

## Capítulo 6

### La Enfermedad Celíaca: Marcadores genéticos

Nora Fernández-Jiménez, Leticia Plaza-Izurieta, Jose Ramón Bilbao

Departamento de Genética, Antropología Física y Fisiología Animal, Universidad del País Vasco (UPV-EHU), Instituto de Investigación BioCruces, Bizkaia.

[immunogenetics.let@gmail.com](mailto:immunogenetics.let@gmail.com)

Doi: <http://dx.doi.org/10.3926/oms.23>

#### Referenciar este capítulo

Fernandez-Jimenez N, Plaza-Izurieta L, Bilbao JR. *La Enfermedad Celíaca: Marcadores genéticos*. En Rodrigo L y Peña AS, editores. *Enfermedad celíaca y sensibilidad al gluten no celíaca*. Barcelona, España: OmniaScience; 2013. p. 103-121.

## Resumen

Aunque el modo de herencia de la enfermedad celíaca (EC) es aún desconocido, existen muchas evidencias a favor de que la genética participe en la predisposición a la enfermedad. Hoy en día, se calcula que la heredabilidad de la EC está cerca del 87%.

Se conoce desde hace mucho tiempo que gran parte del riesgo genético a la EC se debe a la presencia de ciertos alelos HLA. A pesar de su papel determinante en la patogénesis de la enfermedad, su contribución a la herencia de la misma es modesta (<50%) por lo que se especula sobre la existencia de numerosos *loci* de susceptibilidad no ligados al HLA, cada uno de los cuales tendría un efecto muy pequeño sobre el riesgo global.

En consecuencia, durante los últimos años, se han realizado numerosos esfuerzos para localizar e identificar genes de susceptibilidad adicionales que puedan explicar la genética de esta enfermedad. Para ello, se han utilizado los estudios de ligamiento en familias y los estudios de asociación. Además, recientemente, se ha investigado la EC mediante estudios de asociación de genoma completo (GWAS), en los cuales se han analizado miles de Single-Nucleotide Polymorphisms (SNPs). Gracias a estos estudios, se han identificado varios genes asociados a la EC, pero no todas las asociaciones observadas se han confirmado en estudios posteriores. Además, la contribución de los genes identificados sigue siendo modesta, y todavía queda una parte importante de la genética de la EC sin esclarecer.

## Abstract

Although the mode of inheritance of celiac disease (CD) is not completely understood, there is abundant evidence supporting the implication of genetic factors in the susceptibility of CD, and the heritability of CD has been estimated to be around 87%.

It has been known for a long time that certain HLA alleles are the major contributor to CD risk. However, despite playing a determinant role in the pathogenesis of the disease, their contribution to inheritance is modest (<50%) and it is believed that there must exist several non-HLA susceptibility *loci*, each one of them with a very small effect on the overall risk.

Consequently, during the last years, a great amount of effort has been made to locate and identify those additional susceptibility genes that might explain the genetics of the disease. Linkage studies in families, candidate gene association studies and more recently, genome wide association studies (GWAS) analyzing hundred of thousand of Single-Nucleotide Polymorphisms (SNPs) have been performed. These approaches have identified several genes that are associated with CD, but not all of them have been confirmed in subsequent studies. Besides, the contribution of the genes identified remains modest, and a large part of the genetics of CD remains to be clarified.

## 1. Introducción

Aunque el modelo de herencia de la enfermedad celíaca (EC) es aún desconocido, hace tiempo que se conoce que la herencia participa en la predisposición a la enfermedad. Los estudios de prevalencia en familias afectadas, y sobre todo aquellos que comparan parejas de gemelos, han sido de gran utilidad para estimar las proporciones en las que los factores de riesgo genéticos y ambientales contribuyen en el desarrollo de las enfermedades. De acuerdo con estos estudios, la Genética juega un papel importante tanto en el inicio como en el posterior desarrollo de la EC. En general, se acepta que la proporción de parejas de gemelos monocigóticos o idénticos en los que ambos padecen la enfermedad se sitúa en torno al 75-86%, mientras que, entre los gemelos dicigóticos o mellizos (que como todos los hermanos, comparten un promedio del 50% del genoma) esta concordancia se reduce hasta el 16-20%. Esta diferencia entre gemelos mono y dicigóticos sirve para calcular el tamaño del componente genético en la EC, que es mayor que en otras patologías complejas de origen inmunológico, como la diabetes tipo 1 (alrededor de un 30% de concordancia entre gemelos idénticos y un 6% en los mellizos).<sup>1</sup> Además, en la EC, la concordancia entre parejas de hermanos y de mellizos es prácticamente la misma, con lo que el componente ambiental tendría un efecto mínimo sobre el riesgo de desarrollar la enfermedad. Todo lo anterior apoya la idea de que existe un fuerte componente genético en el desarrollo de la celiacía. Hoy en día, se calcula que la heredabilidad de la EC (proporción del riesgo de padecer una enfermedad que es atribuible a factores genéticos, frente a los ambientales) está cerca del 87%.<sup>2</sup>

Hace tiempo que se sabe que gran parte del riesgo genético a padecer la EC se debe a la presencia de ciertos alelos del antígeno leucocitario humano o HLA. A pesar de su papel determinante en la patogénesis de la enfermedad, su contribución a la herencia de la misma es modesta, por lo que se ha especulado sobre la existencia de numerosos *loci* de susceptibilidad no ligados al HLA, cada uno de los cuales tendría un efecto muy pequeño sobre el riesgo global.

## 2. La región HLA y la Enfermedad Celíaca

### 2.1. Región HLA

El Antígeno Leucocitario Humano ó HLA (acrónimo inglés de *Human Leucocyte Antigen*) es el nombre que recibe el Complejo Mayor de Histocompatibilidad (MHC) en humanos. Es un *superlocus* localizado en el brazo corto del cromosoma 6 y contiene un gran número de genes relacionados con el sistema inmune. Los genes HLA son los responsables de codificar las proteínas presentadoras de antígenos que se expresan en la superficie de la mayoría de las células del ser humano y constituyen una pieza fundamental en la capacidad de discernir entre lo propio y lo extraño.

Los genes HLA influyen en la aparición de numerosas enfermedades inflamatorias y autoinmunes, así como en la susceptibilidad a desarrollar enfermedades infecciosas como la malaria o el SIDA. Sin embargo, debido a la complejidad de la región, se desconocen los componentes genéticos y los mecanismos patogénicos concretos para la mayoría de estas enfermedades. La región HLA es una de las regiones con mayor densidad génica del genoma.

