinmune crónica desencadenada por la ingesta de gluten en personas genéticamente predispostas, causando inflamación y daño en el intestino delgado, lo que afecta la absorción de nutrientes y puede generar complicaciones si no se trata. Recientemente, se han logrado avances en la identificación de biomarcadores para la EC, especialmente a través de la proteómica y el análisis de microRNA (miRNA). La proteómica ha identificado proteínas diferencialmente expresadas en muestras de duodeno y plasma, sugiriendo alteraciones metabólicas relevantes para el diagnóstico. Por otro lado, los miRNAs, pequeñas secuencias de ARN no codificantes, juegan un papel clave en la regulación inmunológica y la patogénesis de la EC. Los miRNAs circulantes, por su estabilidad en fluidos biológicos, son prometedores como biomarcadores no invasivos, facilitando el diagnóstico y seguimiento de la enfermedad. En definitiva, la bioinformática ha sido esencial para entender las complejas redes moleculares asociadas a la EC. Gracias a los análisis bioinformáticos, se han identificado genes clave involucrados en las vías de señalización inmunológica y autoinmune, proporcionando una comprensión más profunda de la fisiopatología de la enfermedad. Estas herramientas permiten procesar grandes volúmenes de datos genómicos y proteómicos, facilitando la identificación de biomarcadores potenciales que pueden mejorar el diagnóstico y tratamiento de la EC. Sin embargo, se necesitan más estudios para validar estos hallazgos en cohortes más grandes y explorar su aplicabilidad clínica.

**Palabras clave:** Enfermedad celíaca, bioinformática, proteómica, miRNA

## Abstract

Celiac disease (CD) is a chronic autoimmune condition triggered by the ingestion of gluten in genetically predisposed individuals, causing inflammation and damage in the small intestine, which affects nutrient absorption and can lead to complications if not properly managed. Recently, significant progress has been made in identifying biomarkers for CD, especially through proteomics and microRNA (miRNA) analysis. Proteomics has identified differentially expressed proteins in duodenal and plasma samples, suggesting metabolic alterations relevant for diagnosis. On the other hand, miRNAs, small non-coding RNA sequences, play a key role in regulating the immune system and the pathogenesis of CD. Circulating miRNAs, due to their stability in biological fluids, are promising as non-invasive biomarkers, facilitating the diagnosis and monitoring of the disease.

In conclusion, bioinformatics has been essential in understanding the complex molecular networks associated with CD. Through bioinformatic analyses, key genes involved in immune and autoimmune signalling pathways have been identified, providing a deeper understanding of the disease's pathophysiology. These tools allow the processing of large volumes of genomic and proteomic data, aiding in the identification of potential biomarkers that could improve CD diagnosis and treatment. However, further studies are needed to validate these findings in larger cohorts and explore their clinical applicability.

**Keywords:** Celiac disease, bioinformatics, proteomics, miRNA

## Índice de contenidos

1. Marco teórico .....	8
1.1. Etiopatogenia.....	8
1.1.1. Gluten dietético .....	8
1.1.2. Genética de la EC.....	9
1.1.3. Inmunología de la EC .....	13
1.1.4. Paso de péptidos de gluten a través del epitelio .....	14
1.1.5. Factores inmunológicos: respuesta adaptativa al gluten.....	16
1.1.6. Factores inmunológicos: inmunidad innata .....	18
1.1.7. MICROBIOTA INTESTINAL EN LA EC.....	19
1.2. Diagnóstico de la enfermedad celíaca.....	21
1.2.1. Sintomatología.....	21
1.2.2. Pruebas serológicas .....	22
1.2.3. Biopsia del intestino delgado .....	23
1.2.4. Pruebas complementarias.....	23
1.3. Tratamiento .....	24
1.4. La bioinformática en la nutrición.....	24
1.4.1. La bioinformática en la celiaquía.....	25
2. Justificación .....	26
3. Objetivos .....	26
3.1. objetivo general.....	26
3.2. Objetivos específicos .....	27
4. materiales y métodos.....	27
4.1. Búsqueda inicial .....	27
4.2. Búsqueda sistemática.....	27

5. Resultados .....	31
5.1. selección de estudios.....	31
5.2. Características de los artículos .....	31
5.3. Resultados de los artículos seleccionados .....	35
5.3.1. Proteómica .....	35
5.3.2. microRNA en la EC.....	42
6. Discusión .....	53
7. Conclusiones.....	54
8. Referencias bibliográficas .....	54

## Índice de figuras

Figura 2: Representación esquemática de los principales procesos implicados en la patogenia de la enfermedad celíaca .....	14
Figura 3. Transporte de gliadina a través de la capa epitelial.....	15
Figura 4: Diagrama de flujo PRISMA.....	30
Figura 6: Tres métodos de aprendizaje automático o machine learning para seleccionar moléculas características.....	36
Figura 7: Agrupamiento de proteínas plasmáticas y análisis funcional.....	37
Figura 8: Tres métodos de aprendizaje automático o machine learning para seleccionar moléculas características.....	38
Figura 9: Diagrama de Venn de la intersección entre todos los genes de la EC obtenidos de 3 bases de datos pública .....	39
Figura 10: La red relacionada con los 54 genes de la EC se construyó utilizando el software Cytoscape.....	39
Figura 11: Análisis funcional de la red de interacción establecida por los 11 genes principales obtenidos de GeneMANIA.....	40
Figura 12: Interacción proteína-proteína de los 54 genes en WebGestalt.....	41
Figura 13: Diagrama de Venn de la intersección entre los genes comunes de la EC en 3 bases de datos públicas y los genes en los conjuntos de datos de GEO (datos de microarreglo y RT-PCR).....	42
Figura 14:Análisis PCA y mapa de calor.....	45
Figura 15: Análisis de expresión de miRNA y genes en la enfermedad celíaca (EC).....	46
Figura 16: Red de interacción miRNA-diana en la enfermedad celíaca (EC).....	47
Figura 17: Análisis de enriquecimiento de los genes objetivo, asociados con los miARN individuales dentro de la red de transcripción miARN-objetivo .....	48
Figura 18: Subconjunto de la red de interacción miRNA-transcripción diana.....	49

## Índice de tablas

Tabla 1. Frecuencia de haplotipos y alelos en la población con enfermedad celiaca.....	11
Tabla 2. Patogénesis de la enfermedad celíaca .....	20
Tabla 3: Características de los estudios seleccionados. ....	33
Tabla 4. Resultados de los artículos seleccionados.....	50

## 1. Marco teórico

La enfermedad celíaca (EC) es una enteropatía con base inmunológica que afecta a la mucosa del intestino delgado y es producida por la intolerancia al gluten/gliadina (prolамина) presente en el trigo, la cebada y el centeno. Se caracteriza por ser una enfermedad multifactorial, en el que intervienen factores genéticos y ambientales (1). Su prevalencia ha aumentado en los últimos años (se estima en torno a un 1% de la población), debido principalmente a la mayor implantación del cribado en individuos de riesgo y a la mejora de herramientas de diagnóstico (2).

### 1.1. Etiopatogenia

En el paciente celíaco interactúan 4 componentes bien definidos: el gluten, la transglutaminasa tisular (TG2), el/los haplotipos HLA-DQ2/DQ8 y los linfocitos T(2-4).

#### 1.1.1. Gluten dietético

El gluten es un conjunto de proteínas de origen vegetal que está presente en los cereales como el trigo, la cebada y el centeno, entre otros. Está compuesto por dos fracciones: las prolaminas, que se caracterizan por ser solubles en alcohol; y las gluteninas, que son insolubles en alcohol. Dependiendo del tipo de prolamina que el cereal tenga, se considerará perjudicial o no para las personas con celiaquía. Aquellas prolaminas que están categorizadas como dañinas para los celíacos son la gliadina (trigo), la secalina (centeno) y la hordeína (cebada) (4-9). La avena en la EC es un tema controvertido, pues se conoce que la mayoría de los pacientes la toleran. Se sabe que contiene alrededor del 10% de contenido del gluten que tienen el trigo, por lo que puede desarrollar EC en algunos pacientes (3,4).

Las prolaminas tienen un mecanismo de acción que se distingue por su alta resistencia a la digestión, debido a su elevado contenido en los aminoácidos prolina y glutamina. Como resultado de esta resistencia a la degradación enzimática, en el intestino permanecen fragmentos de gluten de gran tamaño. Estos péptidos atraviesan la barrera intestinal hasta la lámina propia, donde son capturados y procesados por las células dendríticas. En este proceso, la transglutaminasa 2 (TG2), una enzima con múltiples funciones, modifica los residuos de glutamina en posiciones específicas de los péptidos del gluten mediante un proceso de deaminación. Esta alteración aumenta la afinidad de los péptidos por los alelos de

susceptibilidad del complejo HLA de clase II. De hecho, existen algoritmos diseñados para predecir secuencias tóxicas basándose en la capacidad de los péptidos de interactuar con dichos alelos, lo que implica la necesidad de la deaminación de glutaminas por la TG2. La combinación de estos algoritmos con estudios sobre la reactividad de los linfocitos T sugiere que la respuesta adaptativa de estas células se dirige principalmente a ciertos péptidos del gluten que cumplen con los criterios de alta afinidad por los alelos de susceptibilidad HLA y la modificación selectiva por la TG2(4).

Se considera la gliadina como la prolamina con el mayor componente tóxico para los pacientes celíacos, en concreto predomina la reactividad de los linfocitos T frente a las  $\alpha$ -gliadinas. Los péptidos inmuno-dominantes, tales como la  $\alpha$ -gliadina p56-89, inducen respuestas inmunes específicas en todos los pacientes con EC. La  $\alpha$ -gliadina está compuesta por 33 aminoácidos (33-mer) capaz de resistir a las proteasas gástricas, pancreáticas e intestinales (4,7,8).

Gracias a los avances en el conocimiento de la patogenia de la enfermedad celíaca, se han identificado dos tipos de péptidos dañinos: los péptidos tóxicos (19-mer), que inducen una respuesta innata con un efecto directo sobre el epitelio, provocando cambios rápidos en la mucosa intestinal, y los péptidos inmunogénicos (33-mer), que activan una respuesta adaptativa al estimular los linfocitos T del intestino o la sangre periférica de pacientes celíacos con restricción HLA-DQ2/DQ8. Ambas vías se interrelacionan y refuerzan mutuamente, contribuyendo al mantenimiento del daño intestinal crónico. Sin embargo, aún es necesario identificar otras secuencias de péptidos tóxicos, determinar sus receptores, comprender mejor su interacción con los enterocitos y el mecanismo de transporte transepitelial de estos péptidos(4).

En conclusión, estos péptidos actúan como antígenos estimulando reacciones inmunitarias adversas en individuos genéticamente susceptibles. Debido a estas reacciones inmunitarias, se produce una respuesta inflamatoria y daño en la mucosa intestinal (4-9). Estudios recientes han demostrado que el gluten afecta a la flora del tracto digestivo y disbiosis microbiana oral (10).

### 1.1.2. Genética de la EC

La enfermedad celíaca (EC) es una condición multifactorial influida por factores genéticos, inmunológicos y ambientales, aunque los factores genéticos tienen un peso mucho mayor en

el riesgo de desarrollarla (11). Diversos estudios han demostrado que la EC tiene un fuerte componente hereditario. De hecho, se estima que aproximadamente el 87% del riesgo de padecer EC está relacionado con la genética, mientras que el impacto de los factores ambientales es significativamente menor (12).

El descubrimiento de la diferencia entre gemelos mono y dicigóticos han permitido a los científicos estimar el elevado componente genético de la EC. Este componente es más elevado comparado con otras enfermedades inmunológicas, tales como la diabetes tipo 1. Así, se ha estimado una concordancia para la EC de alrededor del 75-86% para gemelos monocigóticos, mientras que, en los gemelos dicigóticos la proporción es del 16-20% (12,13).

La mayor porción del riesgo genético de desarrollar EC proviene de la presencia de ciertos alelos del HLA, aunque se han identificado otros genes no HLA que confieren una predisposición a la enfermedad. Además, estos genes están involucrados en las respuestas inflamatorias e inmunes (4,11,14).

#### 1.1.2.1. Genes tipo HLA

El HLA es la denominación del Complejo Mayor de Histocompatibilidad (MHC, por sus siglas en inglés) en los seres humanos. Se ubica en la región cromosómica 6p21 y agrupa numerosos genes implicados en la respuesta inmune. Las proteínas codificadas por estos genes desempeñan un papel clave en la presentación de antígenos, ya que se expresan en la mayoría de las células humanas y permiten al organismo diferenciar entre moléculas propias y externas (4).

El sistema HLA comprende distintas clases de moléculas, pero las de clase II son las más relevantes en la enfermedad celíaca, destacando especialmente HLA-DQ2 y HLA-DQ8. Estas moléculas se expresan en la superficie de células presentadoras de antígenos (APC), como las células dendríticas, los macrófagos y los linfocitos B. Su función principal es la presentación de antígenos, como el gluten, al sistema inmunológico, específicamente a los linfocitos T auxiliares (CD4+) (4).

La molécula DQ está compuesta por una cadena α (codificada por el gen HLA-DQA1) y otra cadena β (codificada por el gen HLA-DQB1) y conforma un heterodímero. Según la combinación de alelos de DQA1 y DQB1, esta molécula estará asociada o no a la EC. Si en el DQA1 se encuentra el alelo 05:01 y en el DQB1 se tiene el alelo 02:01, la combinación dará

como resultado el haplotipo HLA-DQ2. Por otro lado, si encontramos en el DQA1 el alelo 03:01 y en el DQB1 encontramos el alelo 03:02, la combinación dará como resultado el haplotipo HLA-DQ8. También encontramos el haplotipo Half DQ2, utilizado para describir la situación genética en la que la persona es portadora de alguno de los alelos del haplotipo DQ2 por separado (DQA1\*0501 o DQB1\*0201) (4,11,12,14,15).

El diagnóstico genético de los genes HLA-DQ2 y HLA-DQ8 son cruciales para el diagnóstico de la celiaquía. Casi todos los pacientes celíacos (99%) son portadores de uno de estos haplotipos. Se ha estimado que el 90% de los pacientes con EC expresan el haplotipo HLA-DQ2, alrededor del 5% de los pacientes con EC expresan el haplotipo HLA-DQ8 y casi el 5% de los individuos restantes son portadores de al menos uno de los dos genes (Half DQ2)(14).

**Tabla 1. Frecuencia de haplotipos y alelos en la población con enfermedad celiaca**

Haplótipos	Alelos	Frecuencia en pacientes con EC
DQ2	DQA1*0501 y DQB1*0201	Aprox. 90%
DQ8	DQA1*0301 y DQB1*0302	Aprox. 5-10%
Half DQ	DQA1*0501 o DQB1*0201	

Tabla 1: Frecuencia de haplotipos y alelos en la población con enfermedad celíaca. Fuente:  
Elaboración propia

Los haplotipos DR3 y DR4 pertenecen al HLA de clase II, específicamente al gen HLA-DRB1. Estos haplotipos se heredan junto con otros genes que afectan a HLA-DR, HLA-DQ y HLA-DP. Las variantes HLA-DQ2 y HLA-DQ8 están en desequilibrio de ligamiento con DR3 y DR4, respectivamente, por lo que suelen denominarse haplotipos de riesgo DR3-DQ2 y DR4-DQ8.

En algunos casos, como en el haplotipo DR3-DQ2, los dos alelos que conforman el heterodímero HLA-DQ2.5 (DQA105:01 y DQB10201) se encuentran en el mismo cromosoma, lo que se conoce como codificación en cis. En cambio, en individuos heterocigotos con los haplotipos DR5-DQ7 y DR7-DQ2, los dos alelos del heterodímero de riesgo están ubicados en cromosomas distintos, lo que se denomina codificación en trans.

Los haplotipos HLA-DQ2.5, tanto en cis como en trans, representan un alto riesgo de desarrollar enfermedad celíaca. Sin embargo, la variante HLA-DQ2, específicamente el dímero HLA-DQ2.2, se asocia con un riesgo significativamente menor (Figura 1) (15).

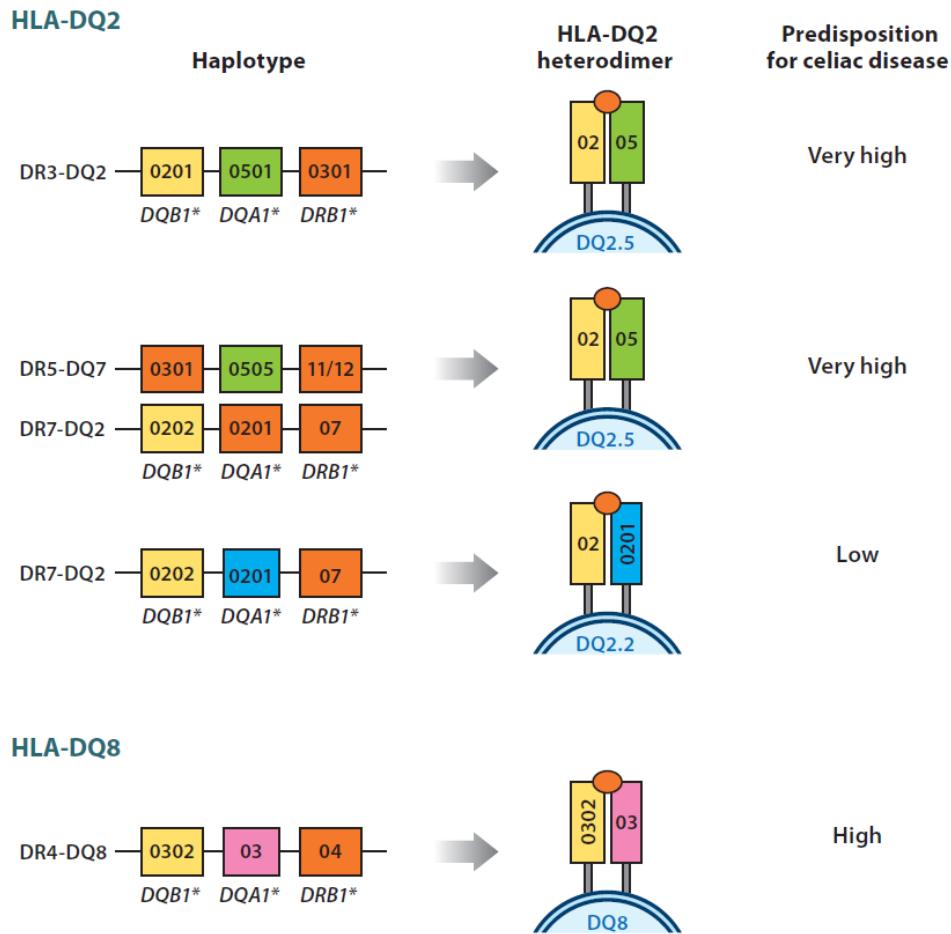


Figura 1: Relación entre la enfermedad celíaca (EC) y el locus HLA. La molécula HLA-DQ2 es el principal determinante genético de riesgo para la EC. La mayoría de los pacientes celíacos presentan el heterodímero HLA-DQ2.5, el cual está codificado por los alelos HLA-DQA105 (cadena  $\alpha$ ) y HLA-DQB102 (cadena  $\beta$ ), pudiendo encontrarse en cis en el haplotipo DR3-DQ2 o en trans en individuos heterocigotos DR5-DQ7 y DR7-DQ2.2. Una variante de HLA-DQ2, el dímero HLA-DQ2.2, codificado por los alelos HLA-DQA102:01 y HLA-DQB102:02, se asocia con un riesgo bajo de desarrollar la enfermedad. En los casos en los que DQ2 está ausente, la mayoría de los pacientes expresan HLA-DQ8, codificado por el haplotipo DR4-DQ8 (15).

La susceptibilidad a la EC también está relacionada con la cantidad de heterodímeros DQ2.5 presentes en el organismo. Aquellos individuos que son homocigotos para el haplotipo DR3-DQ2, así como los heterocigotos con la combinación DR3-DQ2/DR7-DQ2, presentan los niveles

más elevados de estos heterodímeros, lo que se asocia con el mayor riesgo genético de desarrollar la enfermedad (4,15).

