

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

GWAS GENE HAUTAGAIEN INPLIKAZIO FUNTZIONALA ERITASUN ZELIAKOAREN PATOGENESIAN

DOCTOREGO TESIA

Leticia Plaza Izurieta

2016

DOKTOREGO TESIA

**GWAS GENE HAUTAGAIEN INPLIKAZIO
FUNTZIONALA ERITASUN ZELIAKOAREN
PATOGENESIAN**

Leticia Plaza Izurieta

2016

Zuzendariak:

Jose Ramón Bilbao Catalá

Juan Carlos Vitoria Cormenzana

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

Laburdurak.....	1
Argitalpen originalen zerrenda	3
Proiektuaren justifikazioa eta garrantzia.....	5
Sarrera.....	7
1. Eritasun zeliakoa	9
1.1. Ezaugarri klinikoak eta diagnostika.....	9
1.2. Epidemiologia.....	10
2. Eritasun zeliakoaren patogenesisia	11
2.1. Glutena	12
2.2. Transglutaminasa	13
2.3. Sistema immune adaptatiboa.....	14
2.4. Sistema immune innatoa	14
3. Eritasun zeliakoaren genetika.....	15
3.1. HLA eremua eta eritasun zeliakoa	16
3.1.1. HLA eremua	16
3.1.2. Suszeptibilitate geneak eta arrisku genetikoa.....	16
3.1.3. HLA-ren rola eritasun zeliakoaren patogenesisian.....	18
3.2. Genoma osoko asoziazio azterketak CDn	19
3.2.1. Lehenengo GWASaren emaitzak.....	21
3.2.2. Bigarren GWASaren emaitzak	22
3.2.3. Immunochip proiektua.....	23
3.2.4. GWAS azterketen jarraipena	24
3.2.5. GWAS emaitzen azterketa funtzionalak	25
Helburuak	27
Materialak eta Metodoak	31
1. Pazienteak.....	33
2. Etikari dagokion adierazpena	33
3. SNPen genotipaketa.....	34
3.1. DNA laginak	34
3.1.1. GWAS erreplikazio azterketa	34

3.1.2.	Immunochipa	34
3.2.	DNA erauzketa	35
3.2.1.	GWAS erreplikazio azterketa	35
3.2.2.	Immunochipa	35
3.3.	Genoma osoaren anplifikazioa	36
3.4.	SNPen aukeraketa	36
3.4.1.	GWAS erreplikazio azterketa	36
3.4.2.	Immunochipa	36
3.5.	SNPen genotipaketa	37
3.5.1.	GWAS erreplikazio azterketa	37
3.5.2.	Immunochipa	38
3.6.	Datuen analisia	38
3.6.1.	SNP genotipoak	38
3.6.2.	Immunochiparen analisi estatistikoak	39
4.	Gene hautagaien analisi funtzionalak	39
4.1.	Biopsiak	39
4.1.1.	GWAS erreplikazio azterketak	39
4.1.2.	Immunochipa	39
4.2.	Biopsietako zelula populazioak	40
4.3.	RNA erauzketa	41
4.3.1.	GWAS erreplikazioa azterketa	41
4.3.2.	Immunochipa	41
4.4.	Gene adierazpena: RT-PCR	42
4.4.1.	GWAS erreplikazioa azterketa	42
4.4.2.	Immunochipa	43
4.5.	Datuen analisia eta estatistika	45
4.5.1.	GWAS erreplikazioa azterketa	45
4.5.2.	Immunochipa	45
4.6.	Koadierazpenaren analisia	45
4.7.	Genotipo-fenotipo korrelazioa	45

1 kapitulua:	47
1. Aurrekariak	49
2. Metodoak	51
3. Emaitzak eta eztabaida	53
2 kapitulua:	59
1. Aurrekariak	61
2. Materiala eta metodoak	65
2.1. Pazienteak eta biopsiak	65
2.2. RNA laginak eta gene adierazpena	65
2.3. SNPen genotipaketa	66
2.4. Koadierazpen analisiak	66
3. Emaitzak	67
3.1. CDn derberdin adierazten diren geneak	67
3.2. Genotipoaren efektua geneen adierazpenean	69
3.3. Koadierazpen gene-patroiak CDn	73
4. Eztabaida	76
3 kapitulua:	79
1. Aurrekariak	81
2. Materiala eta metodoak	83
2.1. Pazienteak eta laginak	83
2.2. Zelula populazioen bereizketa hesteetako biopsietatik	83
2.3. Odol-periferikoko zelula mononuklearren isolamendua	84
2.4. RNA laginak eta gene adierazpena	84
2.5. Asoziazio azterketa	84
3. Emaitzak	85
4. Eztabaida	95
Azken hausnarketak	97
Ondorioak	103
Material osagarria	107
Bibliografia	115