Una explicación para este fenómeno sería que en esta región se favorece un nivel alto de expresión.<sup>3</sup>

## 2.2. Contribución al riesgo genético y genes de susceptibilidad

Como se ha mencionado anteriormente, la región HLA es el *locus* de susceptibilidad más importante en la EC y explica alrededor del 50% de la heredabilidad de la enfermedad. Las primeras evidencias de asociación entre la región HLA y la EC se publicaron en 1972 y se encontraron mediante el uso de métodos serológicos. Debido al alto grado de desequilibrio de ligamiento en la zona, los primeros estudios identificaron las variantes HLA-A1, HLA-B8 y HLA-DR3 como las variantes etiológicas en la región, pero los estudios moleculares han evidenciado que los factores directamente implicados son los genes HLA de clase II que codifican para las moléculas HLA-DQ2 y HLA-DQ8. La asociación de HLA-DQ2 con la enfermedad es la más fuerte y así, alrededor del 90% de los pacientes celíacos presentan al menos una copia del heterodímero HLA-DQ2.5 (formado por la combinación de los alelos DQA1\*05 y DQB1\*02, encargados de codificar las cadenas  $\alpha$  y  $\beta$  del heterodímero, respectivamente). Por otro lado, un 20-30% de la población no celíaca también es portadora de esta variante de riesgo, lo que demuestra que aún siendo muy importante, HLA-DQ2 es por sí solo insuficiente para desarrollar la enfermedad. La gran mayoría de los pacientes con EC que carecen de HLA-DQ2 son portadores de la variante DQ8, presente en el haplotipo formado por los alelos DQA1\*0301 y DQB1\*0302.<sup>4</sup> Una proporción muy pequeña de los pacientes son negativos tanto para DQ2 como para DQ8, pero se ha observado que en la mayoría de los casos, estos individuos presentan al menos uno de los dos alelos que codifican la molécula DQ2, es decir, DQA1\*05 o DQB1\*02.<sup>4,5</sup>

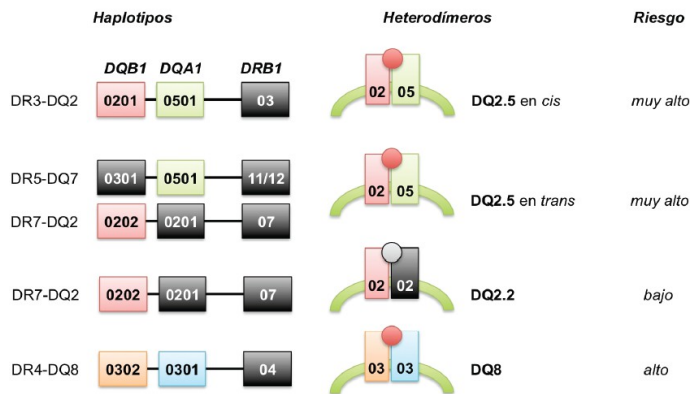


Figura 1. Asociación del locus HLA con la EC. La molécula HLA-DQ2 es el factor principal de riesgo genético a EC. La mayor parte de los individuos celíacos expresan el heterodímero HLA-DQ2.5 codificado por los alelos HLA-DQA1\*05 (cadena  $\alpha$ ) y HLA-DQB1\*02 (cadena  $\beta$ ), que pueden encontrarse en cis en el haplotipo DR3-DQ2 o en trans en los heterocigotos DR5-DQ7 y DR7-DQ2.2. El dímero HLA-DQ2.2, variante de HLA-DQ2 (codificada por los alelos HLA-DQA1\*0201 y HLA-DQB1\*02) confiere un riesgo bajo de desarrollar EC. La mayoría de pacientes DQ2-negativos expresan HLA-DQ8, codificado por el haplotipo DR4-DQ8

Las variantes de riesgo DQ2 y DQ8, están en desequilibrio de ligamiento (estrechamente asociadas) con las variantes del gen *HLA-DRB1*, DR3 y DR4, respectivamente. Por ello, al referirse a estas variantes de riesgo, se habla de los haplotipos DR3-DQ2 y DR4-DQ8.<sup>6</sup> Los haplotipos que codifican el heterodímero de riesgo HLA-DQ2.5 se han asociado con la EC en la mayoría de las poblaciones (Figura 1). En algunos haplotipos, como el DR3-DQ2, ambos alelos del heterodímero HLA-DQ2.5 (DQA1\*0501 and DQB1\*0201) se encuentran en el mismo cromosoma y están codificados en *cis*. En los individuos heterocigotos para los haplotipos DR5-DQ7 y DR7-DQ2, las dos moléculas se codifican en cromosomas diferentes, o en *trans* (Figura 1). Las diferencias entre ambos heterodímeros HLA-DQ2.5 afectan a un aminoácido del péptido señal de las cadenas DQ $\alpha$  (DQA1\*501 versus DQA1\*0505) y otro residuo de la región de membrana de las cadenas DQ $\beta$  (DQB1\*0201 versus DQB1\*0202) y no parece tener consecuencias funcionales, por lo que se les atribuye un riesgo similar. No obstante, el riesgo conferido por otra variante de la molécula HLA-DQ2, el dímero HLA-DQ2.2 es muy bajo (Figura 1).<sup>7,8</sup>

El grado de susceptibilidad a EC está relacionado con la abundancia del heterodímero DQ2.5. Los individuos homocigotos para el haplotipo DR3-DQ2 o heterocigotos DR3- DQ2/DR7-DQ2 expresan los niveles más elevados de heterodímeros DQ2.5 y presentan el máximo riesgo genético de desarrollar EC.<sup>8-10</sup> En este sentido, cabe destacar que los pacientes con EC refractaria, que no responde a la dieta sin gluten presentan un grado de homocigosidad DR3-DQ2 (44–62%) mayor que otros pacientes celíacos (20–24%). Un efecto de dosis alélica similar se ha sugerido para las moléculas DQ8.

Junto con los genes que codifican las moléculas DQ, la región HLA contiene otros genes que participan en la respuesta inmune y que podrían influir en la susceptibilidad a EC. Diversos estudios han postulado que polimorfismos en genes como *MICA*, *MICB* o *TNF* podrían contribuir al riesgo de desarrollar la enfermedad. No obstante, la mayoría de los estudios no han tenido en cuenta el elevado desequilibrio de ligamiento entre dichos genes y HLA-DQ y los resultados no son concluyentes. La secuenciación y el mapeo exhaustivo de la región HLA ayudarán a determinar si contiene otros factores de susceptibilidad.

A pesar de la importante contribución de los genes HLA en el riesgo genético, la concordancia de la enfermedad entre hermanos idénticos para HLA es tan solo de alrededor del 30%, por lo que se puede concluir que los genes HLA son importantes pero no suficientes para desarrollar EC.<sup>7</sup>

### 2.3. Implicación en la patogénesis

La fuerte asociación de los genes HLA de clase II con la EC se explica por el papel fundamental que juegan los linfocitos T CD4+ en la patogenésis de la enfermedad. De hecho, existen células T CD4+ que reconocen los péptidos del gluten en la mucosa intestinal de pacientes celíacos, pero no en la de individuos sanos. Estas células T CD4+ presentes en el intestino de los celíacos, están normalmente caracterizadas por las moléculas HLA-DQ2 o -DQ8.<sup>9</sup>

Cuando individuos genéticamente susceptibles (que expresan las moléculas HLA-DQ2 o -DQ8) son expuestos a ciertos epítomos del gluten, estos epítomos son presentados por las células presentadoras de antígenos, estimulando la proliferación de células T CD4+ gluten-específicas.

Un hito importante en el conocimiento de las bases moleculares de la asociación del HLA con la EC fue el descubrimiento de que la unión de las moléculas HLA-DQ2 y -DQ8 al gluten depende de que dichos péptidos hayan sido modificados enzimáticamente por una enzima denominada transglutaminasa (TG2). Esta enzima cataliza una reacción que provoca que los epítomos del

gluten adquieran mayor carga negativa, favoreciendo su unión con las moléculas HLA-DQ2 y -DQ8 y provocando la presentación de los péptidos del gluten a las células T.

Dada la importancia de las moléculas HLA en el proceso de activación de las células T autoreactivas contra el gluten, es lógico pensar que cualquier modificación en la secuencia que las codifica, pueda provocar una alteración en cualquiera de los pasos de este proceso. Así, los polimorfismos en la secuencia que codifica la parte de unión al antígeno podrían provocar un cambio en la afinidad de unión, favoreciendo el reconocimiento de los péptidos del gluten. Por otro lado, ciertos polimorfismos localizados en zonas reguladoras pueden provocar una sub-expresión o sobre-expresión de las moléculas HLA, disminuyendo o aumentando la respuesta inmune contra del gluten.