Por otro lado, los pacientes con EC refractaria (pacientes que no mejoran con una dieta estricta sin gluten o DSG) tienen una mayor frecuencia de homocigosidad para DR3-DQ2 (44-62%) en comparación con otros pacientes celíacos (20-24%). Además, se ha propuesto que un efecto similar, dependiente de la dosis, también podría darse en el caso de las moléculas DQ8 (4).

Existe, por tanto, una fuerte asociación entre los genes que codifican las moléculas HLA de clase II y la EC, sin embargo, la presencia de dichos genes no es suficiente para desarrollar la EC. Un 30% de la población general presenta los alelos de riesgo, pero sólo el 1% de este grupo es diagnosticado con EC(16). Por tanto, un estudio genético DQ2/DQ8 positivo no implica un diagnóstico de EC. Así, el papel del HLA queda relegado a casos dudosos y grupos de riesgo, ya que la prueba genética posee un alto valor predictivo negativo y la ausencia de DQ2 y DQ8 prácticamente excluye la EC (17,18).

### 1.1.3. Inmunología de la EC

El mecanismo de patogenia en la EC aún es desconocido, pero el modelo más aceptado sostiene que depende de la inmunidad adaptativa tras la estimulación de los linfocitos T CD4+. Estos son estimulados por péptidos del gluten que han sido modificados por la transglutaminasa tisular y la restricción por moléculas HLA-DQ2/DQ8, que producen citocinas proinflamatorias. Cabe mencionar, que el gluten también activa la inmunidad innata y la citotoxicidad epitelial mediada por los linfocitos intraepiteliales (3,4).

Sin embargo, aún no se tiene explicación de por qué sólo algunos individuos portadores del genotipo HLA de riesgo (HLA-DQ2 o HLA-DQ8) desarrollan la EC. Algunos estudios han hablado de la influencia de los factores ambientales (por ejemplo, la composición de la flora intestinal) en la capacidad de inducir una respuesta inmune innata, aumentando la probabilidad del individuo de desarrollar la EC(3,10).

El modelo patogénico combina diversos factores, incluyendo los que afectan al epitelio de la lámina propia (LP), la digestión parcial y el transporte transepitelial de péptidos, el efecto directo del gluten sobre el epitelio, la activación y proliferación de linfocitos intraepiteliales

(LIE), y el reconocimiento de péptidos de gluten por linfocitos T específicos con restricción HLA-DQ2, tras su modificación por la transglutaminasa tisular (TG2) (Figura 2) (3,4).

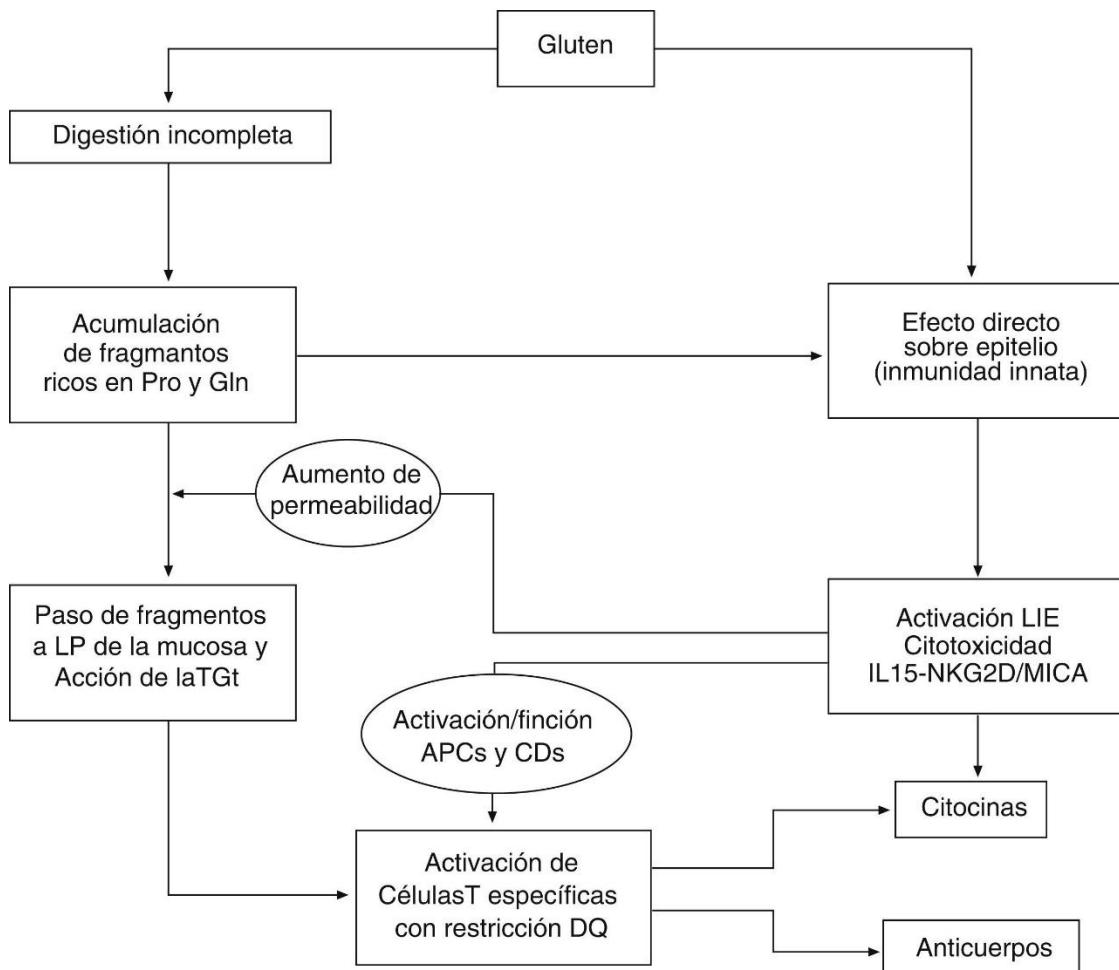


Figura 2: Representación esquemática de los principales procesos implicados en la patogenia de la enfermedad celíaca(3).

#### 1.1.4. Paso de péptidos de gluten a través del epitelio

En condiciones normales, el paso de los péptidos del gluten a través del epitelio intestinal se produce porque las proteínas son previamente hidrolizadas por enzimas (peptidasas gástricas y pancreáticas), generando péptidos de menor tamaño o aminoácidos libres. Estos atraviesan el epitelio intestinal mediante el contrtransporte dependiente de iones de hidrógeno y del transporte activo secundario acoplado con sodio. Algunos péptidos, debido a su tamaño, no se pueden absorber y se acumulan en la luz intestinal. Para que puedan atravesar el epitelio lo podrán hacer mediante una de las 4 rutas alternativas (Figura 3) (3,4):

- Vía paracelular: Ocurre por efecto de la zonulina, una proteína que modifica la estructura del citoesqueleto celular y compromete las uniones estrechas del epitelio intestinal. La gliadina estimula su liberación, lo que debilita la conexión entre los enterocitos y permite el tránsito de péptidos a través de la barrera intestinal.
- Vía transcelular: Involucra la endocitosis de péptidos por los enterocitos, su degradación en los lisosomas y su posterior transporte hacia la membrana basal. En la enfermedad celíaca, esta ruta parece estar alterada, facilitando el tránsito de péptidos no degradados hacia la LP.
- Retrotranscitosis: Se fundamenta en la función del receptor CD71 de transferrina, presente en los enterocitos. Este receptor favorece el transporte de complejos formados por péptidos de gliadina y anticuerpos IgA secretora (sIgA1), evitando su degradación enzimática a lo largo de su paso por el epitelio.
- Captación directa: Ciertas células dendríticas mieloides de la LP, que presentan HLA-DQ y TG2 en su membrana, son capaces de captar estos péptidos directamente y promover su procesamiento por el sistema inmunológico.

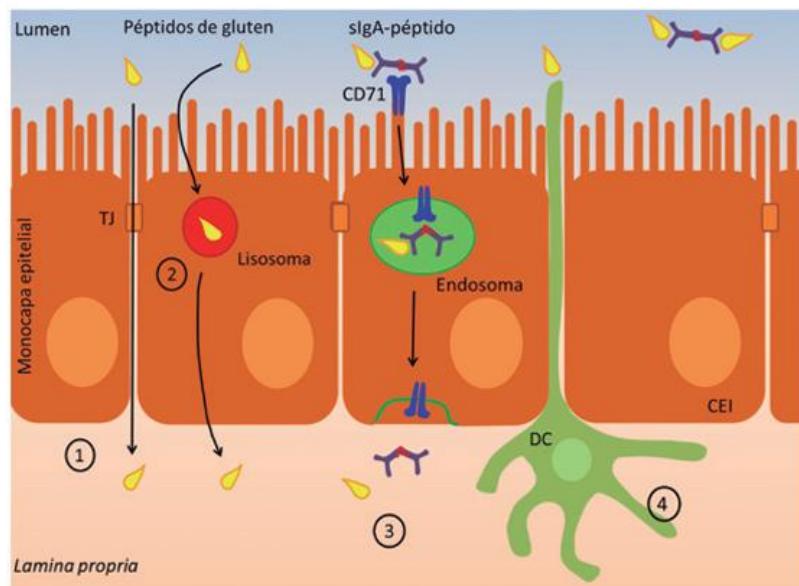


Figura 3. Transporte de gliadina a través de la capa epitelial. 1) Ruta paracelular. 2) Ruta transcelular. 3) Retrotranscitosis. 4) Captación directa. TJ: tight-junctions (uniones estrechas); CEI: células epiteliales in testinales; DC: célula dendrítica; sIgA: inmunoglobulina A secretora(4).

El paso de péptidos del gluten a través del epitelio no solo está determinado por la función de la barrera intestinal, sino también por la expresión génica y las señales de fosforilación involucradas en el metabolismo, la proliferación y la adhesión celular. En pacientes con enfermedad celíaca activa, se ha observado un incremento en el transporte transepitelial mediado por el interferón  $\gamma$  (INF- $\gamma$ ), liberado por células T activadas. Además, el procesamiento de estos péptidos en el epitelio está alterado, lo que podría permitir el paso de péptidos tóxicos (19-mer) e inmunogénicos (33-mer) a la LP (4).

#### 1.1.5. Factores inmunológicos: respuesta adaptativa al gluten

##### 1.1.5.1. Transglutaminasa tisular (TG2)

La activación de la respuesta inmune adaptativa al gluten en la EC se produce por la transglutaminasa tisular (TG2). La TG2 induce la modificación enzimática de los

La transglutaminasa tisular (TG2) es el componente que explica la activación de la respuesta inmune adaptativa al gluten. La TG2 provoca la modificación enzimática de los péptidos de gliadina inmunodominantes, lo que resulta en la adición de cargas negativas en determinados residuos de aminoácidos. Este cambio aumenta la afinidad de los péptidos por las moléculas HLA-DQ2/DQ8 (4,19,20)(Figura 4).

La TG2 es una enzima que participa en los procesos de apoptosis celular y remodelación de la matriz extracelular durante la reparación de tejidos. La TG2 se encuentra en algunos órganos y tejidos del cuerpo. Generalmente la encontramos a nivel intracelular, pero aparece fuera de las células cuando se produce daño tisular. En condiciones normales, la TG2 se expresa en el subepitelio, en la mucos de la LP y en el tejido conectivo que circula por las criptas. Pero, en pacientes con EC, se expresa adicionalmente en la superficie apical de los enterocitos (4,19,21).

Como se ha mencionado antes, esta enzima está involucrada en el mecanismo de transporte por retrotranscitosis para el paso de los péptidos de gliadina a través del epitelio. Se ha demostrado que la TG2 puede interactuar con el receptor de la transferrina CD71 y con las IgAs en la superficie apical de los enterocitos (4).

Es importante señalar que los efectos de la TG2 sobre los péptidos del gluten también ocurren en condiciones con pH inferior a 7.0, es decir, en situaciones no fisiológicas. Bajo estas condiciones, la gliadina, que contiene más del 30% de glutamina, es susceptible a las

modificaciones inducidas por la enzima. Este proceso es clave en la enfermedad celíaca, ya que los péptidos desaminados tienen una mayor afinidad por las moléculas HLA-DQ, especialmente hacia las moléculas HLA-DQ2 (4,22).

#### 1.1.5.2. Linfocitos T CD4+

La respuesta adaptativa en la EC está mediada por las células T específicas de la LP. La activación de la respuesta requiere de la presentación de antígenos por las células presentadoras de antígenos (CPAs), que expresan en sus membranas las moléculas HLA-DQ2/DQ8 (4). Estas moléculas poseen un sitio de unión a péptidos con características especiales que les permiten adaptar secuencias peptídicas: HLA-DQ2 favorece los aminoácidos de carga negativa en posiciones centrales (P4, P6 o P7), mientras que HLA-DQ8 prefiere los aminoácidos en posiciones más externas (P1 o P9) (3).

En los individuos que no desarrollan la EC, aproximadamente el 80% de los macrófagos suelen expresar las moléculas HLA-DQ en sus membranas y se caracterizan por tener un fenotipo tolerogénico CD103+CD11c+. En cambio, en los pacientes con EC, la mayoría de CD parecen provenir de monocitos de sangre periférica presentando un fenotipo proinflamatorio (CD14+CD11c+). Se piensa que la diferenciación hacia dicho fenotipo proinflamatorio se debe a la presencia de altos niveles de INF- $\alpha$  en la mucosa intestinal de los pacientes con EC (23).

Las moléculas HLA-DQ2 y DQ8 presentan nuevos epítopos complejos formados por TG2-gluten-péptidos a las células T CD4+ en la mucosa de la lámina propia (LP). Los linfocitos T activados provocan una respuesta proinflamatoria, caracterizada por la liberación de citoquinas Th1, con predominio de IFN- $\gamma$ , IFN- $\alpha$ , IL-18 e IL-21, junto con una disminución de las citoquinas reguladoras IL-10 y TGF- $\beta$ . Este proceso inflamatorio puede ser responsable de las lesiones típicas observadas en la enfermedad celíaca primaria (4,24,25).

#### 1.1.5.3. Respuesta al gluten de las células B

La enfermedad celíaca (EC) se distingue por la presencia de diversos autoanticuerpos séricos y moléculas extrañas. Se ha identificado la TG2 como el principal autoantígeno, con una fuerte reactividad contra el anticuerpo anti-endomisio(26).

Los anticuerpos IgA Anti-TG2, producidos por células plasmáticas en la lámina propia del duodeno, aumentan en los pacientes con enfermedad celíaca activa, especialmente en las áreas lesionadas. Estos depósitos de IgA específicos para TG2 se encuentran en todas las

etapas de la enfermedad, incluso antes de la aparición de síntomas o lesiones intestinales. (26,27).

Las células B son células presentadoras de antígenos profesionales que interactúan con el antígeno a través del receptor BCR. En condiciones normales (sin enfermedad celíaca activa), el intestino contiene pocas células B vírgenes o de memoria, siendo la mayoría plasmoblastos o células plasmáticas en la lámina propia, con baja expresión de moléculas HLA-II(4).

Durante la EC activa, los linfocitos B juegan un papel crucial como células presentadoras de antígenos en los nódulos linfáticos mesentéricos, donde amplifican la respuesta de las células T al gluten. Lo logran al presentar complejos TG2-gluten a las células T CD4+, lo que favorece la producción adicional de anticuerpos (26,27).

Además de su función como células presentadoras de antígenos, los linfocitos B se diferencian en células plasmáticas productoras de anticuerpos, como el anti-TG2 (IgA e IgG), cuyos niveles están elevados en la mucosa duodenal durante la enfermedad celíaca activa y pueden detectarse incluso antes de la aparición de síntomas o daño tisular. También producen anticuerpos anti-gliadina y anti-endomisio, que contribuyen a la respuesta inmune (4,26,27).

Los depósitos de IgA específicos para TG2 en el intestino pueden agravar la inflamación al favorecer la absorción del gluten y activar los receptores Fc en los granulocitos locales. Este proceso se revierte al eliminar el gluten de la dieta, lo que reduce los niveles de anticuerpos anti-TG2 y disminuye la respuesta inflamatoria (4).

#### 1.1.6. Factores inmunológicos: inmunidad innata

En la enfermedad celíaca, algunos péptidos de la gliadina pueden activar la respuesta inmune innata al interactuar con células epiteliales y dendríticas, aunque aún no se comprende completamente su mecanismo de acción. Péptidos como p31-43 y p31-49 no son reconocidos por células T CD4+, pero afectan el tráfico intracelular y provocan estrés celular (4).

Se ha observado un aumento de IL-15, COX-2, CD25 y CD83 en la lámina propia, lo que indica activación inmune. Además, en las células epiteliales, la mayor expresión de moléculas MHC clase I sugiere una intensa respuesta inmunológica (4).

Algunos péptidos de gliadina pueden imitar al EGF, prolongando la activación del EGFR y favoreciendo el daño epitelial, mientras que otros activan células dendríticas a través del TLR4, desencadenando inflamación. También pueden estabilizar HLA-E, afectando la regulación

inmune, y aumentar la permeabilidad intestinal al unirse a CXCR3, facilitando la entrada de más péptidos inmunogénicos (4).

Estos mecanismos muestran cómo la gliadina no solo activa la respuesta adaptativa, sino que también altera la barrera intestinal y promueve inflamación, contribuyendo al daño característico de la enfermedad celíaca (4).

#### 1.1.6.1. Linfocitos Intraepiteliales (LIE)

Los linfocitos intraepiteliales (LIE) son un grupo diverso de células inmunitarias situadas en la base de los enterocitos a lo largo del intestino. Se dividen en LIE naturales (T TCR $\alpha\beta$  y T TCR $\gamma\delta$ ) y LIE inducidos (T TCR $\alpha\beta$  CD4+ y T TCR $\alpha\beta$  CD8 $\alpha\beta+$ ), desempeñando funciones clave en la defensa contra infecciones, la memoria inmunitaria y el mantenimiento de la integridad epitelial (4,28).

A pesar de su papel protector, los LIE pueden contribuir a enfermedades como la celiaquía y la inflamación intestinal. En la enfermedad celíaca (EC), el aumento de células T con TCR $\alpha\beta$  se asocia con atrofia vellositaria y un fenotipo citotóxico. Además, hay un incremento de LIE con TCR $\gamma\delta+$ , una alteración persistente incluso con una dieta sin gluten. Estos linfocitos pueden activarse en presencia de inflamación, interactuar con células epiteliales a través de receptores como NKG2D y CD94, y desencadenar la destrucción de enterocitos, favoreciendo la atrofia vellositaria y la hiperplasia de criptas (4,28).

Otras células inmunitarias también pueden estar implicadas en la EC. Las células natural killer (NK), esenciales en la respuesta contra virus y células tumorales, se encuentran reducidas en pacientes con EC activa, lo que sugiere una alteración en su función (4,28).

Las células NKT, que combinan características de las NK y las células T al expresar TCR y receptores CD3 e IgG, pueden producir IL-10 y otras citoquinas con funciones inmunorreguladoras. Sin embargo, su papel en la EC no está del todo claro, ya que podrían influir en la modulación de la respuesta inmune en la enfermedad (4,28).

#### 1.1.7. MICROBIOTA INTESTINAL EN LA EC

La ingesta de gluten es el principal desencadenante ambiental responsable de las manifestaciones clínicas de la enfermedad, sin embargo, la exposición al gluten no explica completamente su desarrollo(4). Por tanto, se han investigado otros factores ambientales que pueden contribuir a desarrollar la EC como es la composición de la microbiota intestinal(10).

Los estudios realizados sobre la influencia de la microbiota intestinal en la EC son limitados, pero se ha sugerido que una disbiosis intestinal podría influir en la aparición y la progresión de la EC. El efecto de la composición de la microbiota intestinal sobre la etiopatogenia de la EX, podría estar relacionado con su actividad proteolítica y su capacidad para generar péptidos tóxicos e inmunogénicos (10,29–32).