Laburdurak

AIDS	(acquired immune deficiency syndrome) Hartutako defizientzia immunitarioaren sindromea
APC	(antigen presenting cell) zelula antigeno aurkezlea
CD	(Celiac disease) eritasun zeliakoa
CEGEC	(Spanish Consortium for Genetics of Celiac Disease) Eritasun zeliakoaren azterketa genetikoentzako espainiar kontsoztzia
EGFR	(epidermal growth factor receptor) Hazkunde faktore epidermikoaren hartzailea
EMA	(anti-endomysium autoantibodies) Anti-endomisio autoantigorputzak
EpCAM	(Epithelial cell adhesion molecule) Zelula epitelialen atxikitze molekula
ESPGHAN	(European Society for Pediatric Gastroenterology, Hepatology and Nutrition) Gastroenterologia, Hepatologia eta Elikadura Pediatrikoaren Elkarte Europarra
FBS	(Fetal bovine serum) Behi-serum fetala
GFD	(gluten free diet) Glutenik gabeko dieta
GWAS	(Genome Wide Association Studies) Genoma Osoko Asoziazio Azterketak
HLA	(Human Leucocyte Antigen) Giza Antigeno Leukozitarioa
IELs	(Intraepithelial lymphocytes) Linfozito intraepiteliala
KIR	(Killer Immunoglobulin-like receptor) Immunoglobulina bezalako hartzaile hiltzailea

LD	(Linkage disequilibrium) Lotura desoreka
MHC	(Major Histocompatibility Complex) Histoconpatibilitate Konplexu Nagusia
NFkB	(Nuclear kappa B transcription factor) Nukleoko Kappa B transkripzio faktorea
NK	Natural killer
OR	Odds ratio
PCR	(Polymerase chain reaction) Polimerasaren kate-erreakzioa
RA	(rheumatoid arthritis) Artritis erreumatoide
RPLPO	(large ribosomal protein) erribosomako proteina handia
RT-PCR	(Real-time quantitative reverse transcription PCR)
SD	(standard deviation) desbideraketa estandarra
SNP	(Single Nucleotide Polymorphism) Nukleotido bakarreko Polimosfismoa
T1D	(type 1 diabetes) 1 motako diabetesa
TF	(transcription factor) transkripzioa faktorea
TG2	(transglutaminase) transglutaminasa
TGA	(transglutaminase autoantibodies) transglutaminasa autoantigorputzak
TLR	(Toll-like receptor) Toll bezalako hartzailea
WGA	(Whole genome amplification) Genoma osoaren anplifikazioa

Argitalpen originalen zerrenda

1. kapitulua

Plaza-Izurieta L, Castellanos-Rubio A, Irastorza I, Fernandez-Jimenez N, Gutierrez G, CEGEC, Bilbao JR. *Revisiting genome wide associaton studies in coeliac disease: replication study in Spanish population and expression analysis of candidate genes*. Journal of medical genetics. 2011, 48:493-496.

2. kapitulua

Plaza-Izurieta L, Fernandez-Jimenez N, Irastorza I, Jauregui-Miguel A, Romero-Garmendia I, Vitoria JC, Bilbao JR. *Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in coeliac disease*. European Journal of Human Genetics. 2015, 23(8):1100-5.

3. kapitulua

Plaza-Izurieta L, Jauregui-Miguel A., Romero-Garmendia I, Garcia-Etxebarria K., Legarda M., Irastorza I., Bilbao JR. *ImmunoChip candidate genes study in CD intestinal cell populations*. (prestatzen)

Proiektuaren justifikazioa eta garrantzia

Eritasun zeliakoa (CD), prebalentzia altuko gaixotasun immunitario kronikoa da. Eritasun honen desagerpenerako ezinbestekoa izango da prebentzioa, eta helburu horrekin, diagnostiko goiztiarra baimenduko duten predikziorako mekanismo eraginkorrak garatu behar dira. Eskala tenporalean, sintoma klinikoen agerpenak eritasunaren progresio prozesuaren egoera aurreratua kontsideratu daiteke. Eritasun aktiboaren fase honen aurretik, predispozizio genetikoa duten indibiduotan erantzun immunitarioak eragindako heste-ehunaren suntzipenaren adierazle izango liratekeen markatzaileen presentzia detektatu daiteke. Honen adibide izan daitezke tTGren (ehunetako transglutaminasa) aurkako autoantigorputz zirkulatzailak.

Hau dela eta, eritasunarekiko suszeptibilitatearen atzean dauden geneak identifikatzea ezinbestekoa da. Alde batetik, CDren garapenean eragina duten mekanismo patogenikoak ulertzeko eta beste aldetik, eritasuna pairatzeko arriskua duten indibiduok identifikatzeko gaitasuna emango liguketuen markatzaile genetikoak lortzeko. Markatzaile hauek lortuz gero diagnostiko prediktiboa burutzea posible izango zen, erantzun autoimmunitarioa hasi baino lehen, eta honek, prebentzio saiakuntzak burutzeko hautagaien aukeraketa erraztuko luke.

Eritasun autoimmunitario konplexu honen genetika ulertzeko helburuarekin, proiektu honetan eritasun zeliakoan burutako genoma osoko asoziazio azterketetan (GWAS) eta haietatik proposatutako gene hautagaietan oinarritu gara, eritasun zeliakoan eragina izan dezaketen partaide funtzionalak aztertzeko.