### 3. Búsqueda de genes de susceptibilidad genética en EC

Durante los últimos años, se han realizado numerosos esfuerzos para localizar e identificar genes de susceptibilidad no localizados en la región HLA que puedan explicar la genética de la EC. Para ello, se han utilizado fundamentalmente dos métodos de análisis: los estudios de ligamiento en familias y los estudios de asociación. Además, recientemente, se ha investigado la EC mediante estudios de asociación de genoma completo (*Genome Wide Association Studies, GWAS*), en los cuales se han analizado miles de polimorfismos de un único nucleótido o SNPs. Gracias a estos estudios, se han identificado varios genes asociados a la EC, pero no todas las asociaciones observadas se han confirmado en estudios posteriores.

#### 3.1. Regiones de ligamiento y genes candidato posicionales

Los estudios de ligamiento en familias permiten identificar regiones cromosómicas repetida y consistentemente heredadas por los afectados de una enfermedad en varias generaciones de una familia. Mediante este tpo de análisis, podemos acotar las regiones del genoma potencialmente implicadas en la patogénesis de las enfermedades. Los genes localizados en estas regiones constituyen genes candidato posicionales, debido a que es su ubicación la que les confiere la sospecha de estar participando en la patogénesis de la enfermedad.



Figura 2. Regiones de ligamiento replicadas en diferentes estudios familiares

Hasta la fecha, se han identificado cuatro regiones candidatas ligadas a la EC: la primera es la región HLA o CELIAC1, que constituye el componente genético más relevante en EC y que ya ha sido tratada en profundidad anteriormente. Las otras tres regiones se denominan CELIAC2, CELIAC3 y CELIAC4 (Figura 2), pero los análisis de estos *loci* no siempre han sido concluyentes y consistentes.

## CELIAC2

La región CELIAC2 está situada en el cromosoma 5q31-33 y fue identificada por primera vez por Greco y colaboradores en 1998.<sup>11</sup> La replicación de este *locus* no ha sido universal, y aún no se ha identificado ningún gen funcionalmente implicado en la enfermedad. Esta región contiene un grupo de genes que codifican para algunas citoquinas, las cuales podrían jugar un papel importante en la regulación del sistema inmune y la inflamación.<sup>12</sup> De todas maneras, aún no se ha identificado ningún gen concreto asociado con EC.

Varios estudios se han centrado en candidatos específicos como por ejemplo *IL12B* o los genes de la familia *SPINK*, pero no se han encontrado asociaciones consistentes en ninguno de ellos.<sup>13</sup> Trece variantes potencialmente funcionales de los genes *IL4*, *IL5*, *IL9*, *IL13*, *IL17B* y *NR3C1*, todas en el *locus* CELIAC2, fueron genotipadas en población irlandesa, pero ninguna de las variantes ni los haplotipos presentaron asociación con la enfermedad.<sup>14</sup>

Por otro lado, en un estudio en el que se seleccionaron genes diferencialmente expresados en la enfermedad que estaban localizados en regiones de ligamiento, se observó evidencia de asociación con el gen *YIPF5*, también localizado en esta región.<sup>15</sup> En un estudio posterior realizado en poblaciones finlandesa y húngara se confirmó el ligamiento de esta región con la EC y se observó de nuevo evidencia de asociación con el gen *YIPF5* aunque no de forma demasiado consistente.<sup>16</sup>

A pesar de ser un *locus* de riesgo importante que ha sido descrito en varios estudios de ligamiento, no se ha encontrado ningún gen que explique su asociación con la enfermedad.

## CELIAC3

CELIAC3 fue identificada por primera vez en 1999 por Holopainen y colaboradores.<sup>17</sup> Esta región está situada en la región cromosómica 2q33 y contiene, entre otros, los genes reguladores de la respuesta linfocitaria *CD28*, *CTLA4* e *ICOS*, que serán discutidos más tarde.

En este primer estudio, se utilizaron 7 marcadores genéticos diferentes que fueron analizados en un total de 100 familias. El microsatélite D2S116 presentó el valor de ligamiento no paramétrico más alto del estudio y se observó, además, que la asociación entre el marcador y la enfermedad es significativa. El ligamiento entre la EC y este *locus* se ha replicado en varios estudios posteriores utilizando diferentes marcadores genéticos (microsatélites y SNPs), además del microsatélite antes mencionado.

El *locus* CELIAC3 contiene los genes *CD28*, *CTLA4* e *ICOS* que están localizados en un bloque de alrededor de 300kb que controla diferentes aspectos de la respuesta de las células T. La unión de *CD28* e *ICOS* con sus respectivos ligandos crea una señal positiva de proliferación y activación de citoquinas, mientras que la unión de *CTLA4* crea una señal negativa que regula la activación de células T. La asociación del gen *CTLA4* con la EC se ha descrito en varias poblaciones, pero los resultados no siempre han sido positivos. Un estudio en el que se analizaron todos los SNPs de este gen sugiere que son los haplotipos, más que los SNPs, los que están más fuertemente asociados a la enfermedad. Sin embargo, se necesitan datos de la función de estas variables y/o haplotipos en la enfermedad para determinar si la asociación es con el gen *CTLA4* o con otro gen vecino.<sup>18</sup>

En el estudio mencionado anteriormente, en el que se analizaron genes diferencialmente expresados en la enfermedad localizados en regiones de ligamiento,<sup>13</sup> el gen que presentó mayor

asociación con la enfermedad fue el gen *SERPINE2*. Este gen es importante en los estadios iniciales de la formación de la matriz extracelular, proceso que está alterado en la EC. Un estudio posterior no fue capaz de replicar la asociación del gen *SERPINE2* con la enfermedad, por lo que, a pesar de los múltiples intentos, el factor genético que confiere el riesgo en la región CELIAC3 aun no ha sido identificado.<sup>19</sup>

#### *CELIAC4*

El *locus* CELIAC4 está situado en la región cromosómica 19p13.1 y fue identificado en primer lugar por van Belzen y colaboradores en el año 2003.<sup>20</sup> Esta región contiene más de 140 genes y algunos forman parte de la respuesta inmune y la inflamación.

En el estudio en el que CELIAC4 fue identificado, se analizaron 82 familias con miembros afectados. Se observó que el microsatélite D19S899 presenta un pico de ligamiento significativo con EC, con un LOD score (logaritmo de la razón de probabilidades) de 4,31. Además, este marcador genético presentó una asociación significativa con la enfermedad cuando, con la intención de confirmar los resultados obtenidos en el estudio de ligamiento, se analizaron 216 pacientes de EC y 216 controles. No obstante, no todos los estudios de replicación posteriores han conseguido resultados positivos sobre el ligamiento de esta región y la enfermedad celíaca.<sup>21</sup>

El mejor candidato de la región CELIAC4 es el gen de la miosina IXB (*MYO9B*), ya que codifica para una molécula de miosina que probablemente participa en la remodelación de actina de los enterocitos. No se conoce la función específica de *MYO9B* pero se sabe que contiene un dominio proteico similar al de los genes involucrados en las uniones estrechas o *tight junction*, por lo que se ha hipotetizado que las variaciones en este gen pueden resultar en la disrupción de la barrera intestinal que permite el paso de péptidos inmunogénicos.<sup>22</sup> Sin embargo, no todos los estudios de asociación realizados han encontrado asociación positiva con *MYO9B*. En esta región hay alrededor de 140 genes adicionales, algunos de los cuales participan en la inmunidad y la inflamación (*CYP4F3*, *HSH2D*, *IL12RB1*, *IFI30* y *KIR*, por ejemplo) y podrían ser buenos candidatos. En un estudio en el que se analizaron diez de los genes de esta región en población holandesa, se encontró evidencia de asociación con los genes *CYP4F3* y *CYP4F2*, ambos involucrados en la inhibición del leucotrieno, potente mediador de la inflamación.<sup>23</sup> El gen *ICAM-1* localizado en esta región es importante en la adhesión intercelular, también presentó asociación en población francesa.<sup>24</sup> Estas asociaciones débiles deben ser replicadas en poblaciones independientes con el fin de determinar la contribución de estos genes al desarrollo de la enfermedad.