Se han detectado alteraciones en la microbiota intestinal en pacientes con EC en tratamiento de DSG, lo que sugiere que tales desequilibrios no son solamente una consecuencia secundaria de la EC. Los estudios epidemiológicos publicados hasta ahora indican que factores ambientales perinatales y postnatales que influyen en el riesgo de desarrollar EC, también influyen en la composición de la microbiota intestinal. Dichos factores incluyen el tipo de parto (parto natural o parto por cesárea), alimentación del lactante (duración de la lactancia materna), la edad gestacional, infecciones repetidas en etapas tempranas y la ingesta de antibióticos(30,32). Por otro lado, se ha revelado que el genotipo HLA-DQ junto con los factores ambientales podrían influir sobre la composición de la microbiota (29,31). Aunque estos estudios se centran en bebés con alto riesgo de padecer EC, por lo que se puede llegar a conclusiones no generalizables a la población (10).

Se ha propuesto el tratamiento de con probióticos en la EC. Los resultados de los estudios de intervención con bacterias probióticas realizados en humanos son prometedores, pero se necesitan más estudios de un mayor alcance para poder confirmar este tipo de intervenciones en pacientes con EC(33–35).

Por otro lado, se está estudiando la detección de nuevos biomarcadores para el diagnóstico y tratamiento de la EC mediante el estudio de la metabolómica y la lipidómica usando herramientas bioinformáticas. Aunque se recalca la necesidad de más estudios en la población para ampliar el conocimiento actual sobre la metabolómica y la lipidómica en la EC(36).

En la tabla 1 se recoge un resumen de la patogénesis de la enfermedad celíaca(6).

**Tabla 2. Patogénesis de la enfermedad celíaca**

Factores	Fisiopatología
Gluten dietético	<ol style="list-style-type: none"><li>1. Gluten de trigo, centeno y cebada. Las licodinas y las gluteninas son ricas en prolina, lo que las hace resistentes a la proteólisis por enzimas gástricas y pancreáticas. Varios péptidos de gliadina largo activan el sistema inmunitario. Los péptidos no digeridos también pueden afectar la microbiota intestinal.</li></ol>

---

<b>Genética</b>	<ol style="list-style-type: none"><li>1. Predisposición genética: HLA-DQ2 y HLA-DQ8 contribuyen al 20% y 40%, respectivamente, del riesgo genético. Se trata de MHC de clase II expresados por células presentadoras de antígenos (APC).</li><li>2. 42 regiones no-HLA se han asociado con la enfermedad celíaca. Se estima que representan el 15% del riesgo genético: <i>IL18R1</i>, <i>IL18RAP</i>, <i>STAT4STAT4</i>, <i>CD28</i>, <i>CTLA4</i>, <i>ICOS</i>, <i>CCR4CCR1CCR1</i>, <i>CCR1</i>, <i>CCR2</i>, <i>CCR3CD3E</i>, <i>IL1R1IL1R1</i>, <i>IL12A</i>, <i>IL2</i>, <i>IL21</i>, <i>TNFAIP3TNFAIP3</i>, <i>ELMO1</i>, <i>PRKCQ</i>, <i>SOCS1SOCS1</i>, <i>ICOSLG</i> e <i>IRAK1</i>. Estos genes pertenecen a la activación del receptor de citoquinas-citoquinas, la vía JAK-STAT, la señalización de los receptores de células T, la red inmune intestinal para la producción de IgA, la señalización NF-KB y las moléculas de adhesión celular. Cabe destacar que muchos de estos genes pertenecen al punto de control inmunitario y a la vía de inmunooncología.</li></ol>
<b>Inmune</b>	<ol style="list-style-type: none"><li>1. Generación de respuestas de células T específicas de gluten: presencia de linfocitos T CD4+ específicos de gluten, anticuerpos contra gliadina y de enzima TG2, y citocinas proinflamatorias.</li><li>2. Generación de autoanticuerpos: activación y diferenciación en células plasmáticas de linfocitos B específicos de gluten y TG2, generación de autoanticuerpos que circulan y se depositan en la mucosa. Estos autoanticuerpos son responsables del aumento de la permeabilidad de la barrera epitelial.</li><li>3. Citoquinas en el sistema inmunitario de la mucosa intestinal: IFN gamma e IL-21 se producen por linfocitos T CD4+ específicos del gluten. Secreción de IL-15, IL-18 e inhibición de linfocitos T regulatorios FOXP3 positivos (Tregs).</li><li>4. Linfocitos intraepiteliales (IEL): aumento en la enfermedad celíaca y su cantidad se correlaciona con la atrofia mucosa. Los IELs muestran la transformación citotóxica e inducen apoptosis de células epitelial intestinales a través de FAS-L, perforina, gránia B y NKG2D. NKG2D interactúa con MICA en células epiteliales.</li><li>5. Activación innata: desregulación de la producción de IL-15 y activación de la respuesta inmune innata, incluyendo la inducción del estrés epitelial.</li></ol>
<b>Medioambiente</b>	<ol style="list-style-type: none"><li>1. Microorganismos: disbiosis intestinal (microbiota intestinal desequilibrada) y mayor prevalencia de genes específicos de virulencia microbiana aislados de pacientes con enfermedad celíaca.</li><li>2. Otros, fumar</li></ol>

---

La patogenia de la enfermedad celíaca es multifactorial e incluye el gluten de la dieta y factores genéticos, inmunes y ambientales.

Adaptación de Carreras, 2022(6)

## 1.2. DIAGNÓSTICO DE LA ENFERMEDAD CELÍACA

Debido a que la EC es una enfermedad multifactorial, su diagnóstico resulta complicada, lo que hace que sea una enfermedad que se diagnostique con menos frecuencia de lo que se debería. Los métodos diagnósticos de la EC es una combinación de signos y síntomas clínicos, pruebas serológicas y biopsias del intestino delgado (37,38).

### 1.2.1. Sintomatología

La enfermedad celíaca (EC) presenta una gran variabilidad clínica. Los síntomas más frecuentes son los trastornos gastrointestinales tanto en niños como en adultos, aunque

también pueden aparecer manifestaciones extraintestinales, como dermatitis herpetiforme, artritis, neuropatía periférica, anemia por deficiencia de hierro y deficiencia de folatos. En algunos casos, la enfermedad puede ser asintomática y detectarse únicamente mediante cribado en grupos de riesgo, como familiares de pacientes con EC o personas con enfermedades autoinmunes, como la diabetes tipo 1 (37,38).

Existen cinco fenotipos principales de EC (38):

- EC clásica: Se manifiesta principalmente con síntomas digestivos y problemas de absorción intestinal. Afecta a cerca del 50% de los pacientes y se confirma mediante pruebas serológicas positivas y la presencia de atrofia de vellosidades en la biopsia intestinal.
- EC atípica: Predominan síntomas fuera del sistema digestivo, como osteoporosis, trastornos neurológicos y anemia ferropénica, mientras que las molestias digestivas son leves o inexistentes. Actualmente, es la forma más común de la enfermedad.
- EC silenciosa: No provoca síntomas evidentes, pero los pacientes presentan pruebas serológicas positivas y atrofia de vellosidades en la biopsia intestinal.
- EC latente: Se detecta en personas con serología positiva, pero sin daño intestinal visible en la biopsia. Sin embargo, tienen un mayor riesgo de desarrollar síntomas en el futuro.
- EC refractaria: Ocurre cuando los síntomas y el daño intestinal persisten a pesar de una dieta estricta sin gluten. Es poco frecuente y supone un reto en su diagnóstico y tratamiento.

#### 1.2.2. Pruebas serológicas

El diagnóstico de la EC se basa en la combinación de pruebas serológicas y biopsia intestinal. Entre los marcadores más utilizados están los anticuerpos anti-Endomisio (EmA) y los anticuerpos anti-transglutaminasa tisular (TG2), que presentan alta sensibilidad y especificidad. Aunque los anticuerpos EmA son considerados el estándar de oro, su uso es más costoso y subjetivo. Por ello, las pruebas automatizadas de ELISA para anticuerpos anti-TG2 son más comunes en la práctica clínica (37,38).

Las pruebas de anticuerpos anti-gliadina han sido reemplazadas por los ensayos de péptidos de gliadina desamidada (DGP), que pueden detectar casos que no se identifican con las

pruebas convencionales. Sin embargo, estas pruebas no se utilizan ampliamente. En pacientes con deficiencia selectiva de IgA, es necesario realizar pruebas en el isotipo IgG. Aproximadamente un 10% de los pacientes con EC son seronegativos, por lo que su diagnóstico depende de la biopsia y la respuesta a una dieta sin gluten. También existen pruebas rápidas comerciales para detectar anticuerpos anti-DGP y anti-TG2, aunque su uso clínico aún requiere más validación(37,38).

#### 1.2.3. Biopsia del intestino delgado

El análisis histológico del intestino delgado es clave en el diagnóstico de la EC, ya que permite identificar atrofia vellosa, linfocitosis intraepitelial e hiperplasia de criptas. Sin embargo, estos hallazgos no son exclusivos de la enfermedad celíaca y pueden estar presentes en otras condiciones, como infecciones, enfermedades autoinmunes o efectos secundarios de ciertos fármacos. En algunos casos, los pacientes pueden tener anticuerpos positivos sin evidencia clara de daño intestinal en la biopsia, lo que se conoce como enfermedad celíaca potencial. Además, las lesiones pueden ser irregulares y limitarse a áreas específicas del intestino, lo que dificulta la interpretación de las muestras. Para obtener un diagnóstico preciso, es fundamental garantizar la calidad de las biopsias y su correcta evaluación (37,38).

#### 1.2.4. Pruebas complementarias

En situaciones en las que el diagnóstico no es concluyente, pueden emplearse pruebas adicionales, como (38):

- Tipaje genético HLA: La ausencia de los haplotipos HLA-DQ2 o HLA-DQ8 prácticamente descarta la enfermedad celíaca.
- Inmunohistoquímica: La cuantificación de linfocitos intraepiteliales CD3+ y γδ+ en la mucosa intestinal puede ser útil en casos dudosos.
- Detección de autoanticuerpos IgA en la mucosa intestinal: Se pueden identificar depósitos de IgA anti-TG2 en biopsias congeladas.
- Pruebas de respuesta inmunológica al gluten: Métodos como ELISPOT y citometría de flujo permiten detectar células T específicas para gluten en sangre, incluso en pacientes que siguen una dieta sin gluten.

Aunque algunas de estas herramientas son prometedoras, su uso clínico sigue siendo limitado y, en la mayoría de los casos, el diagnóstico se basa en la serología, la biopsia y la respuesta a una dieta libre de gluten(39).

### 1.3. TRATAMIENTO

La EC no tratada tiene consecuencias y complicaciones que pueden llegar a ser graves, incluyendo fracturas patológicas, linfoma intestinal, neoplasias epiteliales (laringe, esófago) y EC refractaria. La EC refractaria es aquella en la que persisten los síntomas y la inflamación o atrofia intestinal a pesar de la dieta sin gluten. Es una forma infrecuente de EC que constituye un reto diagnóstico y terapéutico (38). Diversos estudios han demostrado un aumento en la mortalidad a largo plazo de los pacientes no tratados(39).

El único tratamiento que se conoce actualmente para la EC es una dieta libre de gluten. Si la dieta se lleva de manera adecuada, el 70% de los pacientes mejoran de manera importante en las primeras dos semanas. Sin embargo, los anticuerpos se negativizan entre 6 y 12 meses después de iniciar la dieta. En cuanto a la recuperación de la histología en el intestino puede tardar años. Además de la dieta, los pacientes tienen que prestar atención a las deficiencias vitamínicas(38,39).

Actualmente se encuentran en investigación y desarrollo diferentes tratamientos y avances tecnológicos para el tratamiento complementario junto con la dieta sin gluten. La necesidad de tratamientos alternativos se debe a que la dieta libre en gluten es ineficaz para algunos pacientes (entre el 2 y el 5%). Los enfoques innovadores abarcan un amplio espectro de estrategias, incluida la degradación enzimática del gluten, la inhibición de la permeabilidad intestinal, la modulación de la respuesta inmunitaria, la inhibición de la enzima transglutaminasa 2 (TG2), el bloqueo de la presentación de antígenos por HLA-DQ2/8 y la inducción de tolerancia (40,41).

### 1.4. La bioinformática en la nutrición

La bioinformática es una disciplina que une la Informática y las ciencias biológicas. Dicha disciplina apareció durante la última década del siglo XX gracias a los avances de la ingeniería genética y las nuevas tecnologías de la información (42).

La necesidad de la bioinformática surgió del alto volumen de datos de los resultados experimentales de hoy en día. Mientras, que hace unos años, los resultados de experimentos podían interpretarse sobre el cuaderno de laboratorio, hoy se necesitan bases de datos y técnicas de visualización sólo para almacenarlos y comenzar a estudiarlos (42,43).

Actualmente vivimos en la llamada "era post-genómica", donde no solo la genómica, sino también otras disciplinas "ómicas", dependen de la bioinformática para integrar y analizar los enormes volúmenes de datos que generan. La bioinformática, como campo transversal, reúne múltiples disciplinas científicas y se enfoca en el desarrollo de herramientas que permitan comprender el flujo de información desde los genes hasta las estructuras moleculares. Este conocimiento abarca desde sus funciones bioquímicas y fisiológicas hasta su impacto en la salud y las enfermedades, posicionándola como un pilar esencial en la biología moderna y la medicina de precisión (44).

Uno de los campos en los que se está empezando a incluir la bioinformática es en el campo de la nutrición. Científicos, en los que se incluyen nutricionistas, hacen hincapié en el potencial de la bioinformática en la nutrición de precisión y su papel, clave para mejorar la salud y prevenir enfermedades crónicas no transmisibles como las cardiovasculares, la diabetes o el síndrome metabólico (45).

#### 1.4.1. La bioinformática en la celiaquía

Actualmente, la EC se diagnostica mediante una combinación de los síntomas de los pacientes, pruebas serológicas, biopsias endoscópicas del intestino delgado (métodos invasivos) y, a veces, el tipaje de los genes HLA-DQ2/DQ8. Hasta el momento, no se ha desarrollado una prueba diagnóstica que excluya totalmente los métodos invasivos y la nueva tendencia es el uso de la bioinformática con el objetivo de descubrir nuevos marcadores biológicos para el diagnóstico de la EC y la evaluación pronóstica(36).

Las herramientas bioinformáticas permiten analizar grandes volúmenes de datos genómicos y microbiológicos, facilitando la identificación de nuevas variantes genéticas asociadas a la enfermedad, así como de alelos de riesgo diferentes al HLA-DQ2/DQ8, como por ejemplo, el estudio de microRNA (miARN) (46). La identificación de nuevos biomarcadores en la EC mediante herramientas bioinformáticas también incluye el estudio de la metabolómica, la lipidómica o la proteómica(36,47).

En conclusión, estas técnicas están ayudando a mejorar la precisión en la detección de alelos HLA y no HLA, así como a identificar patrones microbianos que podrían influir en el desarrollo de la enfermedad, abriendo nuevas vías para terapias dirigidas (6).

## 2. Justificación

El estudio de la EC a través del uso de herramientas bioinformáticas se justifica por su incremento en la incidencia de la enfermedad. Se piensa que los cambios en los patrones de alimentación actuales son la causa de este incremento. Algunos de estos cambios son el aumento en la consumición de comida ultraprocesada y de alimentos que contienen alto contenido en gluten. Por otro lado, la EC se muestra con diferentes patrones sintomatológicos y las intervenciones médicas para la enfermedad comienza después de la aparición de dichos síntomas(48).

Actualmente, el único tratamiento disponible es la dieta sin gluten (DSG); sin embargo, no es efectiva en todos los pacientes. El manejo a lo largo de la vida de esta enfermedad representa una carga económica significativa para los sistemas de salud. Ante esta situación, los esfuerzos terapéuticos deberían enfocarse en la Medicina Predictiva, Preventiva y Personalizada (MPPP)(49). Una comprensión profunda de los mecanismos moleculares de la enfermedad es un pilar fundamental en este enfoque, y la bioinformática juega un papel clave en este proceso, ya que permite vincular hallazgos y patrones de cada caso con los de otros (38,39).

Las investigaciones mediante herramientas bioinformáticas han logrado integrar datos de estudios moleculares, genómicos y de conectividad biológica. Este enfoque ha revelado nuevos niveles de complejidad en los sistemas biológicos, lo que contribuye a una mejor comprensión de los mecanismos de las enfermedades, al desarrollo de métodos de diagnóstico más precisos y a la creación de estrategias terapéuticas y preventivas innovadoras(50).

## 3. Objetivos

### 3.1. Objetivo general

El objetivo principal de este trabajo es identificar nuevos biomarcadores para la EC mediante técnicas bioinformáticas.

### 3.2. Objetivos específicos

En cuanto a los objetivos específicos se propone:

- Seleccionar aquella información disponible en los últimos 10 años acerca de la búsqueda de nuevos biomarcadores para la EC mediante el uso de herramientas bioinformáticas.
- Conocer qué herramientas y técnicas bioinformáticas son las que se están usando actualmente para el estudio de la EC.
- Seleccionar artículos científicos que hayan comprobado de manera experimental los resultados obtenidos con las herramientas bioinformáticas.
- Comprobar si los nuevos biomarcadores pueden ser objetivo diana para otro tipo de tratamiento en la celiaquía complementarios a una DSG.
- Comprobar si existen biomarcadores potenciales para el diagnóstico de la EC para evitar técnicas invasivas como las biopsias del duodeno.

## 4. Materiales y métodos

### 4.1. Búsqueda inicial

Las primeras búsquedas se realizaron en octubre de 2024 combinando los términos ‘celiac disease’ y ‘coeliac disease’ en las bases de datos PubMed y Science Direct. También se realizó una búsqueda en la base de datos española Dialnet usando como términos clave ‘enfermedad celíaca’ y ‘celiaquía’. Posteriormente se amplió la búsqueda con una combinación usando los operadores booleanos AND y OR según conviniera, de los términos ‘bioinformatics’, ‘biomarkers’, ‘biological markers’, ‘nutrigenomic’, ‘genetic’, ‘HLA’, ‘HLA-DQ2’, ‘HLA-DQ8’, ‘treatment’, ‘diagnosis’, ‘prevention’. En el caso de la base de datos de Dialnet, se realizó la misma búsqueda, pero usando los términos en español. Estas búsquedas arrojaron una cantidad considerable de resultados, muchos repetidos entre las bases de datos y poco útiles para la revisión, pero nos dieron una visión amplia del tema de la EC.

### 4.2. Búsqueda sistemática

La búsqueda sistemática se resume en el diagrama de flujo siguiendo las declaraciones PRISMA (Figura 4). La búsqueda sistemática se realizó nuevamente en enero de 2025, en

Pubmed y ScienceDirect, acotando los resultados a las publicaciones realizadas en los últimos 10 años, es decir desde 2014, inclusive, hasta la actualidad.

La combinación de términos que se usaron para la búsqueda en Pubmed fue la siguiente: ("celiac disease" OR "coeliac disease") AND (bioinformatics) AND (biomarkers OR "biological biomarkers"). Los resultados obtenidos fueron 38 artículos.

Por otro lado, en la base de datos de ScienceDirect, la búsqueda con la misma combinación de términos utilizada en PubMed arrojó únicamente 2 artículos, los cuales ya estaban presentes en esta última base de datos. Debido a esta duplicación, la selección final de artículos se llevó a cabo exclusivamente a partir de PubMed.

Antes de proceder a la selección de artículos la revisión, se establecieron criterios de inclusión y exclusión.

Criterios de inclusión:

- Los artículos que se hayan publicado en los últimos 10 años, es decir, entre 2014 y principios de 2025 (exclusión ya tomada en cuenta previamente).
- Artículos de investigaciones empíricas en muestras humanas.
- Artículos que incluyan herramientas bioinformáticas para la detección de nuevos biomarcadores en la EC.
- Artículos que estudien únicamente los biomarcadores relacionados con la EC.

Criterios de exclusión:

- No incluir artículos que sean revisiones sistemáticas, libros, manuales o estudios de casos únicos.
- Artículos que no usen herramientas bioinformáticas para el procesamiento de los datos.
- Investigaciones en muestras no humanas.
- Estudios que incluyan otras enfermedades como la diabetes tipo 1 o la colitis ulcerosa.