Sarrera

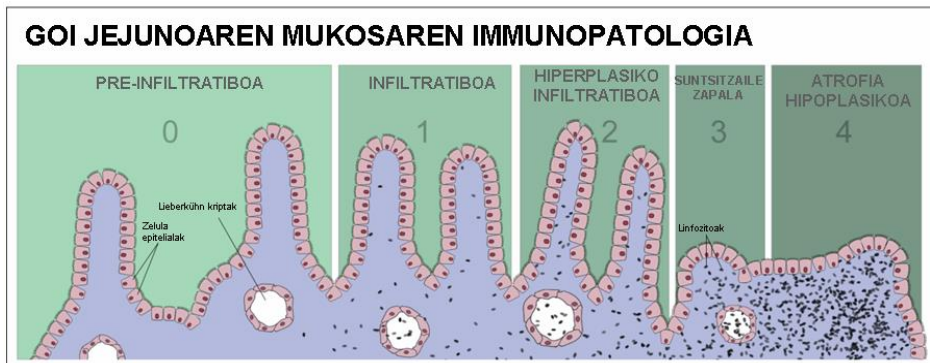
1. Eritasun zeliakoa

Glutenarekiko sentikorra den enteropatia (MIM 212750) edo eritasun zeliakoa (CD) gaixotasun kroniko eta inflamatorioa da, eta bere ezaugarriak nabarmenena heste meharreko biloxken zapaltzea da, dietan hartutako gariaren glutenarekiko zein zekale eta garagarrean agertzen diren antzeko proteinekiko intolerantzia dela medio.

1.1. Ezaugarri klinikoak eta diagnosis

Dietan hartutako glutenaren efektu kaltegarriak 1950era arte antzeman ez ziren arren ¹, eritasun zeliakoaren agerpen klinikoen lehenengo irudi orokorrak 60 urte lehenago ikusi zuen argia ². Tratatzen ez bada, CDk honako sintomak aurkeztu ohi ditu: batetik, heste meharrari dagozkion ezaugarriak, hala nola, beherakoa, distentsio abdominala edota goragalea, eta bestetik, malabsortzioak eragindakoak, garapenari dagozkionak (pisu baxua, gantzarik eza, ilearen fintasuna) edo urritasun psikomotorreari dagozkionak (giharren ahultzea...), besteak beste ³. CDri bestelako sintoma atipiko batzuk ere egotzi izan zaizkio. Hauen artean gertaera neurologikoak, akatsak hortzetako esmaltean, infertilitatea, osteoporosia, giltzadura mailako sintomak eta gibelesko entzimen kontzentrazio altuak dira nabarmenenak ⁴. Ikuspuntu histologikotik, pertsona suszeptibleak glutendun dieta hartzen duenean, heste meharraren endekatze graduala gertatzen da. Prozesu honen emaitza biloxken atrofia eta kripten hiperplasia ezaugarri dituen heste lesioa izaten da (1. irudia).

Glutenak eragindako hesteko mukosaren kaltearen maila eta larritasuna Marsh-Oberhuber sailkapenean dago definituta ⁵. Laburbilduz, hasieran linfozito intraepitelialen infiltrazioa dago biloxken epitelioan (Marsh I). Ondoren, kripten hiperplasia ematen da (Marsh II), biloxken tamaina mantentzen delarik. Estadiorik aurreratuenean (Marsh III), kripten hipertrofiaz gainera, *lamina propria* hanpatuta dago eta biloxken atrofia partziala, subtotaleta edo totala gertatzen da (1. irudia). Hesteko gertaerez gain, CDn glutenaren menpeko diren hainbat autoantigorputz ere agertzen dira, anti-endomisio (EMA) edo anti-transglutaminasa (TGA) antigorputzak kasu ⁶.



1. irudia: Glutenak induzitutako aldaketak estadio ezberdinetan, Marsh sailkapenaren arabera. www.theglutensyndrome.net orrialdetik hartuta.

Ezaugarri patologiko hauek kontutan hartuta, Gastroenterologia, Hepatologia eta Elikadura Pediatrikoaren Elkarte Europarrak (ESPGHAN) irizpide diagnostikoak finkatu ditu. Orain dela gutxi arte, kriterio hauek heste meharreko biopsien kalte histologiko eta emaitza serologiko positiboetan oinarritu izan dira (azken hauek ezinbesteko izan ez badira ere) ⁷, baina badira bi urte irizpide diagnostiko berriak onartu direla. Azken hauen arabera, antigorputzek eta ezaugarri klinikoek lehentasuna dute eta duodenoaren biopsia egiteari utz dakioke baldin eta umea sintomatikoa bada eta IgA motako TGA maila normaltasunaren goiko muga baino 10 aldiz altuagoa badu. Giza Antigeno Leukozitarioaren (HLA) genotipaketa oso baliagarria izan daiteke, CD oso ezohikoa delako arrisku alelorik gabe ⁸.