### **3.2. Genes candidato funcionales**

#### *Genes de la respuesta inmune innata*

Cada vez es más evidente la participación del sistema inmune innato en el desarrollo de la EC, por lo que varios genes de la respuesta innata han sido estudiados en busca de polimorfismos de riesgo. En uno de los estudios se analizaron polimorfismos funcionales localizados en regiones reguladoras de diferentes mediadores proinflamatorios (IL-1alpha, IL-1beta, IL-1RN, IL-18, RANTES y MCP-1). Ninguno de los genes analizados en este estudio, excepto *RANTES*, que presentó una asociación dudosa, fue asociado con riesgo a desarrollar EC.<sup>25</sup>

La familia de genes *KIR* (*Killer Immunoglobulin-like receptors*) también ha sido estudiada por contener genes candidato de la respuesta inmune innata en EC. Estos receptores están localizados en la región 19q13.4, una de las regiones que ha presentado evidencia de ligamiento



con la enfermedad, y codifican para receptores de las células NK (*Natural Killer*) y de ciertas células T que modulan la actividad citolítica mediante las interacciones con los ligandos HLA de clase I, participando en la respuesta inmune innata. Se analizó el contenido génico, los genotipos y haplotipos de los genes KIR en población vasca y se observó que la frecuencia de la combinación KIR2DL5B(+)/KIR2DL5A(-) era significativamente mayor en los individuos enfermos. Esta asociación se replicó en población española (riesgo de odd 3,63) sugiriendo una implicación del gen *KIR2DL5B* con un aumento de riesgo a padecer EC, probablemente debido a la falta de una señal inhibitoria eficiente.<sup>26</sup> Por otro lado, en otro estudio, se observó que el gen inhibidor *KIR3DL1* estaba sobreexpresado en mucosa intestinal en la enfermedad activa, probablemente debido al aumento de subpoblaciones linfocitarias con fenotipo NK.<sup>27</sup>

Los receptores tipo *toll* (TLR), cuya función es el reconocimiento del patógeno y la estimulación de la respuesta inmunitaria, también se han analizado en busca de asociación con la enfermedad. Aunque se ha visto que su expresión está alterada en los enfermos, no se ha encontrado asociación entre los polimorfismos de estos genes y la EC. Del mismo modo, tampoco se ha encontrado asociación del número de copias (*Copy Number Variation*, CNV) de los genes *TLR2* y *TLR4* con la enfermedad.<sup>28</sup>

A su vez, las  $\beta$ -defensinas forman un *cluster* con número de copias variable en la población y forman parte de la respuesta inmune innata, actuando como antibiótico natural. Los genes que forman parte de esta familia han sido relacionados anteriormente con enfermedades autoinmunes e inflamatorias, como la psoriasis o la enfermedad de Crohn. Aunque no se ha detectado ninguna asociación entre los SNPs presentes en estos genes y la EC, sí que se ha demostrado una asociación entre el número de copias del clúster y EC, ya que se observó una menor presencia de números de copia altos (>4) en el grupo de los pacientes, sugiriendo un papel protector de las  $\beta$ -defensinas en la enfermedad.<sup>28</sup>

Como se ha comentado con anterioridad, los genes *MICA* y *MICB* que codifican para moléculas de estrés, también han sido estudiados en busca de variantes de riesgo, pero la localización de estos genes en el *locus* CELIAC1 ha dificultado sacar conclusiones sobre la contribución independiente de estos genes, debido al alto desequilibrio de ligamiento del HLA.<sup>29</sup>

Aunque se ha observado una activación del sistema inmune innato en los pacientes celíacos, ninguno de los genes candidato estudiados ha presentado una fuerte asociación con la enfermedad, por lo que se puede pensar que múltiples genes del sistema inmune innato, cada uno de ellos con un efecto débil, contribuyen al desarrollo de la enfermedad mediante la activación de la respuesta innata.

#### *Genes de la respuesta inmune adaptativa*

La respuesta Th1 es una de las principales respuestas inflamatorias de la EC y la citoquina más característica de esta respuesta es IFNG. Se ha observado que la producción de esta citoquina está significativamente aumentada en la enfermedad activa, llegando hasta niveles 240 veces mayores en situaciones de atrofia total. Se estudió el gen *IFNG* en tres cohortes holandesas y en población finlandesa pero no se encontraron diferencias entre las distribuciones alélicas de casos y controles. Hasta el momento, no hay evidencias de que *IFNG* sea un gen que predisponga a la enfermedad, a pesar de estar altamente sobreexpresado en la mucosa de los pacientes celíacos.<sup>30</sup>

Las células Th17 han sido involucradas en la patogénesis de la EC. La señalización mediante la IL23 y su receptor (IL23R) es un elemento fundamental en la diferenciación de las células T a

células Th17. El gen *IL23R* ha sido asociado a otras enfermedades autoinmunes y/o inflamatorias como psoriasis o colitis ulcerosa. El gen *IL23R* presenta una variante codificante que fue analizada en población holandesa pero que no presentó asociación significativa con la enfermedad.<sup>31</sup> Sin embargo, el análisis de esta misma variante en población española presentó un aumento del alelo menor en los individuos enfermos, de forma opuesta a lo observado en las otras enfermedades.<sup>32</sup> En un estudio posterior se encontró evidencia de ligamiento en la región del gen *IL23R* en poblaciones húngara, finlandesa e italiana, pero no se encontró asociación con los polimorfismos estudiados.<sup>33</sup> De todas formas, un estudio reciente en población española en el que se estudió la asociación de 101 SNPs repartidos por 16 genes relacionados con la respuesta Th17 entre los que se encuentra *IL23R*, señala que no existe asociación con la enfermedad.<sup>34</sup>

Por otro lado, el gen *CIITA* parece ser el mayor regulador de los genes de clase II del HLA. Este gen tiene un patrón de expresión complejo y dos de los polimorfismos localizados en su promotor se han asociado con otras enfermedades autoinmunes. Estos polimorfismos fueron analizados en busca de asociación con la EC en población española pero no se encontraron diferencias significativas entre los enfermos y los controles.<sup>35</sup> Por el contrario, el segundo GWAS muestra asociación entre EC y la región que contiene el gen *CIITA*.

Hasta la fecha ningún gen candidato de la respuesta inmune adaptativa ha sido fuertemente asociado con el riesgo a desarrollar EC.