De los 36 artículos que quedaron eliminando los duplicados, se eliminaron aquellos que realizaban los estudios en muestras no humanas, artículos que estaban en idioma diferente a inglés o español y revisiones sistemáticas. Con estos filtros, se eliminaron 5 artículos. Posteriormente, se descartaron artículos que investigaban otras enfermedades (por ejemplo,

la colitis ulcerosa, la diabetes...) o que no usasen técnicas bioinformáticas. Mediante este cribado se seleccionaron 21 artículos para su elegibilidad.

Mediante la lectura de los *Abstracts* o los artículos completos, sobre todo, haciendo hincapié en la metodología, por el uso de herramientas bioinformáticas, se seleccionaron 5 artículos.

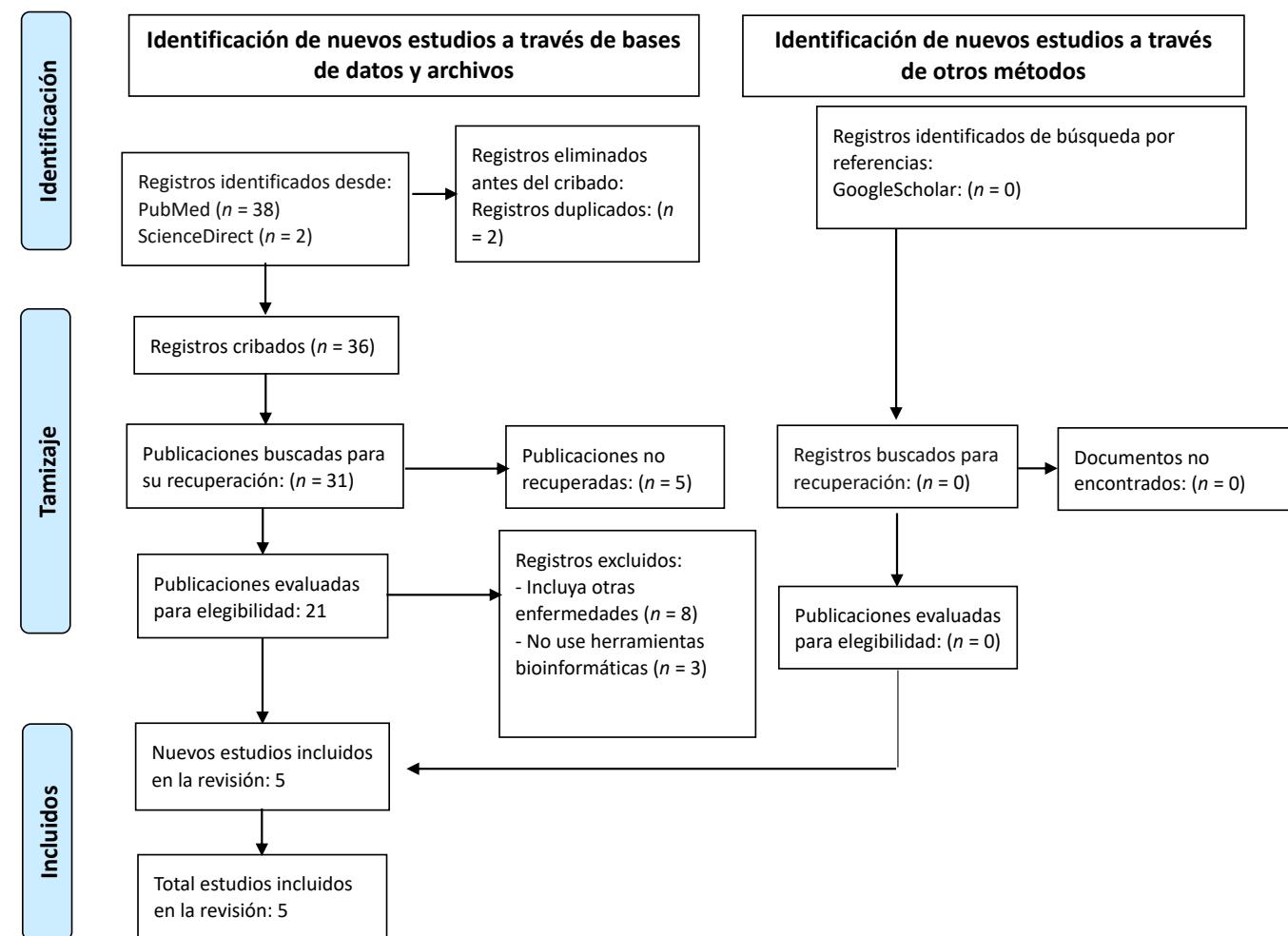


Figura 4: Diagrama de flujo PRISMA(51).

## 5. Resultados

### 5.1. Selección de estudios

Se identificaron un total de cinco estudios que cumplían con los criterios de inclusión y abordaban el uso de herramientas bioinformáticas para la identificación de biomarcadores en la enfermedad celíaca (EC). Estos estudios fueron seleccionados tras una revisión sistemática de la literatura y un análisis detallado de sus metodologías y hallazgos.

### 5.2. Características de los artículos

Las características de los estudios se ven reflejadas en la tabla 3. Los cinco estudios incluidos en esta revisión fueron publicados entre 2015 y 2025 y se enfocaron en el uso de herramientas bioinformáticas para la identificación de biomarcadores en EC. Estos estudios se llevaron a cabo en China, Irán e Italia, empleando metodologías basadas en el análisis de datos ómicos obtenidos de diferentes fuentes biológicas, como plasma, biopsias duodenales y bases de datos genéticas y proteómicas. El diseño de los estudios fue observacional, con enfoques transversales o retrospectivos. Cuatro de ellos analizaron muestras biológicas para identificar biomarcadores a través de técnicas de proteómica y transcriptómica, mientras que un estudio se basó exclusivamente en el análisis de bases de datos genómicas y proteómicas para detectar biomarcadores génicos relacionados con la EC.

En total, los estudios incluyeron 319 participantes, con tamaños muestrales que oscilaron entre 10 y 120 individuos por estudio. Uno de los estudios se realizó en población adulta e incluyó 45 pacientes con EC recién diagnosticados y 45 controles sanos, con edades promedio de  $48.2 \pm 10.8$  y  $47.9 \pm 15.8$  años, respectivamente. Otros tres estudios se centraron en población pediátrica con EC, con tamaños muestrales de entre 33 y 120 participantes, distribuidos en pacientes con diagnóstico confirmado de EC, pacientes en tratamiento con dieta sin gluten (DSG) y controles sanos. Un estudio adicional recopiló datos de bases de datos genéticas y proteómicas sin incluir directamente muestras biológicas de pacientes.

Las muestras biológicas utilizadas en los estudios incluyeron plasma y biopsias duodenales, mientras que el estudio basado en análisis bioinformático recurrió a bases de datos como CTD, GeneCards y DISEASES para la identificación de genes y proteínas implicadas en la EC. Tres estudios utilizaron plasma para la detección de proteínas y microARN circulantes como

biomarcadores potenciales, mientras que otros tres estudios analizaron biopsias duodenales para evaluar la expresión de microARN y proteínas en el tejido intestinal afectado.

Para el análisis de los datos, los estudios emplearon diversas técnicas y herramientas bioinformáticas. Se utilizaron métodos de análisis proteómico y redes de interacción proteína-proteína mediante herramientas como STRING y Cytoscape. Asimismo, se realizaron análisis de expresión diferencial de proteínas y genes a través de técnicas como WGCNA, KEGG y GO Pathways, con el objetivo de determinar la función biológica de los biomarcadores identificados. El análisis de microARN y transcriptómica se llevó a cabo mediante RNA-Seq, qPCR y Western Blot, utilizando software especializado como DESeq2, TargetScan, miRTarBase y GeneMANIA. Además, algunos estudios implementaron algoritmos de aprendizaje automático, como random forest, support vector machines (SVM) y regresión logística, para mejorar la precisión de la identificación de biomarcadores. Finalmente, se realizaron estudios de redes génicas, análisis de enriquecimiento de vías metabólicas e interacciones genéticas con herramientas como Enrich, WebGestalt y GEO2R.

Los estudios incluidos tenían como objetivo principal la identificación de biomarcadores para el diagnóstico y seguimiento de la EC. Dos estudios evaluaron proteínas plasmáticas y tisulares con la finalidad de desarrollar pruebas diagnósticas no invasivas, mientras que tres estudios se enfocaron en la identificación de microARN circulantes y tisulares como biomarcadores de la EC. Además, un estudio utilizó bases de datos bioinformáticas para analizar la expresión génica y los genes clave involucrados en la enfermedad.

Los criterios de inclusión de los estudios contemplaron pacientes con diagnóstico confirmado de EC mediante biopsia duodenal y pruebas serológicas, así como controles sanos sin evidencia de la enfermedad. En los estudios pediátricos, se consideraron tanto pacientes con EC activa como aquellos en tratamiento con DSG. Entre los criterios de exclusión se incluyó la presencia de enfermedades autoinmunes concurrentes distintas a la EC, el uso previo de una dieta sin gluten en estudios que requerían evaluar la expresión de biomarcadores en pacientes con EC activa, y la calidad insuficiente de las muestras biológicas o datos genéticos incompletos en estudios de bioinformática.

**Tabla 3: Características de los estudios seleccionados.**

Autor/a/es	País	Objetivo	Participantes	Muestra	Metodología
(47)	China	Evaluar la fisiopatología de la EC e identificar posibles marcadores circulantes para su diagnóstico mediante el análisis bioinformático y el aprendizaje automático.	45 pacientes recién diagnosticados con EC ( $48.2 \pm 10.8$ años) y 45 pacientes como controles sanos ( $47.9 \pm 15.8$ ).	Plasma y duodeno	<ol style="list-style-type: none"> <li>1. Recolección de muestras y población del estudio en la fase de descubrimiento.</li> <li>2. Análisis mediante cromatografía líquida, espectrometría de masas y bioinformática.</li> <li>3. Identificación de posibles biomarcadores diagnósticos en plasma mediante WGCNA y aprendizaje automático.</li> <li>4. Uso del método Mfuzz y aprendizaje automático para explorar biomarcadores de proteínas plasmáticas en la atrofia de vellosidades intestinales en CeD.</li> <li>5. Recolección de muestras plasmáticas para la validación mediante ELISA.</li> </ol>
(52)	Irán	Detección de nuevos biomarcadores génicos relacionados con la EC mediante el uso de herramientas bioinformáticas para el procesamiento de datos procedentes de diferentes bases de datos.	Pacientes que se han sometido a una endoscopia del tracto gastrointestinal superior con frecuencia y que se les ha realizado una biopsia del duodeno.	Datos genéticos y proteómicos de 3 bases de datos diferentes (CTD, GeneCards y DISEASE).	<ol style="list-style-type: none"> <li>1. Obtención de los datos genéticos procedentes de las bases de datos CTD, GeneCards y DISEASE.</li> <li>2. Búsqueda de los genes compartidos en las 3 bases de datos mediante el software del Diagrama de Venn.</li> <li>3. Análisis funcional mediante la base de datos Enrich.</li> <li>4. Análisis topológico de redes: se usó la base de datos String para construir las redes; las interacciones proteína-proteína (PPI) se construyó mediante el software Cytoscape; análisis de enriquecimiento mediante WebGestalt.</li> <li>5. Análisis funcional de las redes: mediante los genes obtenidos mediante el software de Cytoscape se realizó un análisis funcional mediante la herramienta GeneMania.</li> <li>6. Análisis de la expresión génica mediante el uso de la base de datos GEO mediante la búsqueda con las palabras clave: <i>Celiac Disease + Duodenum</i>.</li> <li>7. Búsqueda de los genes compartidos entre las 3 bases de datos (CTD, DISEASE y GeneCards) + los genes obtenidos en la base de datos GEO, mediante el software del Diagrama de Venn.</li> <li>8. Análisis estadístico y análisis de enriquecimiento en Enrich (Omim, Gene Ontology y vías metabólicas).</li> <li>9. Construcción y análisis de las redes usando el software Cytoscape.</li> <li>10. Análisis funcional mediante la base de datos GeneMania.</li> </ol>

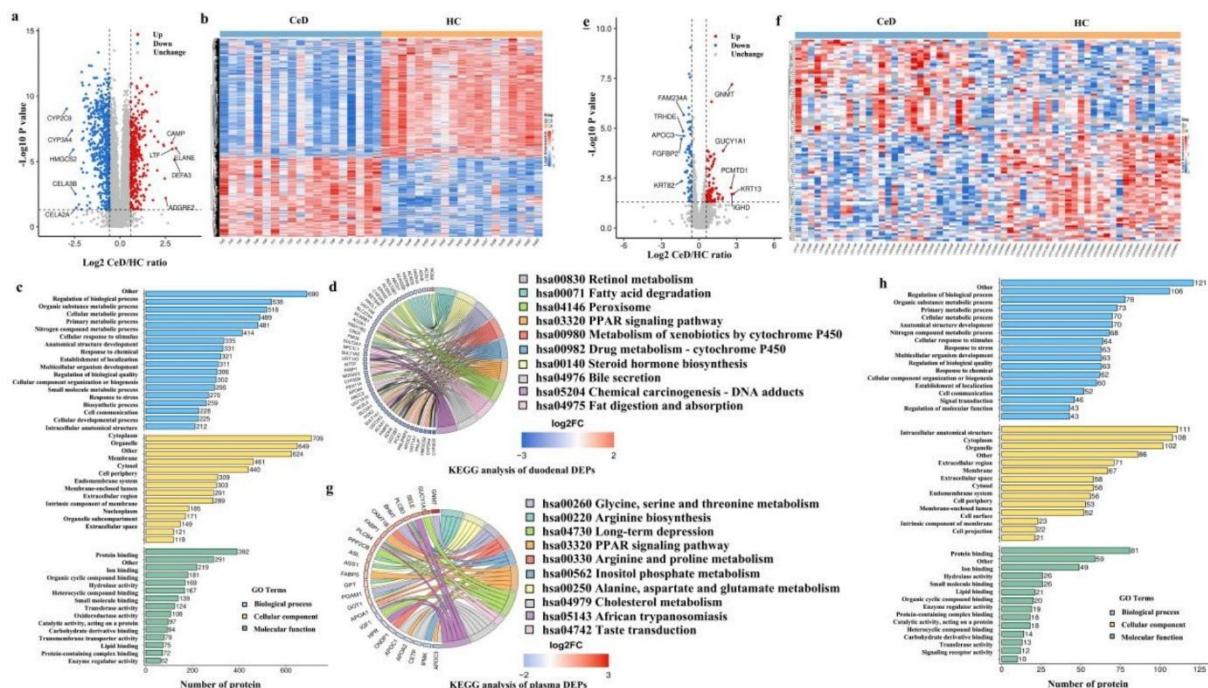
(53)	Italia	Detección de microRNA circulantes como biomarcadores no invasivos para el diagnóstico de la EC pediátrica.	120 participantes entre 3 y 15 años pertenecientes a 3 grupos: - Pacientes con EC (n=40) - Pacientes con EC siguiendo una DSG (n=40) - Grupo control sano (n=40)	Suero	1. Recolección de muestras de sangre de pacientes pediátricos de los 3 grupos (Control, EC, EC en tratamiento con DSG). 2. Aislamiento de RNA. 3. Descubrimiento de microRNA circulantes mediante secuenciación de RNA (Illumina). 4. Análisis bioinformático de RNA-Seq. 5. Validación de miRNAs desregulados por qPCR. 6. Análisis estadístico.
(54)	Italia	Detección de microRNA circulantes y en tejido de duodeno y sus dianas genéticas con la finalidad de descubrir nuevos biomarcadores para la EC.	1) 28 pacientes pediátricos sometidos a endoscopia gastrointestinal (8 controles y 20 pacientes con EC no tratada).  2) 36 pacientes pediátricos (12 controles, 17 pacientes con EC activa y 7 pacientes con EC en tratamiento con DSG.	Biopsias duodeno y plasma	1. Recolección de muestras de biopsias de duodeno y muestras sanguíneas. 2. Extracción del ARN de las biopsias y plasma sanguíneo. 3. PCR cuantitativa en tiempo real. 4. Western Blot para la extracción de proteínas. 5. Análisis estadístico.
(55)	Italia	Explorar el papel de la desregulación del transcriptoma mediada por miRNA en la EC.	Biopsias duodenales de pacientes pediátricos con EC (n=33) y sujetos control (n=10) para la secuenciación de miRNA. Algunos pacientes con EC (n=6) y controles (n=5) estaban disponibles para la secuenciación de ARN.	Biopsias duodeno	1. Recolección de muestras de biopsias duodenales. 2. Aislamiento de ARN, miRNA y secuenciación de ARNm. 3. Análisis de expresión diferencial de la secuenciación de miRNA y del ARN. 4. Construcción de la red de transcripción diana de miRNA. 5. Análisis de enriquecimiento. 6. Análisis de enriquecimiento de tipos celulares.

### 5.3. Resultados de los artículos seleccionados

Los resultados obtenidos en los estudios analizados se reflejan en la tabla 4. Estos estudios se centran en la identificación de biomarcadores en tejido duodenal y plasma/suero sanguíneo con el objetivo de evaluar su potencial como predictores de la enfermedad celíaca (EC). Para ello, se emplearon diversas herramientas bioinformáticas que permitieron analizar datos ómicos, patrones de expresión génica y redes de interacción molecular, facilitando así una mejor comprensión de los mecanismos subyacentes a la enfermedad.

#### 5.3.1. Proteómica

El análisis de Li y colaboradores (2024) se centra en la identificación de biomarcadores diagnósticos no invasivos en pacientes con enfermedad celíaca (EC), mediante un análisis proteómico de plasma y biopsia de duodeno. Para este estudio, se utiliza la proteómica de adquisición dependiente de datos en cuatro dimensiones (4D-DIA), una técnica avanzada de espectrometría de masas que combina varios parámetros para identificar y cuantificar proteínas. Los resultados mostraron proteínas diferencialmente expresadas (PED) reguladas al alza y a la baja en muestras de duodeno y plasma (Figura 5).

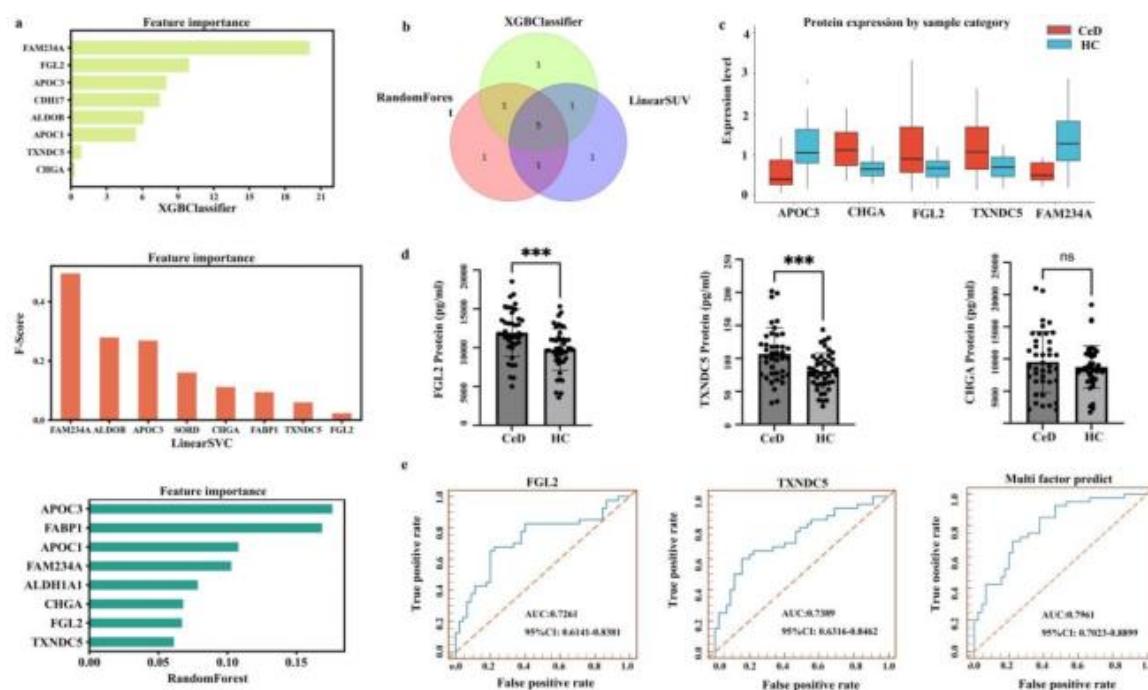


**Figura 5: Análisis proteómico de muestras duodenales (a, b, c, d) y de muestras plasmáticas (e, f, g, h) en pacientes con EC. a) Gráfico de volcán de las PED. b) Mapa de calor de las PED. c) Histograma del análisis de GO. d) Análisis de enriquecimiento de KEGG de las PED. e)**

*Gráfico volcán de las PED. f) Mapa de calor de las PED. g) Histograma del análisis de GO. h) Análisis de enriquecimiento de KEGG de las PED(47).*

En las muestras de duodeno, el análisis de ontología genética (GO) reveló que las PED estaban principalmente involucradas en la regulación de procesos biológicos (Figura 5c). Además, el análisis KEGG indicó que las PED estaban asociadas con importantes vías metabólicas, como la secreción biliar, metabolismo de fármacos (citocromo P450) y señalización del receptor activado por peroxisomas (PPAR) (Figura 5d). Estos resultados sugieren alteraciones metabólicas relevantes en la EC(47).