1.2. Epidemiologia

Aurreko mende bukaerara arte, CD ezohiko gaixotasuntzat hartzen zuen komunitate zientifikoak, europarren artean prebalentzia 1/1000 zela pentsatzen baitzen ³. Dena den, berrikiago argitaratutako zenbait populazio ikerketen arabera, maiztasuna %1 da mendebaldeko europarren artean, desberdintasunak desberdintasun ⁹. Zenbait ikerlanek iradoki duenez, CDren haurtzaroko agerpenaren maiztasuna haziz joan da 1980 eta 1990eko hamarkadetan zehar, eta hau umeen elikadura irizpideekin erabat lotuta dago ¹⁰. Bestetik, diagnostikoek helduengan ere gora egin dute informazioa eskuragarri dagoen lurraldeetan ¹¹⁻¹³. Genetikoak ez diren arrisku faktoreen artean urtaroen arabeko patrioiak dituzten faktoreak eta zenbait infekzio biriko aipatu izan dira ¹⁴.

CD europar jatorriko pertsonen modu ia eskusiboan pairatzen duten gaixotasuna dela uste zen, baina test serologiko berrien hedatzeak CDren maiztasunaren estimazio zehatzagoak ekarri ditu, zoriz aukeratutako Ipar Amerika eta Australiako populazioetan eritasunaren maiztasuna 1:100 eta 1:200 artekoa dela agerian utziz^{15,16}. Bestetik, CD Latinoamerikan¹⁷⁻¹⁹, Afrikako iparraldean^{20,21} eta Ekialde Hurbilean^{22,23} ere arraroa eta ezohikoa dela pentsatu izan da, leku hauetan kasu mugatu eta puntualak soilik behatu izan baitira. Gainera, Ekialde Urrunean (Txina, Japonia, Korea...) existitzen ez zela uste izan da²⁴. Dena den, bahetze edo *screening* berriek munduko leku hauetan guztietan CDren maiztasuna gutxietsi izan dela frogatu dute, agerpen frekuentzia mendebaldeko herrialdeetan erregistratutakoaren antzekoa dela erakutsiz. Mendebaldeko dieta modernoaren zabalkundearekin batera, glutendun zerealak eta bereziki garia, munduko leku guztietara hedatu dira eta CD Osasun Publikoko arazo bihurtu da, baita garapen bidean dauden herrialdeetan ere¹⁶.

Generoaren ikuspuntutik, emakumeak maizago daude afektatuta gizonak baino (proportzioa 3:1 ratiara iristen da garai emankorrean dauden gaixoen artean)²⁵.

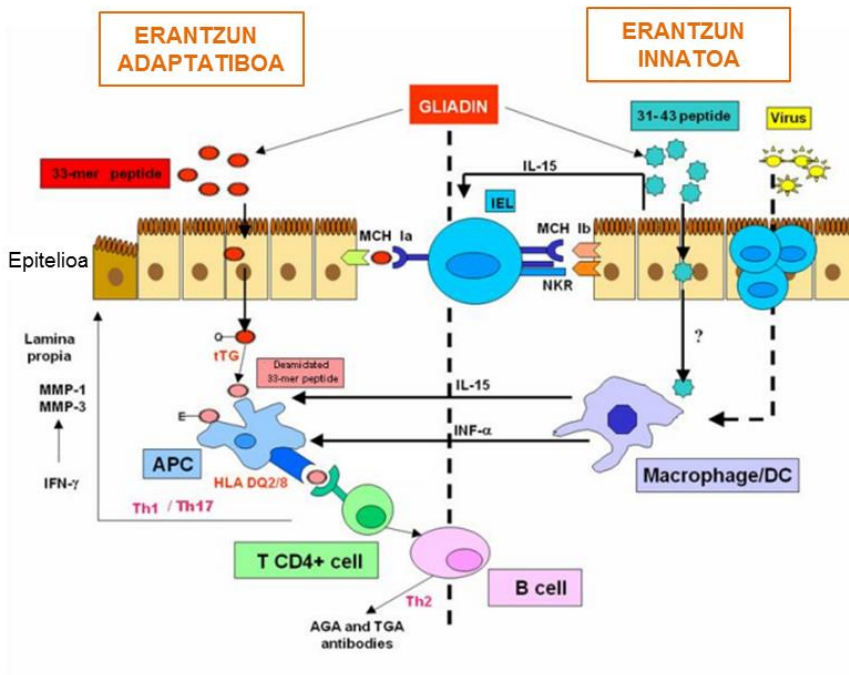
Egun, tratamendu eraginkor bakarra biziarteko glutenik gabeko dieta (GFD) hertsia da, gari, zekale eta garagarraren produktuak dietatik ezabatuz^{3,26}. Nolanahi ere, GFD zorrotza betetzea ez da erraza eta bizi kalitatea gutxitzen du. Tratamendua modu zorrotzean jarraitzen ez bada konplikazioak ager daitezke, hala nola, garaiera baxua, desnutrizioa, osteoporosia, gaixotasun autoimmune sekundarioak, minbiziak eta zailtasunak haurdunaldian²⁷.

2. Eritasun zeliakoaren patogenesisia

Gaixotasun zeliakoaren garapenean parte hartzen duten mekanismoen ulermenean azken urteetan egindako aurrerapenek, CD HLArekin erlazionaturiko gaixotasunen artean hobekien ezagutzen duguna izatea ekarri dute. Hala ere, prozesu patogeniko batzuk oraindik azaltzeke daude.