#### *Genes implicados en la remodelación del epitelio intestinal*

Se ha descrito que la permeabilidad del epitelio intestinal está aumentada en los pacientes celíacos en respuesta a la gliadina. Esta alteración en la barrera intestinal se ha asociado a modificaciones estructurales en las uniones intercelulares. El gen *MYO9B* de la región de ligamiento CELIAC4 ha sido analizado en busca de polimorfismos asociados por su posible implicación en la remodelación del epitelio intestinal.<sup>23</sup> En un estudio realizado en el año 2008 se analizaron 197 SNPs de 41 genes implicados en la comunicación intercelular en población holandesa y británica. Dos de los genes analizados, *PARD3* (2 SNPs) y *MAGI2* (2 SNPs) demostraron asociación débil con la enfermedad en población holandesa. La replicación en población británica dio como resultado asociación en uno de los SNPs de *PARD3*. El análisis conjunto de ambas poblaciones corroboró la asociación para ambos genes con valores de riesgo de odds 1,23 para *PARD3* y 1,19 para *MAGI2*. Estos genes también dieron asociación positiva con colitis ulcerosa, sugiriendo un defecto etiológico común en la barrera intestinal en ambas enfermedades.<sup>36</sup>

#### *Rutas de señalización celular*

Varias rutas de señalización están alteradas en la EC, entre ellas la ruta de señalización Jak-Stat, la ruta de señalización del factor de transcripción kappa B (NFkB), la ruta de señalización MAPK o la ruta de señalización del factor de crecimiento transformante beta (TGFB).<sup>15</sup> Varios genes de estas rutas han sido analizados en busca de asociación con la enfermedad.

Uno de los genes es el gen *STAT1*, cuya expresión está alterada en la enfermedad y además es también un candidato posicional ya que está localizado en el *locus* CELIAC3. Se analizaron cinco polimorfismos *tag* que cubren todo el gen en población holandesa, pero no se encontró ningún indicio de asociación de los polimorfismos con la enfermedad.<sup>37</sup>

El gen *NFKB1* también ha sido estudiado en busca de asociación genética con la EC, pero a pesar de que este factor de transcripción está constitutivamente activo en la mucosa de los enfermos, no se han encontrado polimorfismos que expliquen su aumento de actividad en la enfermedad. Se ha sugerido que los efectos patogénicos atribuidos a este factor de transcripción pueden estar causados por un defecto regulador en lugar de por un polimorfismo en el factor de transcripción mismo. Genes localizados aguas arriba en la cascada biológica pueden ser los responsables del riesgo genético provocando mayor actividad transcripcional dependiente de NFκB. Dos de los genes identificados en un estudio de seguimiento de GWAS (*REL* y *TNFAIP3*) están localizados en esta cascada, pudiendo ser responsables de su desregulación. Recientemente un polimorfismo regulador del gen *UBD* implicado en la activación de NFκB, ha sido asociado a la enfermedad en un estudio realizado en población española. Este gen está sobreexpresado en los enfermos activos y la distribución alélica del polimorfismo asociado presenta una correlación significativa con los niveles de expresión génica.<sup>38</sup>

Las modificaciones observadas en estas rutas biológicas complejas pueden alterar la expresión de genes que están localizados más abajo en la ruta, por lo que el análisis de genes individuales puede dar lugar a error. Un análisis exhaustivo de estas rutas puede ser crucial para la selección de candidatos para estudios de asociación.

#### *Matriz extracelular*

La matriz extracelular se encuentra degradada en el epitelio intestinal de los pacientes celíacos. Las metaloproteinasas son enzimas degradadoras de los componentes de la matriz, y se ha visto que su expresión está incrementada en los estadios activos de la enfermedad, contribuyendo a las alteraciones morfológicas de la mucosa intestinal. Por ello, estos genes han sido estudiados en diversas ocasiones en busca de variantes de susceptibilidad. De todas maneras, polimorfismos funcionales del gen *MMP-1* no han podido ser asociados a la enfermedad.<sup>39</sup>

## **4. Estudios de asociación de genoma completo en enfermedad celíaca**

Los GWAS, permiten un rápido escaneo de marcadores en completos sets de DNA, o genomas de muchos individuos, con el objetivo de encontrar variaciones genéticas asociadas a una enfermedad en particular. Una vez identificadas estas asociaciones genéticas, los investigadores pueden usar esta información para desarrollar nuevas y mejores tecnologías para detectar, tratar y prevenir enfermedades. Estos estudios son especialmente útiles a la hora de encontrar variaciones genéticas que contribuyen al desarrollo de enfermedades comunes y complejas, como podrían ser asma, cáncer, diabetes, y en este caso la EC.

Para realizar un estudio de asociación de genoma completo, los investigadores utilizan dos grupos de participantes: individuos que padecen la enfermedad a estudio e individuos con características similares a los anteriores pero que no padecen dicha enfermedad. Se trata de un estudio de asociación a escala genómica.

El ADN completo o genoma de cada individuo es purificado a partir de una muestra de sangre. Este ADN se coloca en chips y es escaneado automáticamente en el laboratorio. Estos aparatos inspeccionan la muestra de forma estratégica en busca de marcadores de variación genética, tratándose en este caso de SNPs.

Si se descubre que ciertas variaciones genéticas resultan ser significativamente más frecuentes en individuos enfermos que en sanos, se dice que estas variaciones están asociadas a la enfermedad. Estas variaciones genéticas asociadas pueden ser marcadores importantes, ya que podrían indicarnos una región del genoma humano donde podría residir la variación causante de la enfermedad. La variante asociada en sí no tendría por qué ser la causa directa de la enfermedad, podría simplemente estar apuntando a la región en la que seguir buscando la verdadera variante causal. Por esta razón, en la mayoría de los casos se necesita seguir adelante en la investigación, como por ejemplo secuenciando el ADN en esa región en particular, con el objetivo de identificar el cambio genético exacto involucrado en la enfermedad, o haciendo estudios funcionales para encontrar asociación entre las variantes concretas y los niveles de expresión génica.

Los GWAS representan un método para capturar una nueva clase de variantes genéticas asociadas a las enfermedades. Los estudios de asociación basados en pedigríes utilizan familias en las que los *clústers* asociados a la enfermedad son útiles para identificar variantes raras con un gran efecto. Mientras tanto, los GWAS dependen de muestras basadas en poblaciones, por lo que requieren variantes comunes con un efecto más modesto (ya que las variantes raras no se podrán observar), que no se podría observar utilizando un enfoque tradicional basado en el ligamiento.

#### 4.1. Resultados del primer GWAS

En el primer estudio de genoma completo realizado en EC, se estudiaron 778 individuos enfermos de EC y 1.422 controles sanos. Se realizaron análisis de asociación en 310.605 SNPs que demostraron tener una frecuencia mayor de 1% para el alelo menor.<sup>40</sup>

Como era de esperar, la mayor asociación se encontró en torno al *locus* del HLA. El alelo rs2187668-A demostró ser un eficiente marcador para el HLA-DQ2.5*cis*. Este es el haplotipo HLA-DQ2 más común asociado a la EC. En este primer estudio se vio que en el 89,2% de los enfermos del Reino Unido estaban presentes una o dos copias del HLA-DQ2.5*cis*, frente al 25,5% de presencia en la población control.

Fuera de la región HLA se observó un mayor número de SNPs asociados de lo que se podría esperar por azar, 56 SNPs presentaban una asociación con una  $p < 10^{-4}$ . Algunos de estos SNPs se encuentran próximos entre sí, lo que sugiere que el exceso de SNPs con valores de  $p$  bajos podría ocurrir debido a una verdadera asociación de los SNPs que se encuentran en desequilibrio de ligamiento con variantes causales de la enfermedad.

El único SNP fuera del HLA en demostrar asociación significativa fue rs13119723, en la región 4q27, localizada en un bloque de desequilibrio de ligamiento que contiene los genes *IL2* e *IL21*. Estos resultados se repitieron en colecciones de pacientes y controles holandeses e irlandeses.