En el análisis de plasma, las PED estuvieron asociadas con la regulación de procesos biológicos, metabolismo de compuestos nitrogenados y otras vías metabólicas (Figura 5g, 5h). Se utilizaron tres algoritmos de aprendizaje automático (XGBClassifier, LinearSVC y RandomForest) para seleccionar 17 proteínas clave. Posteriormente, se validaron tres de ellas (FGL2, TXNDC5 y CHGA) mediante ELISA, observándose una mayor expresión de FGL2 y TXNDC5 en los pacientes con EC (Figura 6d). Los análisis de curvas ROC indicaron que FGL2 y TXNDC5 tienen un gran potencial como biomarcadores diagnósticos para EC (Figura 6e)(47).



*Figura 6: Tres métodos de aprendizaje automático o machine learning para seleccionar moléculas características. a) Los puntajes de contribución significativos de las características identificadas por XGBClassifier, LinearSVC y RandomForest. b) Diagrama de Venn de los tres algoritmos de aprendizaje automático. c) Diferencia en la distribución de las cuatro*

características entre las diferentes categorías de muestras. d) Histograma de los resultados de ELISA para las proteínas FGL2, TXNDC5 y CHGA. e) Curva ROC de las proteínas FGL2, TXNDC5 y CHGA(47).

Además, se realizó un análisis de proteínas plasmáticas para identificar patrones de expresión relacionados con los diferentes grados de atrofia de las vellosidades del intestino delgado, utilizando el método Mfuzz. De las 1.682 proteínas analizadas, se seleccionaron 460 con expresión significativa, agrupándolas en cuatro grupos. El grupo 1 mostró una asociación con procesos biológicos relevantes, como el cáncer colorrectal y la regulación biosintética (Figura 7). Se identificaron nueve DEP relacionadas con el grado Marsh, y mediante aprendizaje automático se seleccionaron cuatro proteínas clave (FABP5, CPOX, BHMT, PPP2CB) que mostraron un buen potencial para discriminar entre los grados de atrofia intestinal (Figura 8)(47).

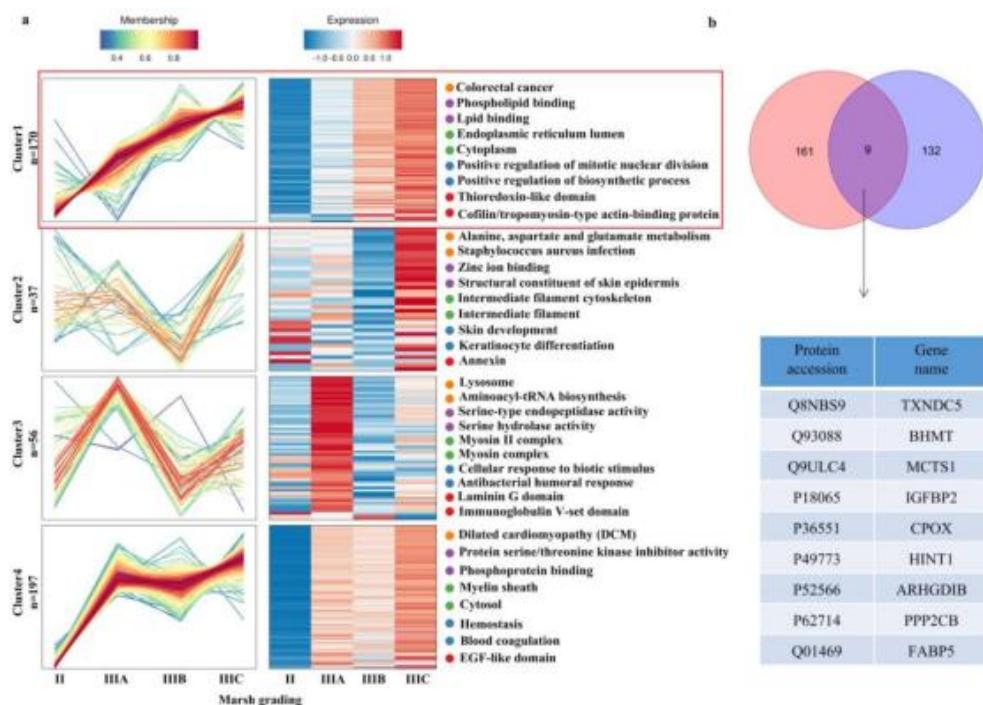
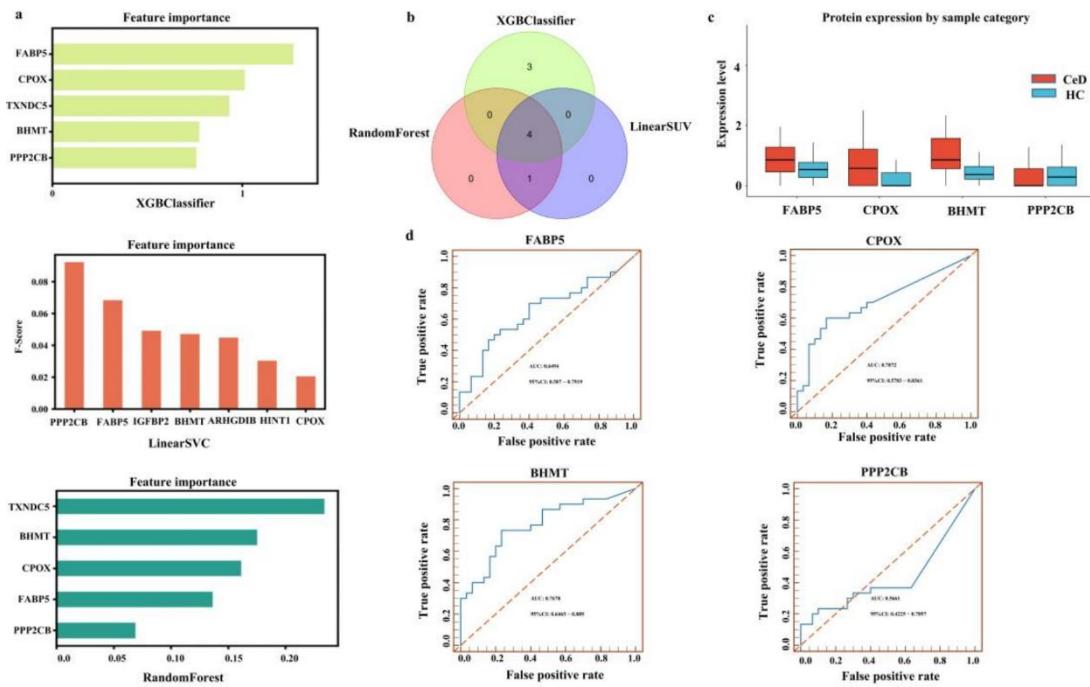


Figura 7: Agrupamiento de proteínas plasmáticas y análisis funcional. a) Gráfico resumen del análisis de agrupamiento de patrones de expresión. b) Diagrama de Venn del grupo 1 y las proteínas diferencialmente expresadas(47).



*Figura 8: Tres métodos de aprendizaje automático o machine learning para seleccionar moléculas características. a) Los puntajes de contribución significativos de las características identificadas por XGBClassifier, LinearSVC y RandomForest. b) Diagrama de Venn de los tres algoritmos de aprendizaje automático. c) Diferencia en la distribución de las cuatro características entre las diferentes categorías de muestras. d) Curva ROC de las proteínas FABP5, CPOX, BHMT y PPP2CB(47).*

En el estudio de las ómicas en la EC también encontramos la investigación realizada por Noori y colaboradores (2020). En este estudio, se utilizaron bases de datos como Comparative Toxicogenomics Database (CTD), DISEASES y GeneCards, las cuales recuperaron un total de 2656 genes relacionados con la EC. De estos, 54 genes fueron identificados como comunes en las tres bases de datos mediante un diagrama de Venn (Figura 9). Posteriormente, se realizaron análisis funcionales utilizando herramientas como Enrich, Cytoscape (Figura 10), GeneMANIA (Figura 11) y WebGestalt (Figura 12). Los resultados mostraron que los genes asociados a la EC están principalmente involucrados en el sistema inmunológico y las vías de señalización relacionadas con enfermedades autoinmunes(52).

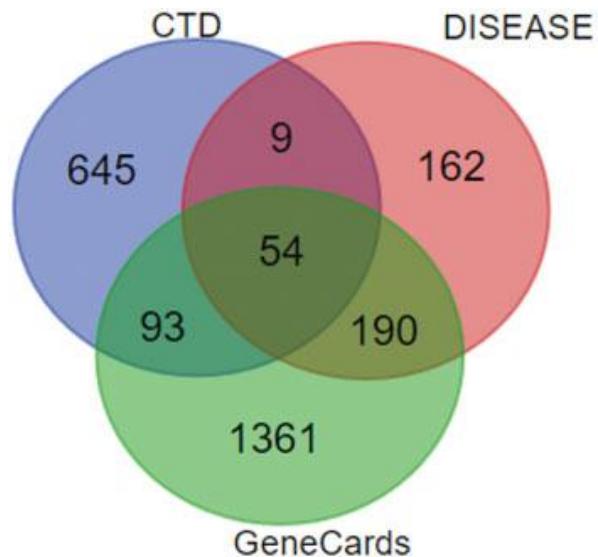


Figura 9: Diagrama de Venn de la intersección entre todos los genes de la EC obtenidos de 3 bases de datos pública(52).

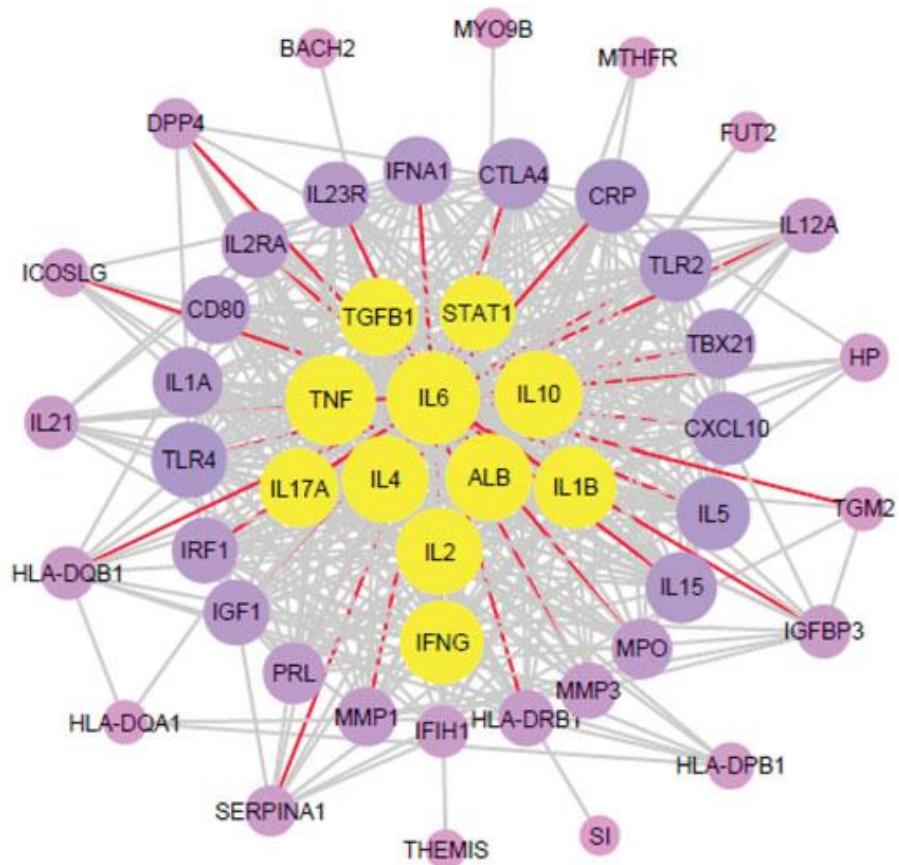
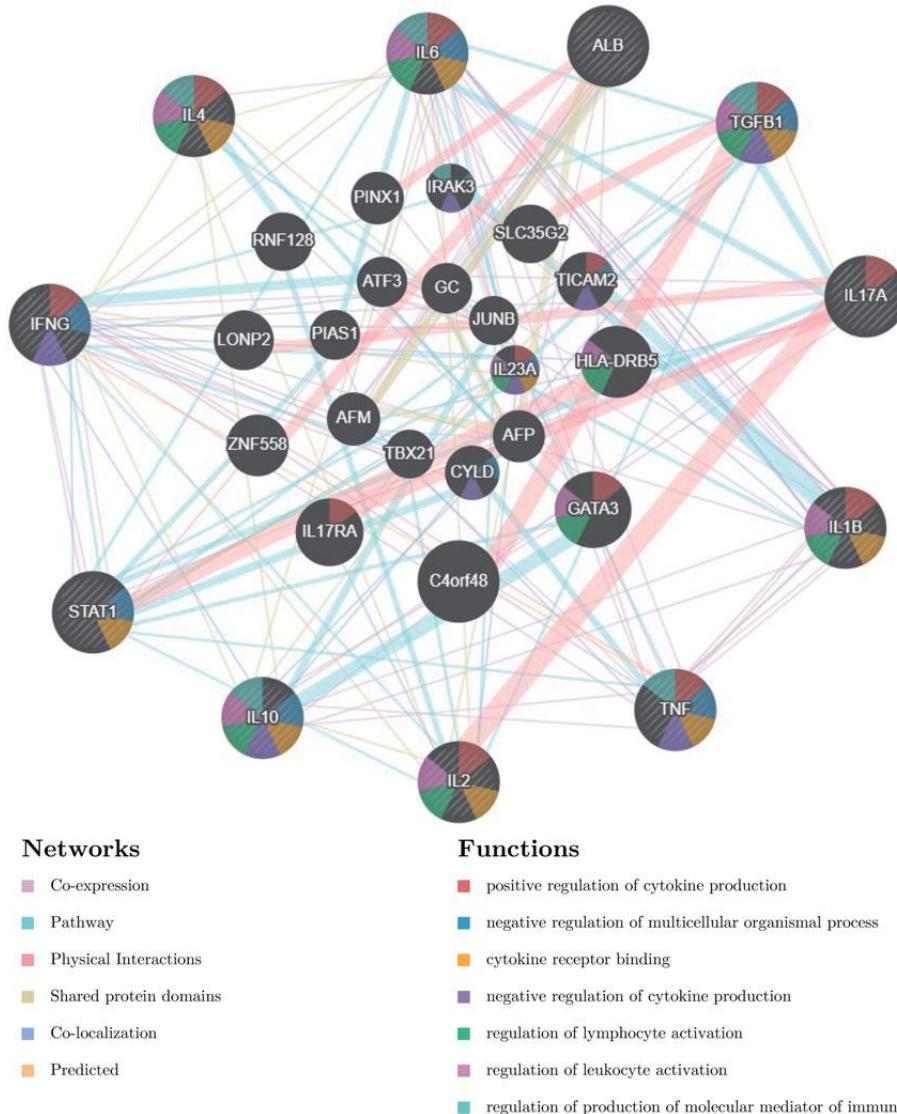


Figura 10: La red relacionada con los 54 genes de la EC se construyó utilizando el software Cytoscape. El tamaño de los nodos está determinado por su grado, y los genes clave (hub genes) se muestran en el centro. Los bordes más oscuros indican, como ejemplo, cómo el nodo central (IL6) interactúa con todos los nodos asociados(52).



*Figura 11: Análisis funcional de la red de interacción establecida por los 11 genes principales obtenidos de GeneMANIA(52).*

Entre los genes más relevantes, se identificaron STAT1, ALB, IL10, IL2, IL4, IL17A, TGFB1, IL1B, IL6, TNF e IFNG, los cuales tienen un papel clave en la regulación del sistema inmunológico y en la patogenia de la EC. Además, se destacó que 25 de los 54 genes fueron clasificados como "genes semilla" en la base de datos WebGestalt (Figura 12)(52).

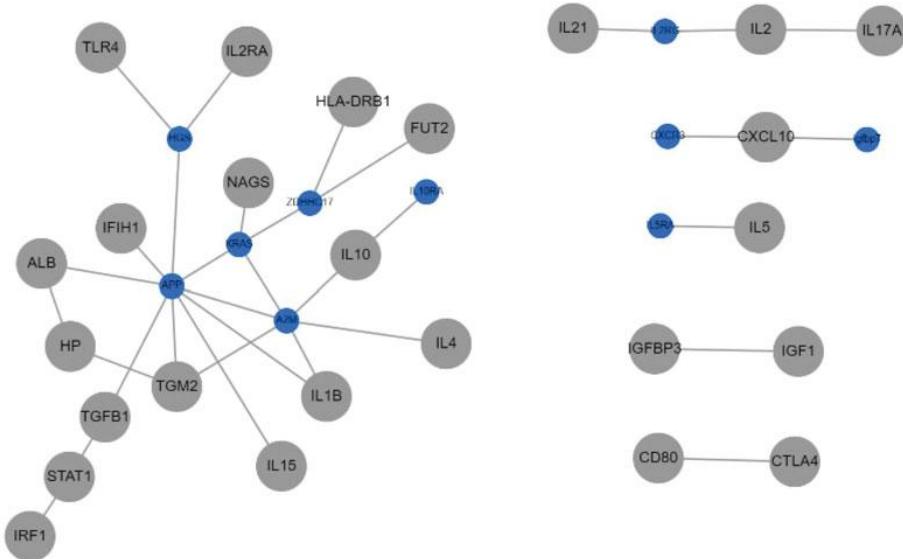


Figura 12: Interacción proteína-proteína de los 54 genes en WebGestalt. Se determinó la subred de genes semilla (nodos grandes) conectados a los primeros nodos conectados (nodos pequeños)(52).

Se realizó un análisis adicional de expresión diferencial mediante la herramienta GEO2R en diversas series de datos del Gene Expression Omnibus (GEO). Los resultados mostraron que, en comparación con los controles, en los pacientes celíacos había genes significativamente expresados en varias series de datos, como GSE76168, GSE87460, GSE87458, GSE87457, GSE112102 y GSE102991. Entre los genes más destacados, IRF1 y STAT1 fueron comunes entre los genes significativos de GEO y los 54 genes relacionados con la EC (Figura 13)(52).