T zelulek CDn parte hartzen dutela aspalditik ezaguna da: glutenaren peptidoak epitelioa zeharkatzen dute eta *lamina propria* iristean, ehunetako transglutaminasak deamidatu egiten ditu, DQ2 eta DQ8 hartzailekun APCek CD4+ T zelula patogenikoei aurkezteko. Honek gerora *lamina propria*ren infiltrazioa, kripten hiperplasia eta biloxken atrofia eragingo dituen Th1 erantzuna ekarriko du²⁸. Dena den, azken hamarkadetako hainbat

azterlanek erakutsi dutenez, erantzun immune innatoak ere berebiziko garrantzia du CDn eta T zelulak bitartekari ez dituen erantzuna aktibatzeneko gai ere bada gliadina^{29,30} (2. irudia).



2. irudia: Eritasun zeliakoaren mekanismo patogenikoak: adar immune adaptatibo era innatoak. (Castellanos-Rubio *et al.* (2010) moldatua³¹).

2.1. Glutena

Glutena alkoholetan disolbagarria diren glutenina monomerikoen eta prolamina aberatsak diren gliadina polimerikoen nahastura da eta gari, zekale eta garagarrean dago. CD suszeptibilitate genetikoaren duten indibiduoetan gertatzen da, dietako glutenaren proteinekiko kontaktua dagoenean. Hesteko proteasek ezin dituzte proteina hauek guztiz endokatu eta 10-15 aminoazidoko peptidoak geratu ohi dira hesteko lumenean. Proteina zati hauek substratu onak dira transglutaminasarentzat, eta honek deamidatu ditzake, HLA-DQ2 eta HLA-DQ8 molekulekiko afinitatea handituz eta ondorioz, CD4+ T zelulek gidatzen duten Th1 erantzuna kitzikatuz. Bestalde, glutenak CD8+ T zelulen erantzuna

ekar dezake, linfozito intraepitelialen populazioa emendatuz, HLArekiko independentea den mekanismo baten bidez^{28,32}.

Heste kulturetan, zelula antigeno aurkezle primarioetan nahiz lerro zelular epitelial eta monozitiko ezberdinetan gauzaturiko zenbait *in vitro* azterketek ideia hau indartu dute. Horrela, gliadinaren 31-43 arteko aminoazidoak hartzen dituen peptido batek sistema immune innatoa estimula dezakeen ikertu da.

HLA-DQ2k ez du peptido hau lotzerik eta ez du ematen T zelulentzat espezifikoa den erantzunik eragiten duenik. Biopsiak bi orduz estimulatu ostean, peptido honek apoptosia induzitu dezake edota hazkunde faktore epidermikoaren hartzailaren (EGFR) bidezko bidezidor endozitikoa inhibitu, glutenaren peptidoaren aurkezpena eta ondorizko CD4+ T zelulen aktibazioa gertatu baino lehen^{32,33}. Ideia honen alde, gliadinarekin inkubaturiko CD biopsiek *MICA* genearen gainadierazpena erakutsi dutela ikusi da. Estres molekula honek, $\gamma\delta$ eta NK zeluletan adierazten den NKG2D hartzailarekin interakzionatu eta erantzun innatoko zitokinak aktiba ditzake gaixotasunaren lehen estadioretan, sistema immune innatoa adaptatiboarekin lotuz³⁴.

2.2. Transglutaminasa

Eritasun zeliakoan, glutenak ehunetako transglutaminasaren aurkako IgA klaseko antigorputzen ekoizpena induzitzen du. TG2 nonahi adierazten den proteina multifuntzionala da eta orokorrean, espazio estrazelularrean da aktibo. Bi proteine arteko glutamina- lisina lotura itzulezin eta kobalenteak katalizatzen ditu^{35,36}. Glutena prolina eta glutaminetan aberatsa da eta gainera, oso aminoazido negatibo gutxi ditu (HLA-DQ2 eta DQ8 molekulekin interakzionatzeko beharrezko ezaugarria). Glutenetik eratorritako peptidoek TG2ren bidez deamidatuak izan behar dute, molekulen eraldaketaren ostean, zitokinen ekoizpen lokal masiboa eragin eta enterozitoen desberdintzapena eta proliferazioa aldaraziko badituzte³⁷. Bestetik, TG2k katalizatutako erreazioan parte hartzen duten peptidoek eta entzimak berak, TG2-gliadina konplexuak eratzen dituzte, autoantigorputzen ekoizpena bultzatuz³⁸. T zelulek ezagutu ditzaketan hiru epitopoeekin partzialki gainezartzen den sei aminoazidoko sekuentziadun 33 aminoazidoko polipeptidoa (33-mer) bakarra da immunogenikoena gliadinatik eratorritakoaren artean, ehuneko transglutaminasak deamidatu ostean.

Anti-transglutaminasa antigorputzek patogenesisian eta hesteko mukosaren ohiko lesioan parte hartzen duten argitzeke dago. Eritasun zeliakoan agertzen diren autoantigorputzek apoptosian, angiogenesisian eta heste iragazkortasunean inpaktu biologikoa dutela ikusi da eta beraz, TG2 patologian paper garrantzitsua izan zezakeela iradoki da ^{38,39}.