Se estimó que la región *IL2-IL21* tan solo podría explicar un 1% del incremento de riesgo familiar para EC, lo que sugiere la existencia de otros genes de susceptibilidad aún no identificados. Por esta razón, se procedió al estudio de los 1.164 SNPs más significativos del primer estudio, en otros 1.643 casos celíacos y 3.406 controles sanos de tres colecciones europeas independientes<sup>41</sup>. Las regiones asociadas tras este nuevo estudio se analizaron en busca de genes candidato que pudieran cumplir alguna función en el desarrollo de la EC, fundamentalmente aquellos genes implicados de alguna forma en la respuesta del sistema inmune (Figura 3).

Es muy importante la replicación de los resultados de descubrimientos genéticos en diferentes poblaciones a la hora de establecer un efecto genético en la predisposición a padecer una enfermedad. Por esta razón, los resultados obtenidos en el primer GWAS se han intentado replicar en varias poblaciones independientes, obteniendo resultados diversos, debidos posiblemente a variaciones poblacionales y al tamaño muestral de cada uno de estos estudios.

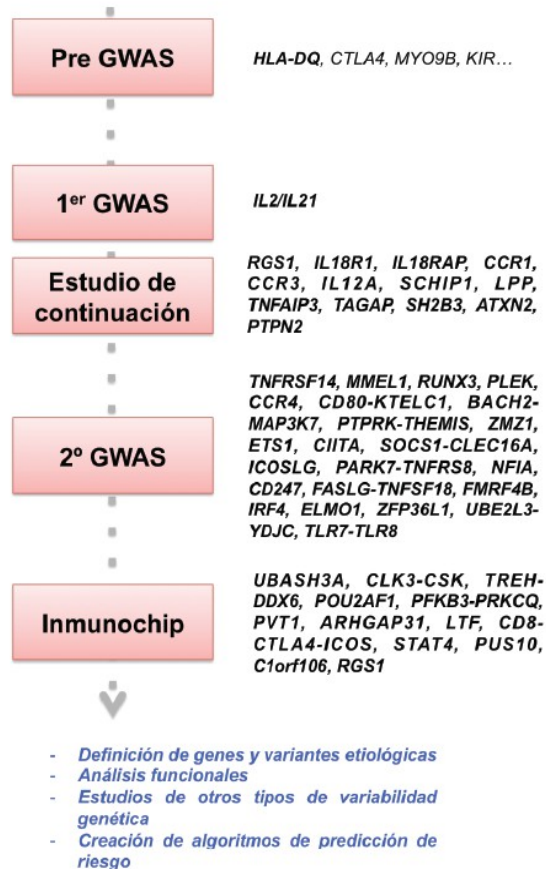


Figura 3. Avances en la Genética de la EC. Tras el estudio del Immunochip, son 40 los loci identificados que contribuyen al riesgo de EC. Ahora es el momento de abordar los estudios funcionales para descubrir las variantes etiológicas y determinar las aplicaciones prácticas de los resultados de asociación (en azul).

#### 4.2. Resultados segundo GWAS

El segundo GWAS en EC se realizó en el año 2009. Para ello se analizaron 292.387 SNPs fuera de la región HLA en muestras de ADN de 4.533 individuos celíacos y 10.750 controles sanos de

origen Europeo. Además, se analizaron también 231.362 SNPs adicionales fuera del HLA en 3.796 individuos enfermos de celiaquía y 8.154 controles sanos.<sup>42</sup>

Se identificaron 13 nuevas regiones de riesgo con evidencias significativas de asociación (Figura 3). En estas regiones se encuentran varios genes con funciones inmunológicas: *BACH2*, *CCR4*, *CD80*, *CIITA-SOCS1-CLEC16A*, *ETS1*, *ICOSLG*, *RUNX3*, *THEMIS*, *TNFRSF14* y *ZMIZ1*. Otras 13 regiones no llegaron a la asociación significativa pero si apuntaron una tendencia. Estas regiones también contienen genes con funciones inmunológicas, incluyendo *CD247*, *FASLG-TNFSF18-TNFSF4*, *IRF4*, *TLR7-TLR8*, *TNFRSF9* e *YDJC*.

### 4.3. ImmunoChip

El último proyecto llevado a cabo con el objetivo de identificar un mayor número de variantes asociadas a la EC y otras enfermedades autoinmunes ha sido el denominado proyecto ImmunoChip. En lo que respecta a la EC se han analizado más de 200.000 variantes en 12.000 pacientes celíacos y 12.000 controles, en muestras que provienen de 7 regiones geográficas.<sup>43</sup>

Se analizaron 183 *loci* relacionados con el sistema inmune y que se encuentran fuera de la región HLA, y 36 mostraron una asociación significativa con EC, las 26 regiones identificadas en los GWAS más 13 nuevos *loci*. Todas las variantes asociadas tienen una frecuencia del alelo menor superior al 5%, es decir, todas ellas son variantes comunes. Tan solo se han detectado variantes de baja frecuencia asociadas a la enfermedad en 4 *loci*. La ventaja del ImmunoChip respecto a los GWAS consiste en la posibilidad de realizar un mapeo fino (*fine-mapping*) que nos permita localizar e identificar señales causales, ya que el genotipado en el ImmunoChip es mucho más denso. Una muestra de ello es que de las 54 señales independientes fuera del HLA que se encuentran en los 36 *loci* genotipados en alta densidad, 29 localizan en torno a un único gen (Figura 3).

Tras la anotación funcional de las regiones asociadas, una de las conclusiones principales a las que se ha llegado ha sido que existen muy pocos marcadores en regiones codificantes de los genes, si bien algunos marcadores se encuentran cerca de lugares de inicio de transcripción, y otros en regiones 3'UTR.

Algunos de los posibles genes causales propuestos, por poseer señales cerca de regiones reguladoras en 5' o 3' son: *THEMIS/PTPRK*, *TAGAP*, *ETS1*, *RUNX3* y *RGS1*. Algunos de ellos ya habían sido propuestos previamente tras los GWAS.

### 4.4. Replicación de los estudios de asociación y análisis funcionales de los genes candidato

En 2011, se replicaron los ocho picos de asociación más significativos del primer GWAS en EC en población española, identificando cuatro genes (*IL12A*, *LPP*, *SCHIP1* y *SH2B3*) cuya expresión en la mucosa intestinal variaba según el estatus de la enfermedad y el genotipo de la variante asociada.<sup>44</sup> Estos resultados sugieren que estos genes podrían estar alterados de forma constitutiva en los celíacos, probablemente desde antes de padecer los síntomas observables de la patología, y que, por tanto, podrían tener un papel primario en la patogénesis de la enfermedad.

Un trabajo aún sin publicar da un paso adelante en la materia, ya que identifica dos genes (*RUNX3* y *THEMIS*), situados en la misma región asociada, que se coexpresan tanto en la

enfermedad activa como en respuesta a la estimulación *in vitro* por medio de gliadina de biopsias intestinales de celíacos inactivos que han seguido la dieta durante al menos dos años. Por lo tanto, parece que las variantes asociadas en esta región influyen en la expresión de diferentes genes, pero no de forma constitutiva, desde el nacimiento del futuro celíaco, sino tras el estímulo tóxico que desencadena la respuesta inmune.

Las implicaciones de este hallazgo son de gran importancia, ya que señalan la existencia de mecanismos reguladores comunes para distintos genes en la secuencia de ADN que sólo tienen efecto ante el estímulo inmunogénico que provoca la enfermedad.