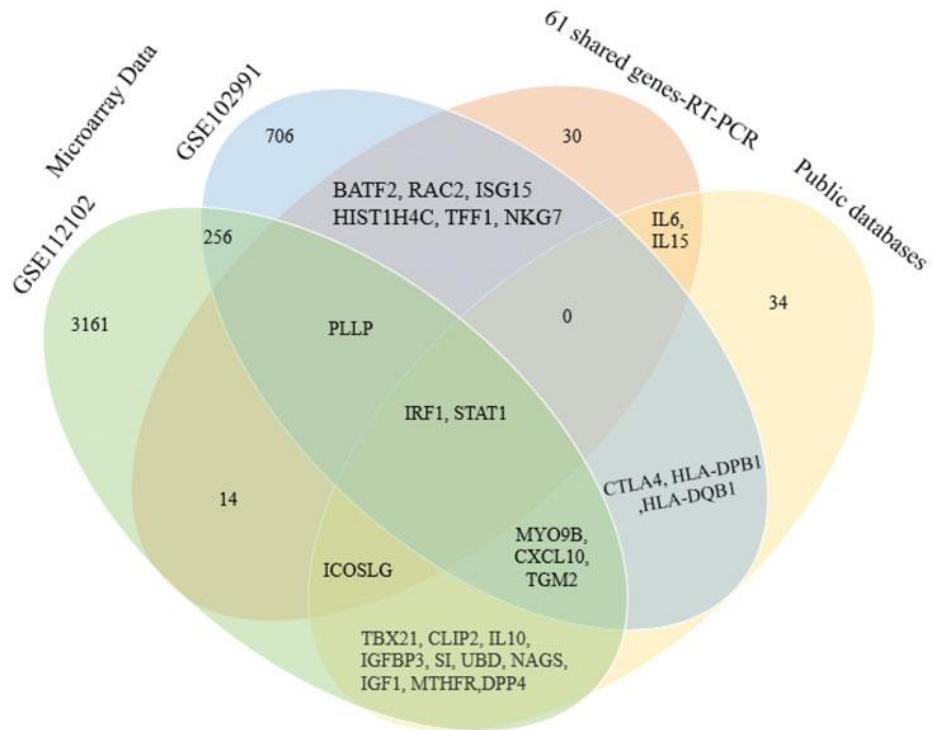


Figura 13: Diagrama de Venn de la intersección entre los genes comunes de la EC en 3 bases de datos públicas y los genes en los conjuntos de datos de GEO (datos de microarreglo y RT-PCR)(52).

A partir de estos hallazgos, se identificaron nueve genes clave (IRF1, STAT1, IL17A, TGFB1, ALB, IL10, IL2, IL4 e IL1B) que están fuertemente asociados con la EC. Estos genes podrían servir como biomarcadores para el diagnóstico, el entendimiento de la patología y el desarrollo de tratamientos más efectivos para la enfermedad celíaca(52).

### 5.3.2. microRNA en la EC

En la búsqueda de nuevos biomarcadores diagnósticos para la EC, encontramos la identificación de microRNA o miRNA circulantes. Se trata de una nueva tendencia en el estudio de enfermedades autoinmunes, con el objetivo de descubrir novedosos biomarcadores. Los miRNA son pequeños ARN no codificantes (18-22 nucleótidos) que regulan la expresión génica a nivel postranscripcional y están implicados en diversos procesos biológicos como el desarrollo, la proliferación y la apoptosis(56–58).

En el estudio de la enfermedad celíaca, se han centrado tanto en los miRNA del tejido del duodeno como en los miRNA circulantes, que son aquellos que se encuentran fuera de las células. Los miRNA circulantes son muy estables, ya que son resistentes a la digestión por ARNasas, condiciones de pH extremos, altas temperaturas, al almacenamiento prolongado y

múltiples ciclos de congelación y descongelación. Por lo tanto, el estudio de miRNA circulantes procedentes de fluidos biológicos como el suero, plasma, saliva u orina resulta muy interesante para emplearse como nuevos biomarcadores no invasivos(56–58).

Los miRNA se expresan de manera diferencial en la EC, donde se han identificado miRNA expresados tanto de manera positiva o negativa en pacientes con EC activa. Su estudio en los últimos años se basa en una combinación de análisis bioinformáticos y datos experimentales. Por lo tanto, los miRNA pueden convertirse en unos biomarcadores importantes para la EC, así como usarse como predictores en pacientes en tratamiento con DSG.

En el estudio de los miRNA, cabe mencionar la investigación de Buoli Comani y colaboradores (2015). El estudio utilizó una combinación de técnicas experimentales y herramientas bioinformáticas para analizar los miRNA en biopsias de duodeno y plasma sanguíneo de pacientes con EC. Utilizaron herramientas como mirTarBase, TargetScan, microRNA.org y Mirò para predecir las dianas de los miRNA y el software Quantity One 4.1.1 para analizar la intensidad de las señales en los experimentos de Western blot. Aunque los resultados en biopsias de duodeno mostraron miRNA desregulados en pacientes con EC, los miRNA circulantes en pacientes tratados con DSG no mostraron resultados significativos. Además, cabe mencionar que los resultados se compararon con otro estudio realizado en adultos, y parecer ser que la desregulación de NOD2 y CXCL2 por la disminución de expresión de miR-192-5p no se producía en los pacientes pediátricos. Por lo tanto, la expresión de los miRNA en pacientes EC difiere según la edad del paciente. Sin embargo, hacen hincapié en que los resultados son prometedores y que se debería analizar un panel más amplio de miRNA plasmáticos para proporcionar información confiable en el diagnóstico de la EC con el objetivo de reducir la necesidad del uso de procedimientos invasivos(54).

El estudio de Felli y colaboradores (2022) también hace un uso extensivo de herramientas bioinformáticas y experimentales. Para el descubrimiento de los miRNA circulantes, los investigadores secuenciaron el ARN y realizaron un análisis bioinformático del RNA-Seq utilizando herramientas como el paquete FastQC y la plataforma Galaxy para asegurar la calidad de las secuencias. Se eliminaron los adaptadores y se recortaron las bases de baja calidad utilizando TrimGalore. Las lecturas se alinearon con el genoma de referencia hg19 usando Bowtie2, y la cantidad de lecturas se contabilizó con HTSeq-count. Los miRNA diferencialmente expresados fueron identificados utilizando el paquete DESeq2, con

corrección de efectos de lote aplicada mediante el método COMBAT. Para representar estos resultados, se utilizó un análisis de componentes principales (PCA) que reveló que el grupo de pacientes con EC tratada con DSG se superponía en su perfil de miRNA con los controles sanos, mientras que los pacientes con EC activa mostraron una clara diferencia. Este análisis permitió identificar 13 miRNA que se expresaban diferencialmente, de los cuales tres (miR-192-5p, miR-215-5p y miR-125b-5p) pudieron discriminar entre los tres grupos de estudio (Figura 14a)(53).

Para validar los resultados obtenidos mediante RNA-Seq, los investigadores también utilizaron la técnica de qPCR. Se identificaron 13 miRNA que fueron obtenidos tanto por RNA-Seq como por qPCR, de los cuales 8 mostraron sobreexpresión y 5 subexpresión en pacientes con EC en comparación con los controles, así como en pacientes con EC en tratamiento con dieta sin gluten (DSG) frente a los controles. Los resultados mostraron una clara separación en el mapa de calor de los miRNA expresados (Figura 14b), donde los miRNA sobreexpresados se ubicaban en la parte superior y los regulados a la baja en la parte inferior. Este patrón de expresión fue claramente distinto en los pacientes con EC en comparación con los controles y los pacientes tratados con DSG(53).

De los 13 miRNA analizados, tres (miR-192-5p, miR-215-5p y miR-125b-5p) fueron capaces de discriminar entre los tres grupos de estudio: control, EC activa y EC tratada con DSG. Además, un algoritmo de modelo lineal generalizado (GLMNET) confirmó que estos miRNA tienen el potencial de ser utilizados para discriminar entre pacientes con EC y los grupos control. Los miRNA identificados están involucrados en procesos biológicos clave como la apoptosis, las transiciones del ciclo celular y la respuesta inmunológica(53).

Este estudio también validó que la DSG puede restaurar los niveles de miRNA desregulados en pacientes con EC, lo que subraya la importancia de los miRNA circulantes como posibles biomarcadores de la EC, además de proporcionar un enfoque robusto utilizando tanto análisis bioinformáticos como experimentales para su validación(53).

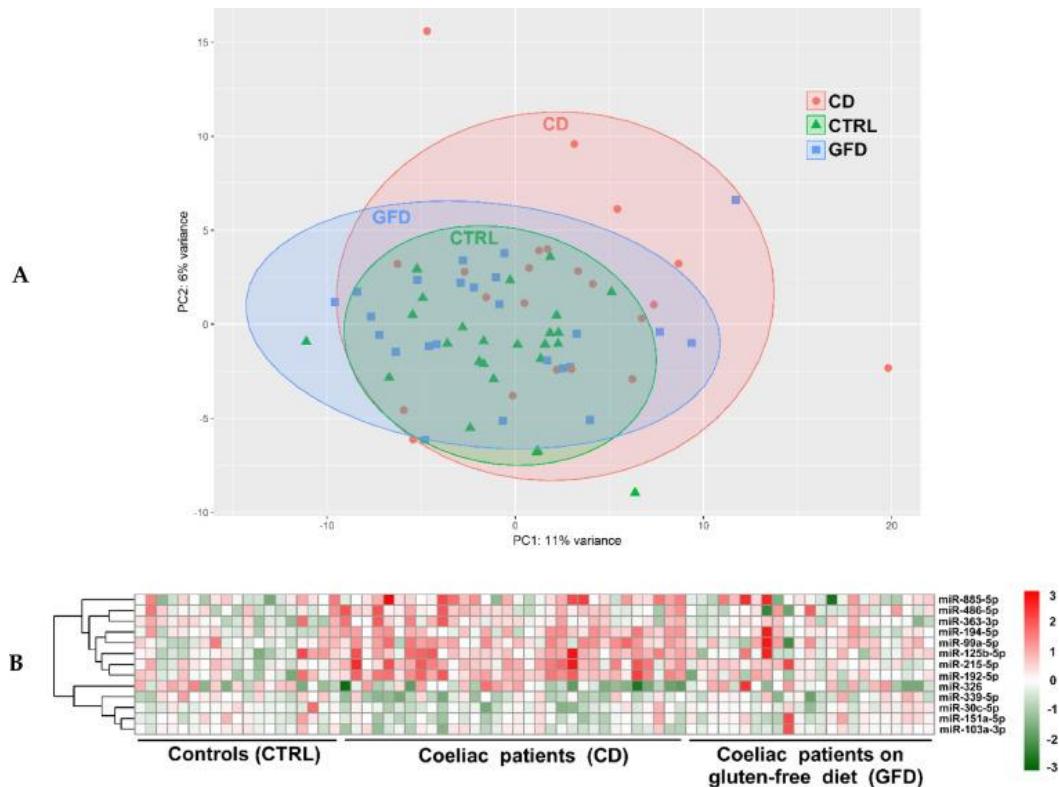
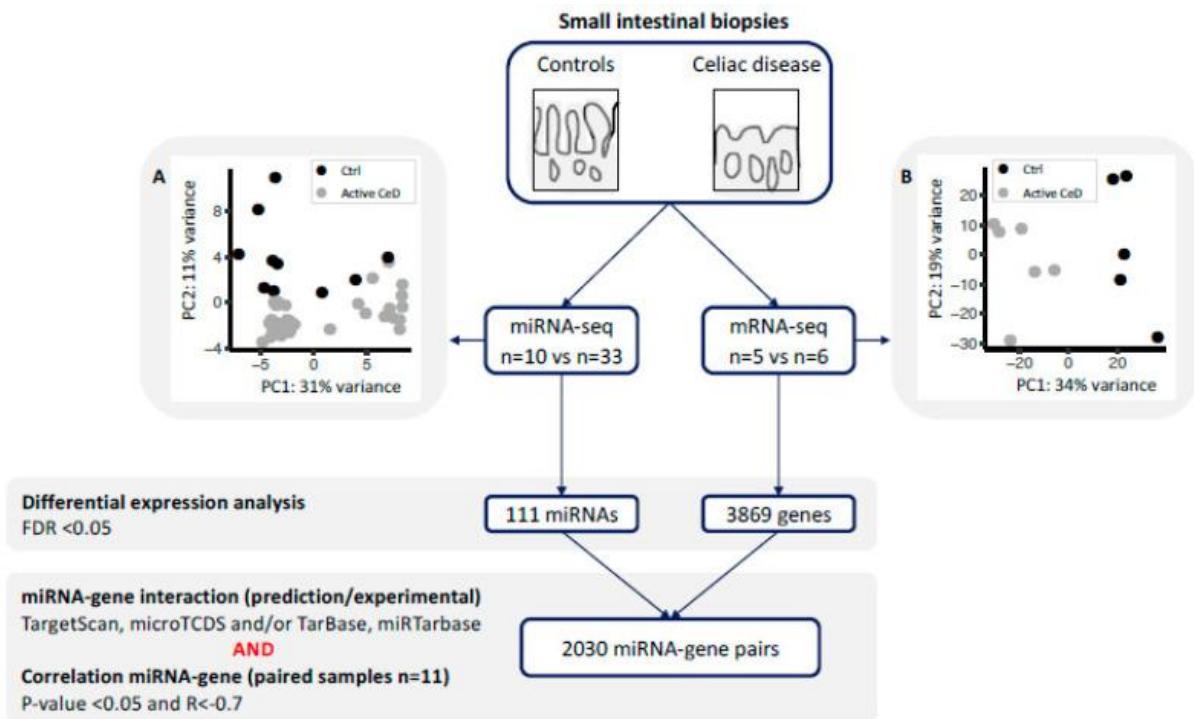


Figura 14: Análisis PCA y mapa de calor. a) Análisis de componentes principales (PCA) de los datos de RNA-Seq para los grupos CTRL (control), CD (pacientes con enfermedad celíaca) y GFD (pacientes en tratamiento con dieta libre de gluten). Las regiones coloradas representan la distribución de los datos para cada categoría, mientras que el cada símbolo (cuadrado, rectángulo y círculo) representa un paciente. b) Mapa de calor de los miRNA circulantes significativos que son capaces de diferenciar los tres grupos analizados (CTRL, CD y GFD) obtenidos por qPCR (53).

Luise Tan y colaboradores (2021) realizaron un estudio sobre la regulación postranscripcional de genes en la enfermedad celíaca (EC) mediante el análisis de miRNA. Utilizando herramientas bioinformáticas, identificaron miRNA diferencialmente expresados y construyeron una red de interacción miRNA-diana. Este análisis destacó el papel de los miRNA en vías inmunitarias y metabólicas, y su contribución al desequilibrio del metabolismo lipídico en el intestino delgado, lo que agrega complejidad a la desregulación del transcriptoma en la EC.

Se realizó un análisis de componentes principales (PCA) para correlacionar el perfil de miRNA en pacientes con EC con los controles sanos. Los resultados mostraron una clara separación significativa entre ambos grupos (Figura 15a). El mismo análisis se aplicó a los perfiles de

expresión de ARNm en biopsias de pacientes con EC y controles, revelando una expresión diferencial entre los grupos (Figura 15b)(55).



*Figura 15: Análisis de expresión de miRNA y genes en la enfermedad celíaca (EC). Se generaron perfiles de expresión a partir de biopsias intestinales mediante secuenciación de nueva generación. El análisis de componentes principales (A) para miRNA-seq y (B) para mRNA-seq muestra una clara separación entre controles (negro) y pacientes con EC activa (gris). Se identificaron 2030 pares miRNA–transcripción con correlación negativa significativa ( $R < -0,7$ ), resaltando su potencial papel en la regulación de genes clave en EC(55).*

En el estudio, se construyó una red de interacción miRNA-diana mediante herramientas bioinformáticas, integrando datos de predicción (TargetScan, microTCDS) y validación experimental (TarBase, miRTarbase). Se identificaron 2030 pares miRNA-diana con correlación negativa significativa, destacando la regulación coordinada de genes clave en la EC y el papel de familias de miRNA en su desregulación transcriptómica (Figura 16)(55).

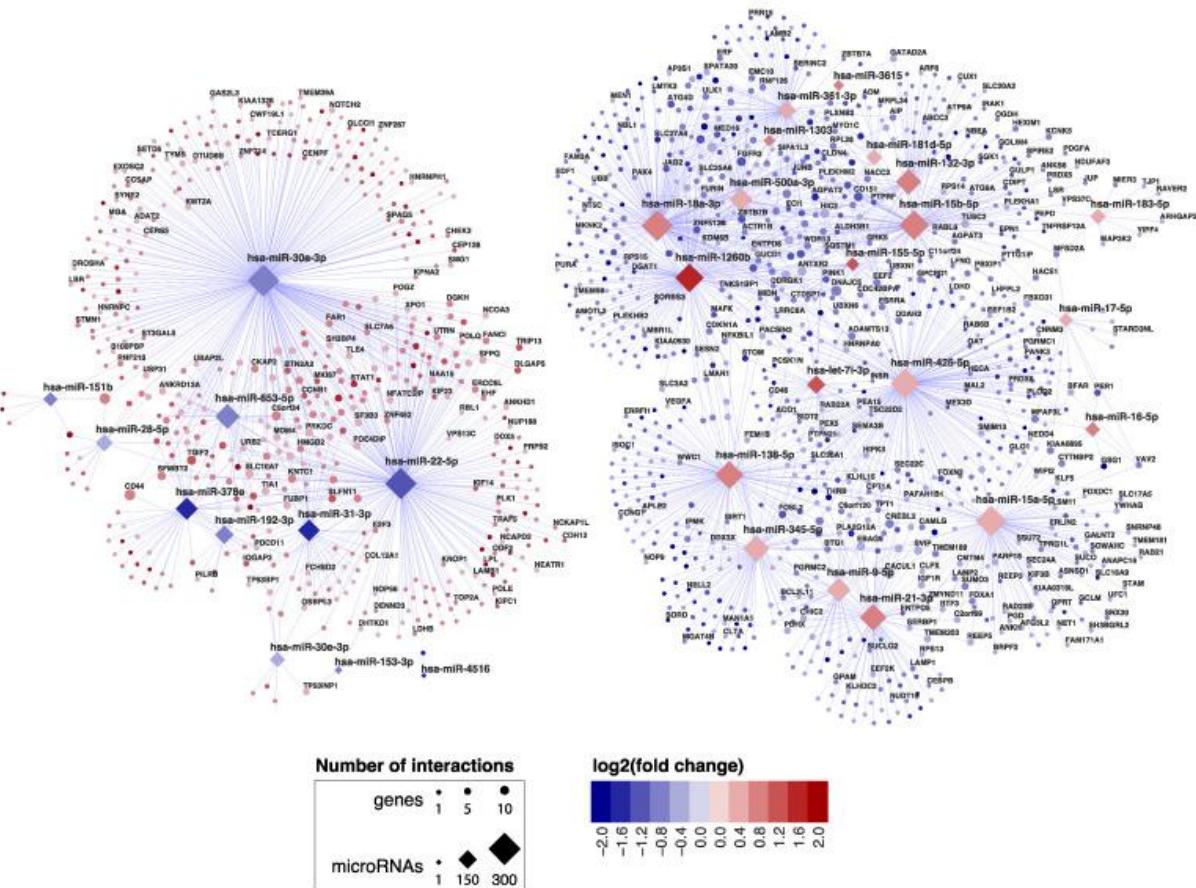
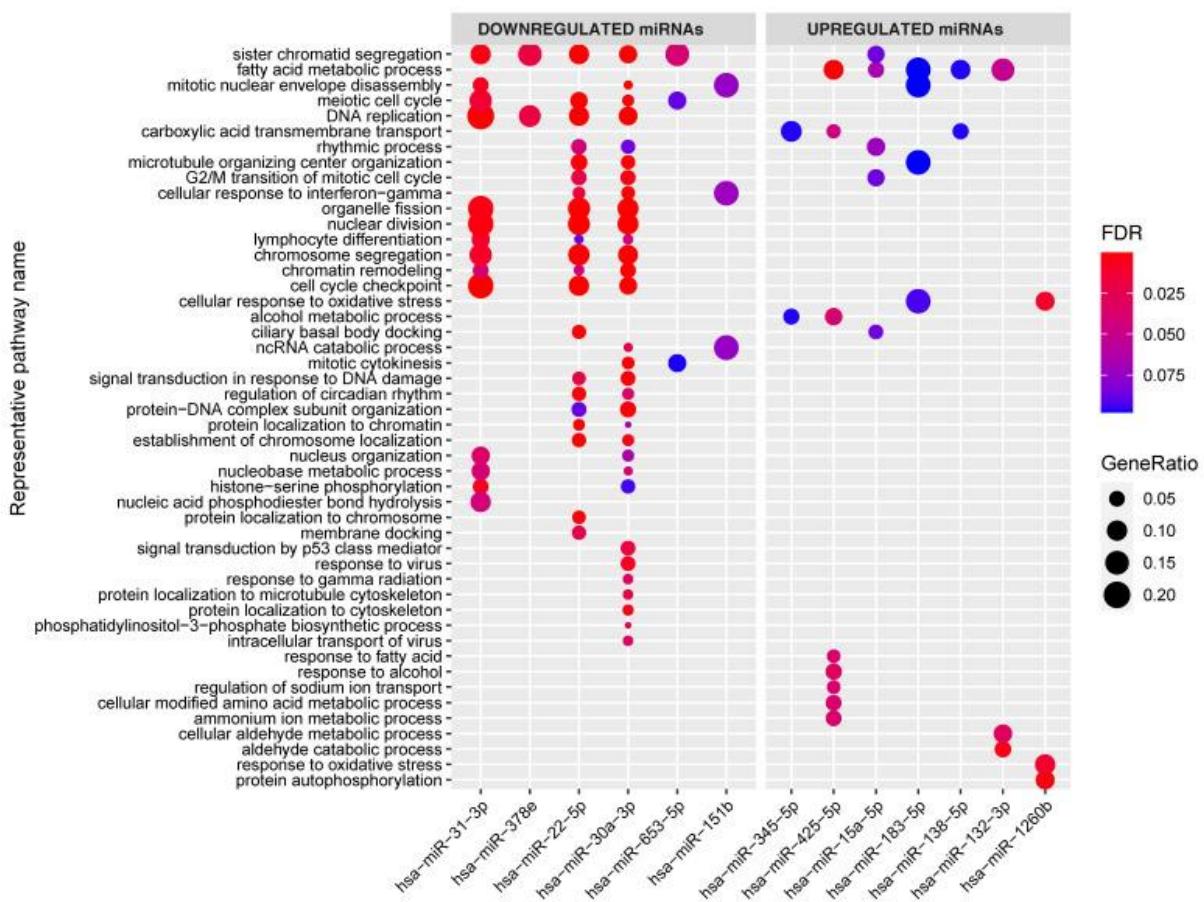


Figura 16: Red de interacción miRNA-diana en la enfermedad celíaca (EC). Se muestran las interacciones entre miRNA diferencialmente expresados y sus genes diana con correlación negativa significativa ( $R < -0,7$ ,  $p < 0,05$ ). Los miRNA (diamantes) y genes (círculos) están coloreados según su  $\log_2(\text{fold change})$ , con tonos rojos indicando sobreexpresión y azules subexpresión en pacientes con EC. El tamaño de los nodos representa el número de interacciones, destacando la regulación coordinada de genes clave en procesos inmunológicos y metabólicos(55).