2.3. Sistema immune adaptatiboa

Sistema immune adaptatiboak T zelulen bidezko immunitatea eta immunitate humorala indutzten ditu, eta biak ala biak CD pazienteen heste meharreko mukosan aktibatzen dira, gliadinaren estimuluz. Hesteko CD4+ T linfozitoek, APC HLA-DQ2 eta HLA-DQ8 heterodimeroek aurkezten dizkieten deamidatutako gliadina peptidoak ezagutzen dituzte ^{40,41}. Gliadinarekiko espezifiko diren mukosa zeliakoko T zelulak Th1 fenotipokoak dira batez ere eta inflamazioa sustatzen duten zitokinak ekoizten dituzte, IFN γ kasu ^{42,43}. IFN γ -z gain, interleukina 18 eta IFN α ere gainadierazita daude ⁴⁴⁻⁴⁶. IL6 eta TGF β ren presentzian ezberdintzatu eta interleukina 17 bultzatzen den bestelako CD4+ T zelula laguntzaileen lerro bat (Th17 linfozitoak) identifikatu da, eta uste denez, lerro hau lehen IL12/ IFN γ sareari egotzen zitzaizkion zenbait funtzioaren erantzulea da ⁴⁷. Th1 eta Th17 erantzunak gaixotasun aktiboan parte hartzen dute, beste zenbait gaixotasun autoimmune batzuetan ikusi den moduan hain zuzen ⁴⁸⁻⁵⁰.

2.4. Sistema immune innatoa

Erantzun immune innatoa patogenoen aurkako lehen harresia da eta agente infekziosoekin topo egin osteko lehenengo estadioetan aktibatzen da. CDn, immunitate innatoak gliadinaren estimulari aurre egiten dio T α β linfozitoekiko modu independentean eta gerora emango den T zelulen aktibazioarako inguru proinflamatorio egokia sortzen laguntzen du. Zenbait *in vivo* azterketek erakutsi duenez, gliadinaren 31-43 aminoazido sekuentziak gaixotasunaren sintomak agertaraz ditzake, eta CDn nabariak diren hainbat aldaketa behatu izan da biopsia kulturetan ⁵¹⁻⁵³. Peptido honek ez du T zelulen bidezko erantzunaren estimulazioa eragiten ^{54,55} eta beraz, pentsatzekoa da 31-43 proteina zatiaren toxizitatea immunitate innatoa aktibatzeke gaitasunari dagokiola.

Zenbait azterketek Toll-like receptor 4 (*TLR4*) hartzailea eta bere seinale transduktore nagusia, MyD88 inplikatu dituzte, monozitoetan, makrofagoetan eta zelula dendritikoetan zerealen proteinekin ematen den erantzunaren hartzaile eta eragile primario modura,

hurrenez hurren²⁸. Glutenak eragindako linfzito intraepitelialen (IEL) aktibazio innatoak *MICA* genearen adierazpena indutzen du hesteko epitelioan, NK eta T $\gamma\delta$ zeluletan eta CD4+ nahiz CD8+ diren zenbait T zelula azpaldetan adierazten den NKG2D hartzailaren ligandoarena, hain zuzen ere. *MICA* epitelialak, gainadierazitako IL15ekin batera, NKG2Dren aktibazioa sustatzen du IELetan, linfzitoak bitartekari modura dituen zitotoxizitate antigeno-espezifikoak bultzatuz. Azkenik, IL15ekin batera diharduen IL21 immunitate innatoaren gidaritzan aipagarria dela argitaratu da^{30,56}.

3. Eritasun zeliakoaren genetika

Eritasun zeliakoaren herentzia patroia oraindik ezezaguna bada ere, genetikak gaixotasun honekiko suszeptibilitatean parte hartzen duela aspalditik jakina da. Familiatan egindako prebalentzia ikerketak eta bereziki, bikiak konparatu dituztenak, berebizikoak izan dira inguru faktoreen eta faktore genetikoen ekarpenak zenbatekoak diren estimatzeko. Azterlan hauen arabera, genetika funtsezkoa da CDren hasieran zein garapenaren fase guztietan.

Orokorrean, komunitate zientifikoak CDren agerpenean bat datozen biki monozigotikoen proportzioa %75-86 dela onartzen du. Biki dizigotikoen kasuan, aldiz, kopuru hau nabarmenki jaisten da, %16-20ra arte, hain zuzen. Biki mota bien konkordantzia desberdintasunek CDren osagarri genetikoa estimatzen lagundu dute eta orain arte bestelako gaixotasun autoimmuneentzat behatutakoa baino pisuzkoagoa dela argi utzi dute. Adibidez, 1 motako diabetesan (T1D) osagarri genetikoa baxuagoa da (konkordantzia biki monozigotiko eta dizigotikoetan %30 eta %6, hurrenez hurren)⁵⁷. Gainera, anai-arreben eta biki dizigotikoen arteko konkordantzia ratioak oso antzekoak dira CDn eta beraz, inguru faktoreek aportazio minimoa dutela ondoriozta daiteke.