Estos y otros trabajos subrayan la necesidad de realizar estudios funcionales y de evitar la elección de hipotéticos genes de susceptibilidad por medio de criterios arbitrarios. Del mismo modo, revelan que queda mucho por descubrir sobre la inmensa complejidad reguladora del genoma y abre las puertas al análisis exhaustivo de las variantes del genoma no codificante, al estudio de las moléculas de ARN no mensajero y al de los niveles de expresión de sus dianas situadas en *trans*, en posiciones alejadas del genoma.

#### **4.5. Conclusiones**

A pesar de los enormes esfuerzos de las últimas décadas, los mecanismos genético-moleculares que subyacen a esta enfermedad aún no se han explicado completamente. Los GWAS y estudios posteriores han comenzado a desentrañar la contribución de la genética a la patogénesis de la EC. A pesar de que desde el punto de vista genético, las enfermedades de etiología inmune muestran amplias diferencias en cuanto al número de *loci* implicados, el efecto de cada uno de éstos y los factores ambientales implicados, es cierto que existe un marcado solapamiento entre este tipo de patologías. Este solapamiento debe implicar la participación de rutas biológicas comunes y sugiere que las estrategias para su tratamiento también podrían ser compartidas. No obstante, la interpretación de los estudios de asociación debe realizarse con cautela ya que es cierto que cada uno de los *loci* identificados contiene más de un gen. Las estrategias para identificar las posibles variantes etilógicas se señalan en la figura 3, y podrán en un futuro identificar las alteraciones funcionales que subyacen a las enfermedades autoinmunes. Con el tiempo estas variantes patogénicas podrán ser incluidas en algoritmos de predicción de riesgo y permitirán un diagnóstico de individuos con una alta predisposición genética antes de la aparición de los síntomas, lo que redundará en una mejora en su calidad de vida y en una disminución de los costes sanitarios. Además se podrán abrir las puertas a nuevas dianas terapéuticas para la propia EC y para otras enfermedades de etiología autoinmune.

## Referencias

1. Sollid LM. Thorsby E. *HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis*. Gastroenterol. 1993; 105: 910-22.
2. Greco L. Romino R. Coto I. et al. *The first large population based twin study of coeliac disease*. Gut. 2002; 50: 624-8. <http://dx.doi.org/10.1136/gut.50.5.624>
3. Horton R. Wilming L. Rand V. Lovering RC. Bruford EA. Khodiyar VK. et al. *Gene map of the extended human MHC*. Nat Rev Genet. 2004; 5: 889-99. <http://dx.doi.org/10.1038/nrg1489>
4. 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. 2003; 64: 469-77. [http://dx.doi.org/10.1016/S0198-8859\(03\)00027-2](http://dx.doi.org/10.1016/S0198-8859(03)00027-2)
5. Spurkland A. Sollid LM. Polanco I. Vartdal F. Thorsby E. *HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7*. Hum Immunol. 1992; 35: 188-92. [http://dx.doi.org/10.1016/0198-8859\(92\)90104-U](http://dx.doi.org/10.1016/0198-8859(92)90104-U)
6. Sollid LM. Thorsby E. *Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer*. J Exp Med. 1989; 169: 345-50. <http://dx.doi.org/10.1084/jem.169.1.345>
7. Sollid LM. *Coeliac disease: dissecting a complex inflammatory disorder*. Nat Rev Immunol. 2002; 2: 647-55. <http://dx.doi.org/10.1038/nri885>
8. Van Belzen MJ. Koeleman BP. Crusius JB. et al. *Defining the contribution of the HLA region to cis DQ2-positive coeliac disease patients*. Genes Immun. 2004; 5: 215-20. <http://dx.doi.org/10.1038/sj.gene.6364061>
9. Ploski R. et al. *HLA-DQ (alpha 1\*0501, beta 1\*0201) associated susceptibility in celiac disease: a possible gene dosage effect of DQB1\*0201*. Tissue Antigens 1993; 41: 173-7. <http://dx.doi.org/10.1111/j.1399-0039.1993.tb01998.x>
10. Lundin KE. Scott, H. Hansen T. Paulsen G. Halstensen TS. Fausa O. et al. *Gliadin-specific, HLA-DQ (alpha 1\*0501, beta 1\*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients*. J Exp Med. 1993; 178: 187-96. <http://dx.doi.org/10.1084/jem.178.1.187>
11. Greco L. Corazza GR. Babron MC. Clot F. Fulchignoni-Lataud MCV. Percopo S. et al. *Genome search in celiac disease*. Am J Hum Genet. 1998; 62: 35-41. <http://dx.doi.org/10.1086/301754>
12. Greco L. Babron MC. Corazza GR. Percopo S. Sica R. Clot F. et al. *Existence of a genetic risk factor on chromosome 5q in Italian coeliac disease families*. Ann Hum Genet. 2001; 65: 35-41. <http://dx.doi.org/10.1046/j.1469-1809.2001.6510035.x>
13. Seegers D. Borm ME. Van Belzen MJ. et al. *IL12B and IRF1 gene polymorphisms and susceptibility to celiac disease*. Eur J Immunogenet. 2003; 30: 421-5. <http://dx.doi.org/10.1111/j.1365-2370.2003.00428.x>
14. Ryan AW. Thornton JM. Brophy K. Daly JS. McLoughlin RM. O'Morain C. et al. *Chromosome 5q candidate genes in coeliac disease: genetic variation at IL4, IL5, IL9, IL13, IL17B and NR3C1*. Tissue Antigens. 2005; 65: 150-5. <http://dx.doi.org/10.1111/j.1399-0039.2005.00354.x>
15. Castellanos-Rubio A. *Combined functional and positional gene information for the identification of susceptibility variants in celiac disease*. Gastroenterol. 2008; 134: 738-46. <http://dx.doi.org/10.1053/j.gastro.2007.11.041>