El análisis de enriquecimiento de vías en este estudio se centró en identificar las vías biológicas asociadas con las transcripciones diana de miRNA, utilizando herramientas bioinformáticas como el paquete R *clusterProfiler* y la plataforma REVIGO para reducir la redundancia de términos GO. Se descubrió que las transcripciones reguladas a la baja estaban relacionadas con el metabolismo lipídico, mientras que las reguladas al alza se asociaban con vías inmunológicas y del ciclo celular. Además, se priorizaron las transcripciones diana con apoyo experimental, destacando que las transcripciones de miRNA como miR-31-3p, miR-22-5p y

miR-30a-3p afectan vías similares, como la diferenciación de linfocitos (Figura 17). Este enfoque bioinformático permitió identificar patrones clave en las interacciones miRNA-transcripción y sus implicaciones funcionales en EC(55).



*Figura 17: Análisis de enriquecimiento de los genes objetivo, asociados con los miARN individuales dentro de la red de transcripción miARN-objetivo. El eje y representa los términos de Gene Ontology (GO), que resumen grupos de términos seleccionados por REVIGO. Los miARN se ordenan según su cambio en el pliegue (fold change) en la comparación entre la enfermedad celíaca (EC) y los controles, desde el cambio en el pliegue más bajo (miR-31) hasta el cambio en el pliegue más alto (miR-1260b). La significancia del enriquecimiento se indica mediante una escala de colores (FDR), y la relación de genes (GeneRatio) refleja la proporción de genes objetivo de cada miARN en relación con el número total de genes en cada vía. Esta figura destaca el impacto funcional de miARN específicos y sus genes objetivo a través de diversos procesos biológicos(55).*

En el análisis de enriquecimiento de tipos celulares se utilizó xCell y GSVA para investigar los tipos celulares enriquecidos en biopsias de pacientes con EC. Los resultados mostraron un

enriquecimiento de células inmunes (células B y T) y un agotamiento de células epiteliales intestinales. Los miRNA regulados negativamente en EC estaban asociados con células epiteliales, mientras que los regulados positivamente estaban asociados con células inmunes (55).

Por último, se realizó un análisis bioinformático que vinculó genes de regiones de riesgo genético asociadas a la EC con una red de transcripción miRNA-diana. De los 118 genes considerados, 102 se expresaron en las biopsias y 13 estaban presentes en la red de transcripción miRNA-diana. Un ejemplo es STAT1, que se asocia con los miRNA miR-22-5p y miR-30a-3p, cuya interacción está respaldada por evidencia experimental (Figura 18)(55).

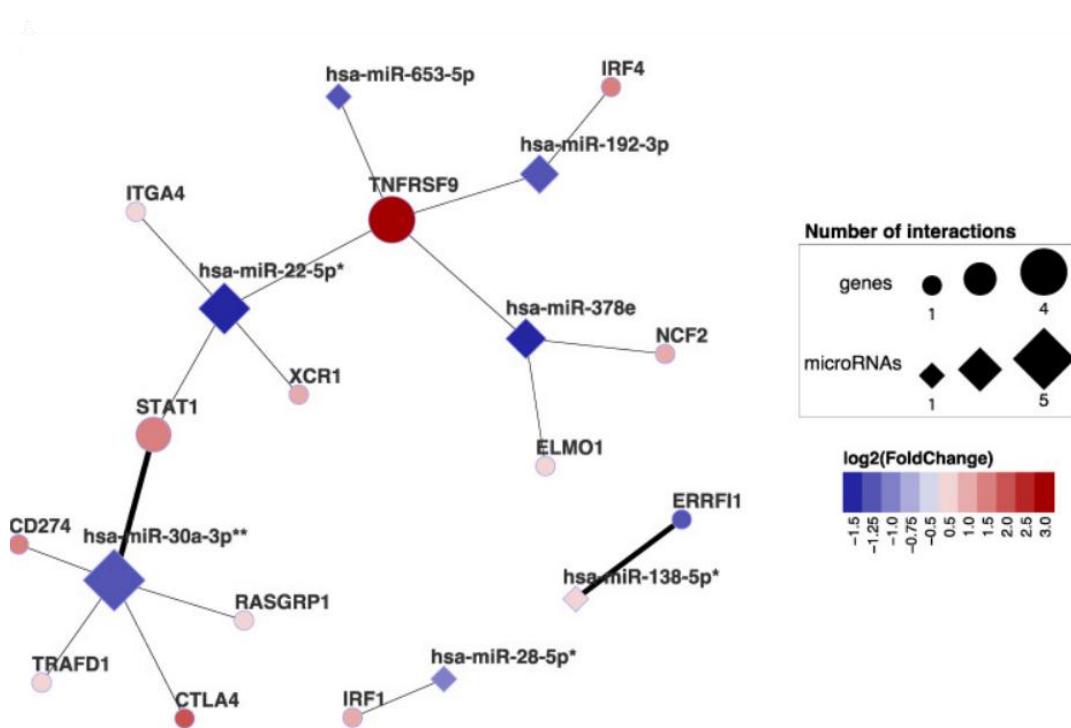


Figura 18: Subconjunto de la red de interacción miRNA-transcripción diana. Solo se muestran los genes que se superponen con los genes EC priorizados. Las líneas de conexión gruesas indican que la interacción se validó experimentalmente previamente. Después de la corrección de la composición del tipo de célula (puntuación de enriquecimiento basada en miRNA para células epiteliales intestinales y células CD3+), los miRNA indicados con un asterisco todavía muestran una tendencia entre EC y controles (FDR: \* 0,05–0,1, \*\* <0,05)(55).

**Tabla 4. Resultados de los artículos seleccionados.**

Autor/a/es	Técnicas/herramientas bioinformáticas	Resultados Biomarcadores	Limitaciones	Conclusiones
(47)	<ul style="list-style-type: none"> <li>- Análisis estadístico → uso de software R: t-Student (variables continuas), Chi-cuadrado (variables categóricas), algoritmoK-nearest neighbors (kNN) para estimar valores faltantes, coeficiente de correlación de Pearson y el índice de Youden para determinar sensibilidad y especificidad.</li> <li>- Análisis de proteínas diferencialmente expresadas mediante GO y KEGG Pathways: saber qué funciones desempeñan las DEPs y cómo interactúan en procesos biológicos.</li> <li>- Construcción de redes de interacción proteína-proteína con STRING y visualización de las redes con Cytoscape.</li> <li>- WGCNA: análisis de redes de coexpresión génica ponderada. Se empleó para identificar módulos de genes altamente correlacionados y explorar relaciones entre genes y fenotipos o rasgos biológicos.</li> <li>- Método de agrupamiento difuso o <i>fuzzy clustering</i> (Mfuzz). Se empleó para el análisis de expresión diferencial de las proteínas plasmáticas en pacientes con EC con diferentes grados de Marsh.</li> <li>- Algoritmos <i>machine learning/ ensemble method</i>: combinación de regresión logística, <i>random forest</i> y <i>support vector machine</i>- SVM.</li> </ul>	<ul style="list-style-type: none"> <li>- Duodeno → Sobreexpresados: CYP2C9, CYP344, HMGCS2, CELA3B, CELA2A; Infraexpresados: CAMP, LTF, ELANE, DEFA3, ADGPE2</li> <li>- Plasma → Sobreexpresados: FAM234A, TRHDE, APOC3, FGFBP2, KRTB2; Infraexpresados: GNMT, GLUCY1A1, PCMTD1, KRT13, IGHD</li> </ul>	<p>Tamaño limitado de la muestra de nuestro estudio.</p>	<p>Las proteínas FGL2 y TXNDC5 como posibles biomarcadores para el diagnóstico no invasivo de la EC. Las proteínas FABP, CPOX, BHMT y PPP2CB como biomarcadores potencialmente asociados con la clasificación patológica de la atrofia vellosa en el intestino delgado.</p>
(52)	<ul style="list-style-type: none"> <li>- Análisis ómicos mediante bases de datos: CTD, DISEASES y GeneCards. Se seleccionaron los genes comunes en las tres bases de datos y se usó un diagrama de Venn para realizar un análisis más profundo.</li> <li>- Herramientas bioinformáticas: Enrich, GeneMANIA, Cytoscape y WebGestalt para evaluar la función de los genes; GEO2R para el análisis de la expresión génica.</li> </ul>	<p>Se identificaron 54 genes de los cuales se seleccionaron algunos para realizar un estudio de redes génicas y análisis bioinformáticos. Se seleccionaron 25 genes para el <i>Seed Genes</i> y se identificaron 11 genes como <i>Hub Genes</i>:</p> <p>- <i>Seed Genes</i> (WebGestalt): IL17A, TGFB1, STAT1, ALB, IL10, IL2, IL4, IL1B, IL21, FUT2, IL5, IGF1, CD80, TLR4, IGFBP3, IL2RA, TGM2, CXCL10, CTLA4, IFIHI, HP, IRF1, IL15, NAGS, HLA-DRB1</p>	<p>No se mencionan.</p>	<p>Los genes IRF1, STAT1, ALB, IL10, IL2, IL4, IL17A, TGFB1, IL1B, IL6, TNF e IFNG como posibles biomarcadores para la EC e involucrados en la regulación del sistema inmunológico. El descubrimiento de nuevos biomarcadores diagnósticos para la EC permite comprender mejor la patología de la enfermedad y desarrollar tratamientos más eficaces.</p>

		<p>- <i>Hub Genes</i> (Cytoscape): STAT1, ALB, IL10, IL2, IL4, IL17A, TGFB1, IL1B, IL6, TNF e IFNG</p>		
(53)	<p>1) Análisis bioinformático de RNA-Seq</p> <ul style="list-style-type: none"><li>- Control de calidad: se usó FastQC en la plataforma Galaxy para evaluar la calidad de los datos de RNA-Seq.</li><li>- Preprocesamiento: eliminación de adaptadores y recorte de bases de baja calidad con TrimGalore.</li><li>- Alineamiento: lecturas mapeadas contra el genoma de referencia hg19 con Bowtie2.</li><li>- Conteo de lecturas: se cuantificaron las lecturas alineadas con HTSeq-count.</li><li>- Análisis diferencial: se identificaron miRNA diferencialmente expresados con DESeq2, aplicando correcciones de lote con COMBAT.</li><li>- Se realizó PCA para visualizar la variabilidad de los datos y distinguir los miRNA entre los tres grupos (Control, EC y DSG).</li></ul> <p>2) Uso de la bioinformática para el estudio estadístico</p> <ul style="list-style-type: none"><li>- Se utilizan herramientas computacionales avanzadas como R Bioconductor, GLMNET, CARET y pROC.</li></ul>	<p>Detección de 13 miRNA capaces de discriminar entre grupos de pacientes con EC, pacientes con EC en tratamiento con DSG y controles sanos. Además, mi-192-5p, miR-215-5p y miR-125b-5p pudieron discriminar entre los 3 grupos con alta precisión y especificidad.</p>	<p>Número limitado de pacientes y controles para el estudio.</p>	<p>El uso de los microRNA circulantes identificados como nuevas herramientas de diagnóstico y seguimiento para la EC pediátrica. Además, comprobaron el efecto de la DSG en la regulación de algunos miRNA.</p>
(54)	<p>1) PCR cuantitativa en tiempo real</p> <ul style="list-style-type: none"><li>- Identificación de objetivos de miRNA: se emplearon bases de datos y herramientas bioinformáticas como miRBase, TargetScan, microRNA.org y Mirò para predecir las posibles dianas de los miRNA.</li><li>- Análisis de la expresión génica y miRNA: normalización y el cálculo de la expresión relativa junto con la técnica de laboratorio experimental qPCR.</li></ul> <p>2) Western Blot</p> <ul style="list-style-type: none"><li>- Para la identificación de proteínas se usó el software Quantity One 4.1.1 para analizar la intensidad de las señales en el Western Blot.</li><li>- Análisis de imágenes de las membranas de Western Blot mediante el FluorS Multilanner. El software procesó las imágenes para cuantificar la expresión de proteínas.</li></ul>	<p>1) Muestras de biopsia de duodeno:</p> <ul style="list-style-type: none"><li>- miRNA regulados positivamente: miR-21-5p, miR-21-3p, miR-486-5p.</li><li>- miRNA regulados negativamente: miR-192-5p, miR-31-5p, miR-338-3p.</li></ul> <p>2) Análisis en plasma: resultados similares observados en biopsias, pero en presencia de DSG no se pudieron detectar valores comparables de maneras significativa con los controles.</p>	<p>Número limitado de pacientes y controles. El estudio se realizó en plasma, que puede contener una mayor cantidad de contaminantes de PCR y no tuvieron en cuenta si había hemólisis.</p>	<p>Los miRNA circulantes y de tejido de duodeno como biomarcadores para la EC en pacientes pediátricos. Sin embargo, recalcan que las alteraciones en la expresión de los miRNA podrían afectar de manera diferente a la de los adultos.</p>

---

(55)	<ol style="list-style-type: none"><li>1) Aislamiento y secuenciación de ARN: Kit miRVana, qiazol, TruSeq Small RNA Sample Prep, HiSeq 2500, miRBase 22, human_g1k_v37 ensemble Release 75.</li><li>2) Expresión diferencial miRNA y RNA: R DESeq2 (v1.26.0), corrección de Benjamini-Hochberg (FDR).</li><li>3) Red de transcripción miRNA-diana: TargetScan (v7.2), microTCDS (v7.0), TarBase (v7.0), miRTarBase (v7.0), correlaciones de Pearson, RedeR.</li><li>4) Análisis de enriquecimiento de vías: R clusterProfiler (v3.14.3), Gene Ontology (GO), REVIGO.</li><li>5) Enriquecimiento de tipos celulares: xCell, R GSVA (v1.34.0), atlas de expresión de miRNA</li></ol>	<p>Construcción de una red que incluye 2030 interacciones de transcripción miRNA-diana, incluidos 423 pares validados experimentalmente. En análisis de vías reveló que las interacciones están implicadas en vías relacionadas con el sistema inmunitario y en vías metabólicas (por ejemplo, el metabolismo de los lípidos). La red incluye 13 genes que se han relacionado con la desregulación por variantes genómicas asociadas a la EC, incluido STAT1.</p>	<p>Una limitación en la investigación de miRNA es que la mayoría de las interacciones miRNA-transcripción diana han sido predichas por algoritmos de predicción de diana pero no han sido validadas experimentalmente.</p>	<p>Los miRNA desregulados en la EC podrían influir en vías metabólicas, del ciclo celular e inmunitarias. Se requieren estudios funcionales para validar su papel en la patogénesis, idealmente con enfoque en tipos celulares específicos. Futuros estudios comparativos con otras enfermedades intestinales (como la enfermedad de Crohn) pueden aportar más información. Comprender mejor estos miRNA podría facilitar la identificación de biomarcadores y nuevas terapias para la EC.</p>
------	---	---	--	--

---

## 6. DISCUSIÓN

Los estudios revisados destacan un avance significativo en la identificación de biomarcadores para la enfermedad celíaca (EC), tanto en el contexto de proteómica como de análisis de miRNA. La proteómica, como se muestra en el análisis de Li y colaboradores (2024), ha revelado proteínas clave que se expresan diferencialmente en muestras de duodeno y plasma, sugiriendo alteraciones metabólicas cruciales en la EC. El uso de técnicas avanzadas de espectrometría, como la proteómica de adquisición dependiente de datos en cuatro dimensiones (4D-DIA), ha permitido identificar proteínas con potencial diagnóstico, como FGL2 y TXNDC5, que podrían usarse para mejorar el diagnóstico de esta enfermedad. Estos biomarcadores también muestran promesas en la diferenciación de los diferentes grados de atrofia intestinal, lo que subraya su relevancia clínica.

Por otro lado, el estudio de los microRNAs (miRNA) en la EC ha demostrado que estos pequeños ARN no codificantes desempeñan un papel esencial en la regulación del sistema inmunológico y en la patogenia de la enfermedad. Los miRNA circulantes son atractivos como biomarcadores no invasivos debido a su estabilidad en fluidos biológicos, lo que facilita su uso en el diagnóstico de la EC. El trabajo de Buoli Comani y colaboradores (2015) y Felli y colaboradores (2022) señala que ciertos miRNA, como miR-192-5p y miR-215-5p, podrían servir para diferenciar entre EC activa y pacientes tratados con dieta sin gluten (DSG). Estos hallazgos resaltan cómo los miRNA podrían reflejar el estado inflamatorio y de reparación intestinal en los pacientes celíacos.

Además, los análisis bioinformáticos realizados en los estudios revisados, como los de Noori y colaboradores (2020) y Luise Tan y colaboradores (2021), evidencian la complejidad de la red de interacciones moleculares asociadas con la EC. El análisis de genes relacionados con la EC revela una fuerte implicación en el sistema inmunológico y en las vías de señalización autoinmunes, lo que resalta la importancia de estudiar las interacciones genéticas para comprender mejor la fisiopatología de la enfermedad. Los enfoques bioinformáticos en combinación con análisis experimentales permiten identificar genes clave como STAT1, IL6 y TNF, que podrían ser utilizados como biomarcadores para el diagnóstico y tratamiento de la EC.

## 7. Conclusiones

La identificación de biomarcadores para la EC, tanto a nivel proteómico como de microRNA, representa un avance significativo en la mejora del diagnóstico y tratamiento de esta enfermedad. Las técnicas avanzadas de proteómica y las herramientas bioinformáticas han permitido identificar proteínas clave como FGL2 y TXNDC5, que podrían servir como biomarcadores diagnósticos no invasivos y como indicadores de la progresión de la enfermedad. Además, los miRNA circulantes, debido a su estabilidad y facilidad de medición en fluidos biológicos, muestran un gran potencial para discriminar entre pacientes con EC activa y aquellos tratados con una dieta sin gluten.