Laburbilduz, azken urtetan jasotako ebidentziek CDk osagarri genetiko oso inportantea duela eta gaixotasunaren heredagarritasuna (faktore genetikoei egotzi ahal zaion CD pairatzeko arriskuaren proportzioa, inguru faktoreei egotzi ahal zaionarekin alderatuta, beti ere) %87 ingurukoa dela aditzera eman dute⁵⁸. Arrisku genetikoen parterik handiena Giza Antigeno Leukozitarioaren (HLA) zenbait aleloren presentziari egotzi ahal zaio. Dena den, eta gaixotasunaren patogenesia HLA molekula hauen papera oinarrizkoa bada ere, herentziari egiten dioten ekarpena ez da oso handia. 2015ean Gutierrez-Achury *et al.* argitaratutako artikuluan, *HLA-DQA1* eta *HLA-DQB1* lokus klasikoen ekarpena arrisku genetikoan %23 dela kalkulatu da, baina beste 5 HLA aldaera berri aurkitu dira

%18ko ekarpenarekin, biak batera CDrekiko arriskuaren %40a azaldu dezaketelarik ⁵⁹. Honek pentsarazten digu, HLA lokusetik kanpo efektu txikiko lokus ezezagun ugari egon beharko litzatekeela.

3.1. HLA eremua eta eritasun zeliakoa

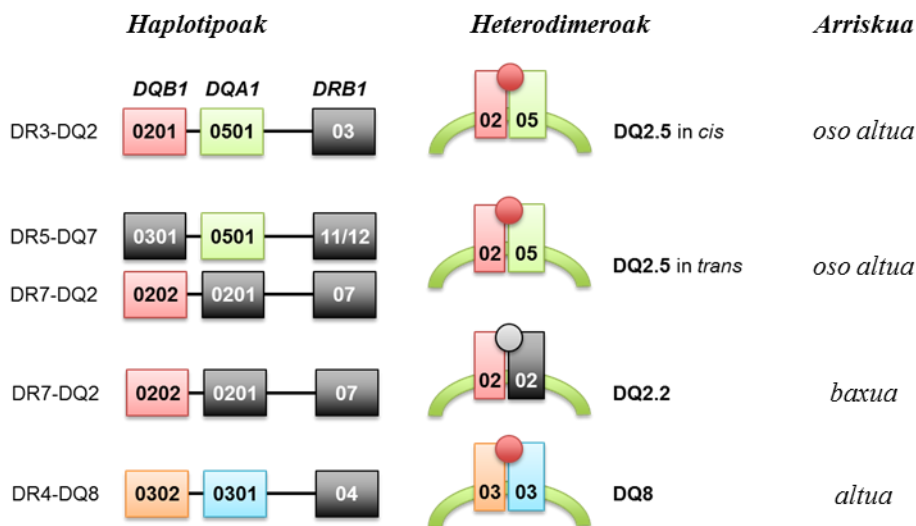
3.1.1. HLA eremua

HLA da giza Histokonpatibilitate Konplexu Nagusiari (MHC) ematen zaion izena; 6p21 eremu kromosomikoko superlokusean dago eta sistema immunearekin zerikusia duten gene ugari dauka. HLA geneek giza zelula gehienetan adierazten diren proteina antigeno aurkezleak kodetzen dituzte. Hauek ezinbestekoak dira molekula exogeno eta endogenoen artean bereizteko.

HLA geneek gaixotasun autoimmune eta inflamatorio askotan parte hartzen dute, eta malaria edo HIESa bezalako eritasun infekziosoak pairatzeko arriskua emenda dezakete. Hala ere, gene talde zabal honek eragindako suszeptibilitatea azaltzen dituzten faktore genetiko zehatz eta mekanismo patogeniko gehienak ezezagunak izaten segitzen dute egun, eremuko konplexutasun genetiko handia dela eta. Izan ere, HLAk genoma osoko gene dentsitaterik altuena eta adierazpen maila goratua ditu ⁶⁰.

3.1.2. Suszeptibilitate geneak eta arrisku genetikoa

Lehen esan bezala, HLA suszeptibilitate eremurik inportanteena da CDn eta gaixotasunaren osagarri genetikoaren %40 (%23 aldaera klasikoak eta %18 berriak) inguru azaltzen du. HLA eta CDren arteko lehenengo asoziazioaren ebidentziek 1973an ikusi zuten argia ⁶¹. Eremuko lotura desoreka indartsua dela eta, lehenengo azterketek HLA-A1, HLA-B8 eta HLA-DR3 identifikatu zituzten aldaera etiologiko modura, baina beranduago publikatutako ikerlanek modu zuzenean inplikaturiko faktoreak HLA-DQ2 eta -DQ8 kodetzen dituzten II klaseko HLA geneak direla erakutsi dute (3. irudia).