16. Koskinen LL. Einarsdottir E. Korponay-Szabo IR. et al. *Fine mapping of the CELIAC2 locus on chromosome 5q31-q33 in the Finnish and Hungarian populations*. Tissue Antigens. 2009; 74: 408-16. <http://dx.doi.org/10.1111/j.1399-0039.2009.01359.x>
17. Holopainen P. Naluai AT. Moodie S. Percopo S. Coto I. Clot F. et al. *Candidate gene region 2q33 in European families with coeliac disease*. Tissue Antigens. 2004; 63: 212-22. <http://dx.doi.org/10.1111/j.1399-0039.2004.00189.x>
18. Brophy K. Ryan AW. Thornton JM. et al. *Haplotypes in the CTLA4 region are associated with coeliac disease in the Irish population*. Genes Immun. 2006; 7: 19-26. <http://dx.doi.org/10.1038/sj.gene.6364265>
19. Dema B. Martínez A. Fernández-Arquero M. Maluenda C. Polanco I. De la Concha EG. et al. *Lack of replication of celiac disease risk variants reported in a Spanish population using an independent Spanish sample*. Genes Immun. 2009; 10: 659-61. <http://dx.doi.org/10.1038/gene.2009.54>
20. Van Belzen MJ. Meijer JWR. Sandkuijl LA. Bardoel AFJ. Mulder CJJ. et al. *A major non-HLA locus in celiac disease maps to chromosome 19*. Gastroenterol. 2003; 125: 1032-41. [http://dx.doi.org/10.1016/S0016-5085\(03\)01205-8](http://dx.doi.org/10.1016/S0016-5085(03)01205-8)
21. Capilla A. Donat E. Planelles D. Espinós C. Ribes-Koninckx C. Palau F. *Genetic analyses of celiac disease in a Spanish population confirm association with CELIAC3 but not with CELIAC4*. Tissue Antigens. 2007; 70: 324-9. <http://dx.doi.org/10.1111/j.1399-0039.2007.00899.x>
22. Monsuur AJ. De Bakker PIW. Alizadeh BZ. Xhernakova A. Bevova MR. Strengman E. et al. *Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect*. Nat Genet. 2005; 37: 1341-4. <http://dx.doi.org/10.1038/ng1680>
23. Curley CR. Monsuur AJ. Wapenaar MC. Rioux JD. Wijmenga C. *A functional candidate screen for coeliac disease genes*. Eur J Hum Genet. 2006; 14: 1215-22. <http://dx.doi.org/10.1038/sj.ejhg.5201687>
24. Abel M. Cellier C. Kumar N. Cerf-Bensussan N. Schmitz J. Caillat-Zucman S. *Adulthood-onset celiac disease is associated with intercellular adhesion molecule-1 (ICAM-1) gene polymorphism*. Hum Immunol. 2006; 67: 612-7. <http://dx.doi.org/10.1016/j.humimm.2006.04.011>
25. Rueda B. Zhernakova A. López-Nevot MA. Martín J. Koeleman BPC. *Association study of functional genetic variants of innate immunity related genes in celiac disease*. BMC Med Genet. 2005; 3: 6-29. <http://dx.doi.org/10.1186/1471-2350-6-29>
26. Santin I. Castellanos-Rubio A. Perez de Nanclares G. Vitoria JC. Castaño L. Bilbao JR. *Association of KIR2DL5B gene with celiac disease supports the susceptibility locus on 19q13.4*. Genes Immun. 2007; 8: 171-6. <http://dx.doi.org/10.1038/sj.gene.6364367>
27. Fernandez-Jimenez N. et al. *Upregulation of KIR3DL1 gene expression in intestinal mucosa in active celiac disease*. Hum Immunol. 2011; 72: 617-20. <http://dx.doi.org/10.1016/j.humimm.2011.04.008>
28. Fernandez-Jimenez N. Santín I. Irastorza I. Plaza-Izurietta L. Castellanos-Rubio A. Vitoria JC. Bilbao JR. *Analysis of beta-defensin and Toll-like receptor gene copy number variation in celiac disease*. Hum Immunol. 2010; 71: 833-6. <http://dx.doi.org/10.1016/j.humimm.2010.05.012>
29. Martín-Pagola A. Pérez-Nanclares G. Ortiz L. Vitoria JC. Hualde I. Zaballa R. et al. *MICA response to gliadin in intestinal mucosa from celiac patients*. Immunogenetics. 2004; 56: 549-54. <http://dx.doi.org/10.1007/s00251-004-0724-8>

30. Wapenaar MC. Van Belzen MJ. Fransen JH. Fariña Sarasqueta A. Houwen RHJ. Meijer JWR. et al. *The interferon gamma gene in celiac disease: augmented expression correlates with tissue damage but no evidence for genetic susceptibility*. J. Autoimmun. 2004; 23: 183-90. <http://dx.doi.org/10.1016/j.jaut.2004.05.004>
31. Weersma RK. Zhernakova A. Nolte IM. Lefebvre C. Rioux JD. Mulder F. et al. *ATG16L1 and IL23R are associated with inflammatory bowel diseases but not with celiac disease in the Netherlands*. Am J Gastroenterol. 2008; 103: 621-7. <http://dx.doi.org/10.1111/j.1572-0241.2007.01660.x>
32. Núñez C. Dema B. Cénit MC. Polanco I. Maluenda C. Arroyo R. et al. *IL23R: a susceptibility locus for celiac disease and multiple sclerosis?* Genes Immun. 2008; 9: 289-93. <http://dx.doi.org/10.1038/gene.2008.16>
33. Einarsdottir E. Koskinen LLE. Dukes E. Kainu K. Suomela S. Lappalainen M. et al. *IL23R in the Swedish, Finnish, Hungarian and Italian populations: association with IBD and psoriasis, and linkage to celiac disease*. BMC Med Genet. 2009; 28: 10-8. <http://dx.doi.org/10.1186/1471-2350-10-8>
34. Medrano LM. García-Magariños M. Dema G. Espino L. Polanco I. Figueredo MA. et al. *Th17-related genes and celiac disease susceptibility*. PLoS One. 2012; 7: e31244. <http://dx.doi.org/10.1371/journal.pone.0031244>
35. Dema B. Martínez A. Fernández-Arquero M. Maluenda C. Polanco I. Figueredo MA. et al. *Autoimmune disease association signals in CIITA and KIAA0350 are not involved in celiac disease susceptibility*. Tissue Antigens. 2009; 73: 326-9. <http://dx.doi.org/10.1111/j.1399-0039.2009.01216.x>
36. Wapenaar MC. Monsuur AJ. Van Bodegraven AA. Weersma RK. Bevova MR. Linskens RK. et al. *Associations with tight junction genes PARD3 and MAGI2 in Dutch patients point to a common barrier defect for coeliac disease and ulcerative colitis*. Gut. 2008; 57: 463-7. <http://dx.doi.org/10.1136/gut.2007.133132>
37. Diosdado B. Monsuur AJ. Mearin ML. Mulder C. Wijmenga C. *The downstream modulator of interferon-gamma, STAT1 is not genetically associated to the Dutch coeliac disease population*. Eur J Hum Genet. 2006; 14: 1120-4. <http://dx.doi.org/10.1038/sj.ejhg.5201667>
38. Castellanos-Rubio A. Santin I. Irastorza I. Sanchez-Valverde F. Castaño L. Vitoria JC. et al. *A regulatory single nucleotide polymorphism in the ubiquitin D gene associated with celiac disease*. Hum Immunol. 2010; 71: 96-9. <http://dx.doi.org/10.1016/j.humimm.2009.09.359>
39. Ciccocioppo R. Di Sabatino A. Bauer M. Della Riccia DN. Bizzini F. Biagi F. et al. *Matrix metalloproteinase pattern in celiac duodenal mucosa*. Lab. Invest. 2005; 85: 397-407. <http://dx.doi.org/10.1038/labinvest.3700225>
40. Van Heel DA. et al. *A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21*. Nat Genet. 2008; 39: 827-9. <http://dx.doi.org/10.1038/ng2058>
41. Hunt KA. Zhernakova A. Turner G. Heap GAR. Franke L. Bruinenberg M. et al. *Newly identified genetic risk variants for celiac disease related to the immune response*. Nat Genet. 2008; 40: 395-402. <http://dx.doi.org/10.1038/ng.102>
42. Dubois PC. Trynka G. Franke L. Hunt KA. Romanos J. Curtotti A. et al. *Multiple common variants for celiac disease influencing immune gene expression*. Nat Genet. 2010; 42: 295-302. <http://dx.doi.org/10.1038/ng.543>

43. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. *Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease*. Nat Genet. 2011; 43: 1193-201. <http://dx.doi.org/10.1038/ng.998>
44. Plaza-Izurieta L, Castellanos-Rubio A, Irastorza I, Fernandez-Jimenez N, Gutierrez G, Bilbao JR. *Revisiting genome wide association studies (GWAS) in coeliac disease: replication study in Spanish population and expression analysis of candidate genes*. J Med Genet. 2011; 48: 493-6. <http://dx.doi.org/10.1136/jmg.2011.089714>