Por otro lado, los análisis funcionales de genes relacionados con la EC y las redes de interacción molecular ofrecen una visión más profunda sobre los mecanismos subyacentes de la enfermedad. Los avances en la investigación de las interacciones entre miRNA y sus genes diana han proporcionado una comprensión más completa de los procesos biológicos alterados en la EC, especialmente aquellos involucrados en la regulación inmunológica y metabólica. Sin embargo, se requieren más estudios para validar estos biomarcadores en cohortes más grandes y explorar su aplicabilidad clínica en el diagnóstico y seguimiento de la enfermedad celíaca.

## 8. Referencias bibliográficas

1. Ferretti G, Bacchetti T, Masciangelo S, Saturni L. Celiac disease, inflammation and oxidative damage: a nutrigenetic approach. *Nutrients* [Internet]. 2012 [cited 2025 Jan 27];4(4):243–57. Available from: <https://pubmed.ncbi.nlm.nih.gov/22606367/>
2. Senosiain BC, Pascual AG, Tirado MR, García JC. Enfermedad celíaca. *Medicine* [Internet]. 2024 Mar 1 [cited 2025 Jan 29];14(5):233–41. Available from: <https://www.medicineonline.es/es-enfermedad-celiaca-articulo-S030454122400060X>
3. Arranz E, Garrote JA. Inmunología de la enfermedad celíaca. *Gastroenterol Hepatol* [Internet]. 2010 Nov 1 [cited 2025 Jan 29];33(9):643–51. Available from: <https://www.elsevier.es/es-revista-gastroenterologia-hepatologia-14-articulo-inmunologia-enfermedad-celiaca-S0210570509005524>

4. Plaza Izurieta L, Fernandez Jimenez N, Bilbao Catalá JR. Genética de la enfermedad celíaca. Genes HLA y NO-HLA. Avances en el conocimiento de las patologías relacionadas con el gluten y evolución de los alimentos sin gluten, 2018, ISBN 978-84-09-04161-9, págs 61-80 [Internet]. 2018 [cited 2025 Jan 28];61–80. Available from: <https://dialnet.unirioja.es/servlet/articulo?codigo=9117048>
5. Wu X, Qian L, Liu K, Wu J, Shan Z. Gastrointestinal microbiome and gluten in celiac disease. Ann Med [Internet]. 2021 [cited 2025 Jan 27];53(1):1797. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8519548/>
6. Carreras J. Artificial Intelligence Analysis of Celiac Disease Using an Autoimmune Discovery Transcriptomic Panel Highlighted Pathogenic Genes including BTLA. Healthcare (Basel) [Internet]. 2022 Aug 1 [cited 2025 Jan 27];10(8). Available from: <https://pubmed.ncbi.nlm.nih.gov/36011206/>
7. Shewry PR, Halford NG, Belton PS, Tatham AS. The structure and properties of gluten: an elastic protein from wheat grain. Philos Trans R Soc Lond B Biol Sci [Internet]. 2002 Feb 28 [cited 2025 Jan 28];357(1418):133–42. Available from: <https://pubmed.ncbi.nlm.nih.gov/11911770/>
8. Shan L, Molberg Ø, Parrot I, Hausch F, Filiz F, Gray GM, et al. Structural basis for gluten intolerance in celiac sprue. Science [Internet]. 2002 Sep 27 [cited 2025 Jan 28];297(5590):2275–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/12351792/>
9. Sollid LM. Coeliac disease: dissecting a complex inflammatory disorder. Nat Rev Immunol [Internet]. 2002 Sep [cited 2025 Jan 28];2(9):647–55. Available from: <https://pubmed.ncbi.nlm.nih.gov/12209133/>
10. Milletich PL, Ahrens AP, Russell JT, Petrone JR, Berryman MA, Agardh D, et al. Gut microbiome markers in subgroups of HLA class II genotyped infants signal future celiac disease in the general population: ABIS study. Front Cell Infect Microbiol [Internet]. 2022 Jul 25 [cited 2025 Jan 28];12. Available from: <https://pubmed.ncbi.nlm.nih.gov/35959362/>
11. Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, Greco L, et al. HLA types in celiac disease patients not carrying the DQA1 \*05-DQB1 \*02 (DQ2) heterodimer: Results from the European genetics cluster on celiac disease. Hum Immunol [Internet]. 2003 Apr 1 [cited

- 2025 Jan 28];64(4):469–77. Available from: <https://pubmed.ncbi.nlm.nih.gov/12651074/>
12. Sollid LM, Thorsby E. HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology* [Internet]. 1993 [cited 2025 Jan 28];105(3):910–22. Available from: <https://pubmed.ncbi.nlm.nih.gov/8359659/>
13. Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, Maglio M, et al. The first large population based twin study of coeliac disease. *Gut* [Internet]. 2002 [cited 2025 Jan 28];50(5):624–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/11950806/>
14. Wolters VM, Wijmenga C. Genetic background of celiac disease and its clinical implications. *Am J Gastroenterol* [Internet]. 2008 Jan [cited 2025 Jan 28];103(1):190–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/18184122/>
15. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol* [Internet]. 2011 Apr 23 [cited 2025 Jan 28];29:493–525. Available from: <https://pubmed.ncbi.nlm.nih.gov/21219178/>
16. Jamnik J, Villa CR, Dhir SB, Jenkins DJA, El-Sohemy A. Prevalence of positive coeliac disease serology and HLA risk genotypes in a multiethnic population of adults in Canada: a cross-sectional study. *BMJ Open* [Internet]. 2017 Oct 1 [cited 2025 Feb 1];7(10). Available from: <https://pubmed.ncbi.nlm.nih.gov/39272232/>
17. Pallav K, Kabbani T, Tariq S, Vanga R, Kelly CP, Leffler DA. Clinical utility of celiac disease-associated HLA testing. *Dig Dis Sci* [Internet]. 2014 Sep 1 [cited 2025 Jan 29];59(9):2199–206. Available from: <https://pubmed.ncbi.nlm.nih.gov/24705698/>
18. Husby S, Koletzko S, Korponay-Szabó I, Kurppa K, Mearin ML, Ribes-Koninckx C, et al. European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. *J Pediatr Gastroenterol Nutr* [Internet]. 2020 Jan 1 [cited 2025 Jan 29];70(1):141–56. Available from: <https://pubmed.ncbi.nlm.nih.gov/31568151/>
19. Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, et al. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat Med* [Internet]. 1997

- Jul [cited 2025 Jan 29];3(7):797–801. Available from: <https://pubmed.ncbi.nlm.nih.gov/9212111/>
20. Arentz-Hansen H, Körner R, Molberg Ø, Quarsten H, Vader W, Kooy YM, et al. The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J Exp Med* [Internet]. 2000 Feb 21 [cited 2025 Jan 29];191(4):603–12. Available from: <https://pubmed.ncbi.nlm.nih.gov/10684852/>
21. Lebreton C, Ménard S, Abed J, Moura IC, Coppo R, Dugave C, et al. Interactions among secretory immunoglobulin A, CD71, and transglutaminase-2 affect permeability of intestinal epithelial cells to gliadin peptides. *Gastroenterology* [Internet]. 2012 [cited 2025 Jan 29];143(3). Available from: <https://pubmed.ncbi.nlm.nih.gov/22750506/>
22. Bruce SE, Bjarnason I, Peters TJ. Human jejunal transglutaminase: demonstration of activity, enzyme kinetics and substrate specificity with special relation to gliadin and coeliac disease. *Clin Sci (Lond)* [Internet]. 1985 [cited 2025 Jan 29];68(5):573–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/2858282/>
23. Beitnes ACR, Ráki M, Brottveit M, Lundin KEA, Jahnson FL, Sollid LM. Rapid Accumulation of CD14+CD11c+ Dendritic Cells in Gut Mucosa of Celiac Disease after in vivo Gluten Challenge. *PLoS One* [Internet]. 2012 Mar 16 [cited 2025 Jan 29];7(3):e33556. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0033556>
24. Fina D, Sarra M, Caruso R, Del Vecchio Blanco G, Pallone F, MacDonald TT, et al. Interleukin 21 contributes to the mucosal T helper cell type 1 response in coeliac disease. *Gut* [Internet]. 2008 Jul [cited 2025 Jan 29];57(7):887–92. Available from: <https://pubmed.ncbi.nlm.nih.gov/17965065/>
25. Salvati VM, MacDonald TT, Bajaj-Elliott M, Borrelli M, Staiano A, Auricchio S, et al. Interleukin 18 and associated markers of T helper cell type 1 activity in coeliac disease. *Gut* [Internet]. 2002 [cited 2025 Jan 29];50(2):186–90. Available from: <https://pubmed.ncbi.nlm.nih.gov/11788557/>
26. Di Niro R, Mesin L, Zheng NY, Stamnaes J, Morrissey M, Lee JH, et al. High abundance of plasma cells secreting transglutaminase 2-specific IgA autoantibodies with limited

- somatic hypermutation in celiac disease intestinal lesions. *Nat Med* [Internet]. 2012 Mar [cited 2025 Jan 29];18(3):441–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/22366952/>
27. Mesin L, Sollid LM, Di Niro R. The intestinal B-cell response in celiac disease. *Front Immunol* [Internet]. 2012 [cited 2025 Jan 29];3(OCT). Available from: <https://pubmed.ncbi.nlm.nih.gov/23060888/>
28. Hüe S, Mention JJ, Monteiro RC, Zhang SL, Cellier C, Schmitz J, et al. A direct role for NKG2D/MICA interaction in villous atrophy during celiac disease. *Immunity* [Internet]. 2004 [cited 2025 Jan 29];21(3):367–77. Available from: <https://pubmed.ncbi.nlm.nih.gov/15357948/>
29. Olivares M, Neef A, Castillejo G, De Palma G, Varea V, Capilla A, et al. The HLA-DQ2 genotype selects for early intestinal microbiota composition in infants at high risk of developing coeliac disease. *Gut* [Internet]. 2015 Mar 1 [cited 2025 Feb 1];64(3):406–17. Available from: <https://pubmed.ncbi.nlm.nih.gov/24939571/>
30. de Palma G, Capilla A, Nova E, Castillejo G, Varea V, Pozo T, et al. Influence of milk-feeding type and genetic risk of developing coeliac disease on intestinal microbiota of infants: the PROFICEL study. *PLoS One* [Internet]. 2012 Feb 3 [cited 2025 Feb 1];7(2). Available from: <https://pubmed.ncbi.nlm.nih.gov/22319588/>
31. Russell JT, Roesch LFW, Ördberg M, Ilonen J, Atkinson MA, Schatz DA, et al. Genetic risk for autoimmunity is associated with distinct changes in the human gut microbiome. *Nat Commun* [Internet]. 2019 Dec 1 [cited 2025 Feb 1];10(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/31399563/>
32. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, et al. Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* [Internet]. 2006 Aug [cited 2025 Feb 1];118(2):511–21. Available from: <https://pubmed.ncbi.nlm.nih.gov/16882802/>
33. Klemenak M, Dolinšek J, Langerholc T, Di Gioia D, Mičetić-Turk D. Administration of *Bifidobacterium breve* Decreases the Production of TNF- $\alpha$  in Children with Celiac Disease. *Dig Dis Sci* [Internet]. 2015 Nov 1 [cited 2025 Feb 1];60(11):3386–92. Available from: <https://pubmed.ncbi.nlm.nih.gov/26134988/>

34. Smecuol E, Hwang HJ, Sugai E, Corso L, Cherñavsky AC, Bellavite FP, et al. Exploratory, randomized, double-blind, placebo-controlled study on the effects of *Bifidobacterium infantis* natren life start strain super strain in active celiac disease. *J Clin Gastroenterol* [Internet]. 2013 Feb [cited 2025 Feb 1];47(2):139–47. Available from: <https://pubmed.ncbi.nlm.nih.gov/23314670/>
35. Olivares M, Castillejo G, Varea V, Sanz Y. Double-blind, randomised, placebo-controlled intervention trial to evaluate the effects of *Bifidobacterium longum* CECT 7347 in children with newly diagnosed coeliac disease. *Br J Nutr* [Internet]. 2014 Jul 14 [cited 2025 Feb 1];112(1):30–40. Available from: <https://pubmed.ncbi.nlm.nih.gov/24774670/>
36. Rostami-Nejad M, Asri N, Bakhtiari S, Khalkhal E, Maleki S, Rezaei-Tavirani M, et al. Metabolomics and lipidomics signature in celiac disease: a narrative review. *Clin Exp Med* [Internet]. 2024 Dec 1 [cited 2025 Feb 4];24(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/38340186/>
37. Lindfors K, Ciacci C, Kurppa K, Lundin KEA, Makharia GK, Mearin ML, et al. Coeliac disease. *Nat Rev Dis Primers* [Internet]. 2019 Dec 1 [cited 2025 Jan 29];5(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/30631077/>
38. Cerdá Contreras E. Enfermedad celíaca: bases moleculares y clínicas. *Rev Gastroenterol Mex* [Internet]. 2010 Nov 1 [cited 2025 Jan 29];75:181–4. Available from: <https://www.revistagastroenterologiamexico.org//es-enfermedad-celiaca-bases-moleculares-clinicas-articulo-X037509061087393X>
39. Niewinski MM. Advances in celiac disease and gluten-free diet. *J Am Diet Assoc* [Internet]. 2008 Apr [cited 2025 Jan 29];108(4):661–72. Available from: <https://pubmed.ncbi.nlm.nih.gov/18375224/>
40. D'heudene M, Vanuytsel T, Wauters L. Celiac disease: Hope for new treatments beyond a gluten-free diet. *Clin Nutr* [Internet]. 2024 Jun 1 [cited 2025 Jan 29];43(6):1240–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/38648685/>
41. Noori E, Hashemi N, Rezaee D, Maleki R, Shams F, Kazemi B, et al. Potential therapeutic options for celiac Disease: An update on Current evidence from Gluten-Free diet to cell

- therapy. *Int Immunopharmacol* [Internet]. 2024 May 30 [cited 2025 Jan 29];133. Available from: <https://pubmed.ncbi.nlm.nih.gov/38608449/>
42. Perezleo Solorzano L, Arencibia Jorge R, Conill González C, Achón Veloz G, Araújo Ruiz JA. Impacto de la Bioinformática en las ciencias biomédicas. *Revista Cubana de Información en Ciencias de la Salud*, ISSN-e 2307-2113, ISSN 1024-9435, Vol 11, N° 4, 2003 [Internet]. 2003 [cited 2025 Jan 29];11(4):6. Available from: <https://dialnet.unirioja.es/servlet/articulo?codigo=3173231&info=resumen&idioma=ENG>
43. Hagen JB. The origins of bioinformatics. *Nat Rev Genet* [Internet]. 2000 [cited 2025 Jan 29];1(3):231–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/11252753/>
44. Lago-Sampedro AM, García-Escobar E, Lago-Sampedro AM, García-Escobar E. Importancia de la bioinformática en la medicina actual. ¿Es realmente necesaria la bioinformática en la práctica clínica? *Nutr Hosp* [Internet]. 2022 May 1 [cited 2025 Jan 29];39(3):487–8. Available from: [https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S0212-16112022000400001&lng=es&nrm=iso&tlng=es](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0212-16112022000400001&lng=es&nrm=iso&tlng=es)
45. La bioinformática y la nutrición de precisión pueden evitar diabetes, enfermedades cardiovasculares y obesidad, según expertos de UNIR | UNIR [Internet]. [cited 2025 Jan 29]. Available from: <https://www.unir.net/actualidad/vida-academica/bioinformatica-y-nutricion-de-precision-para-evitar-enfermedades-cronicas/>
46. Tan IL, Coutinho de Almeida R, Modderman R, Stachurska A, Dekens J, Barisani D, et al. Circulating miRNAs as Potential Biomarkers for Celiac Disease Development. *Front Immunol* [Internet]. 2021 Dec 7 [cited 2025 Feb 5];12. Available from: <https://pubmed.ncbi.nlm.nih.gov/34950132/>
47. Li N, Maimaitireyimu A, Shi T, Feng Y, Liu W, Xue S, et al. Proteomic analysis of plasma and duodenal tissue in celiac disease patients reveals potential noninvasive diagnostic biomarkers. *Sci Rep* [Internet]. 2024 Dec 1 [cited 2025 Feb 5];14(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/39622892/>
48. Lerner A, Matthias T. Changes in intestinal tight junction permeability associated with industrial food additives explain the rising incidence of autoimmune disease.

- Autoimmun Rev [Internet]. 2015 [cited 2025 Mar 1];14(6):479–89. Available from: <https://pubmed.ncbi.nlm.nih.gov/25676324/>
49. Zachariou M, Minadakis G, Oulas A, Afxenti S, Spyrou GM. Integrating multi-source information on a single network to detect disease-related clusters of molecular mechanisms. J Proteomics [Internet]. 2018 Sep 30 [cited 2025 Mar 1];188:15–29. Available from: <https://pubmed.ncbi.nlm.nih.gov/29545169/>
50. Karns R, Tabar S, Bardes EE, Jegga AG, Aronow BJ. How Do Bioinformatics Approaches Apply to the Analysis and Understanding of Disease Pathology? Pathobiology of Human Disease: A Dynamic Encyclopedia of Disease Mechanisms [Internet]. 2014 Jan 1 [cited 2025 Mar 1];4140–57. Available from: [https://www.researchgate.net/publication/300284017\\_How\\_Do\\_Bioinformatics\\_Approaches\\_Apply\\_to\\_the\\_Analysis\\_and\\_Understanding\\_of\\_Disease\\_Pathology](https://www.researchgate.net/publication/300284017_How_Do_Bioinformatics_Approaches_Apply_to_the_Analysis_and_Understanding_of_Disease_Pathology)
51. Yepes-Nuñez JJ, Urrútia G, Romero-García M, Alonso-Fernández S. Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. Rev Esp Cardiol. 2021 Sep 1;74(9):790–9.
52. Noori E, Kazemi B, Bandehpour M, Zali H, Khalesi B, Khalili S. Deciphering crucial genes in coeliac disease by bioinformatics analysis. Autoimmunity [Internet]. 2020 Feb 17 [cited 2025 Feb 10];53(2):102–13. Available from: <https://pubmed.ncbi.nlm.nih.gov/31809599/>
53. Felli C, Baldassarre A, Uva P, Alisi A, Cangelosi D, Ancinelli M, et al. Circulating microRNAs as novel non-invasive biomarkers of paediatric celiac disease and adherence to gluten-free diet. EBioMedicine [Internet]. 2022 Feb 1 [cited 2025 Feb 16];76. Available from: <https://pubmed.ncbi.nlm.nih.gov/35151110/>
54. Buoli Comani G, Panceri R, Dinelli M, Biondi A, Mancuso C, Meneveri R, et al. miRNA-regulated gene expression differs in celiac disease patients according to the age of presentation. Genes Nutr [Internet]. 2015 Sep 6 [cited 2025 Feb 18];10(5):32. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC4522246/>
55. Tan IL, Barisani D, Panceri R, Modderman R, Visschedij M, Weersma RK, et al. A Combined mRNA- and miRNA-Sequencing Approach Reveals miRNAs as Potential Regulators of the Small Intestinal Transcriptome in Celiac Disease. Int J Mol Sci

- [Internet]. 2021 Nov 1 [cited 2025 Feb 23];22(21). Available from: <https://pubmed.ncbi.nlm.nih.gov/34768815/>
56. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* [Internet]. 2008 Jul 29 [cited 2025 Feb 16];105(30):10513–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/18663219/>
57. Chen X, Liang H, Zhang J, Zen K, Zhang CY. Secreted microRNAs: a new form of intercellular communication. *Trends Cell Biol* [Internet]. 2012 Mar [cited 2025 Feb 16];22(3):125–32. Available from: <https://pubmed.ncbi.nlm.nih.gov/22260888/>
58. Li M, Marin-Muller C, Bharadwaj U, Chow KH, Yao Q, Chen C. MicroRNAs: control and loss of control in human physiology and disease. *World J Surg* [Internet]. 2009 Apr [cited 2025 Feb 16];33(4):667–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/19030926/>