3. irudia: HLA lokusaren eta CDren arteko asoziazioa. HLA-DQ2 molekula CD garatzeko arrisku faktorerik garrantzitsuena da. Gaixo zeliako gehienek HLA-DQ2.5 heterodimeroa adierazten dute, HLA-DQA1*05 (α katea) eta HLA-DQB1*02 (β katea) aleloek kodetuta. Alelo hauek cis eran ager daitezke, DR3-DQ2 haplotipoan, edo transen, DR5-DQ7 eta DR7-DQ2.2 heterozigotoetan. HLA-DQ2.2 dimeroak, HLA-DQA1*02:01 eta HLA-DQB1*02:02 aleloek kodeturiko HLA-DQ2ren aldaerak, hain zuzen, gaixotasuna garatzeko arrisku txikia ematen du. DQ2 negatiboak diren paziente gehienek DR4-DQ8 haplotipoak kodetzen duen HLA-DQ8 adierazten dute. (Abadie et al., Annu Rev Immunol (2011) publikaziotik adaptatuta ⁶²).

Eremuko asoziaziorik indartsuena HLA-DQ2rekin aurkitutakoa da eta gaixoen %90ak HLA-DQ2.5 heterodimeroaren kopia bat dauka gutxienez (heterodimeroaren α eta β kateak kodetzen dituzten DQA1*05 eta DQB1*02 aleloek osatuta). Bestalde, zeliakoa ez den populazioaren %20-30ak HLA-DQ2 aldaera hau eramaten du eta beraz, garrantzitsua izan arren, gaixotasuna garatzeko nahikoa ez dela argi utziz. HLA-DQ2 eramaileak ez diren gaixo gehienek HLA-DQ8 dute eta ondorioz, gutxienez, DQA1*03:01 eta DQB1*03:02 aleloak kodetzen dituen haplotipoaren kopia bat daukate ⁴. Pazienteen proportzio txiki bat DQ2 zein DQ8rako negatiboa da, baina kasu gutxi hauetan ere, gaixoak DQ2 molekula kodetzen duten aleloetako baten eramaileak direla ikusi da (DQA1*05 edo DQB1*02) ^{63,64}.

HLA-DQ2 eta HLA-DQ8 aldaerak lotura desorekan daude DR3 eta DR4rekin, hurrenez hurren, eta hori dela eta, askotan, arrisku alelo hauek haplotipoen izenak hartzen dituzte,

DR3-DQ2 eta DR4-DQ8, hain zuzen ere ⁶⁵. Haplotipo batzuetan, DR3-DQ2an kasu, HLA-DQ2.5 heterodimeroaren alelo biak (DQA1*05:01 eta DQB1*02:01) kromosoma berean daude eta beraz, *cis* eran kodetuta. Aldiz, DR5-DQ7 eta DR7-DQ2 haplotipoen eramaile heterozigotoen kasuan, arrisku heterodimeroan parte hartze duten molekulak *trans* eran daude kodetuta, kromosoma ezberdinetan kokatuta daudelako. HLA-DQ2.5 mota bi hauen arteko desberdintasunak DQ α kateko aminoazido batean (DQA1*05:01 *vs.* DQA1*05:05) eta DQ β katearen mintzaren eremuko beste residuo batean (DQB1*02:01 *vs.* DQB1*02:02) datzate, baina ez du ematen ondorio funtzionalik dutenik, eta arrisku efektu berdinarekin asoziatuta daude. Dena den, HLA-DQ2.2k (HLA-DQ2ren beste aldaera bat) ematen duen arriskua oso txikia da ^{57,66}.

CDrekiko suszeptibilitate mailaren eta DQ2.5 heterodimero kopuruaren arteko harremana ere ikusi da. Bi DR3-DQ2 haplotiporen eramaile homozigotoak batetik, eta DR3-DQ2/DR7-DQ2 heterozigotoak bestetik, DQ2.5 mailarik altuenak adierazten dituzten indibiduoak dira eta beraz, CD pairatzeko arriskurik handiena ematen duten genotipoen eramaileak ⁶⁷⁻⁶⁹. Zentzu honetan, aipagarria da CD errefraktarioa duten pazienteek (GFDren aurrean erantzuten ez dutenak, hain zuzen) DR3-DQ2 haplotipoaren homozigositate maiztasun altuagoa aurkezten dutela (%44-62) beste gaixo zeliakoekin alderatuta (%20-24). DQ8 molekulekin ere, dosiaren menpeko antzeko efektua ikusi izan da.

DQ molekulak kodetzen dituzten geneez aparte, HLA eremuak baditu sistema immunean parte hartu eta CD pairatzeko arriskua emenda dezaketen beste gene asko ere. Azterketa ezberdinek *MICA*, *MICB*, *TNF* eta bestelako geneetako polimorfismoek gaixotasuna izateko arriskuan eragina dutela proposatu dute. Hala ere, orokorrean, lan hauek ez dute eremuko lotura desoreka indartsua behar beste kontutan hartu, eta beraz, emaitzak ez dira erabakigarriak. Nahiz eta HLA geneek berebiziko ekarpena egin suszeptibilitate genetikoari, gaixotasunaren presentzia bat datozen anai-arrebak gutxi dira, %30 inguru, eta ondorioz, HLA geneak garrantzitsuak izan arren gaixotasuna garatzeko nahikoak ez direla esan daiteke ⁶⁶.

3.1.3. HLA-ren rola eritasun zeliakoaren patogenesisian

CDren eta II klaseko HLA geneen arteko asoziazioa guztiz lotuta dago T zelula CD4+ek gaixotasunaren patogenesisian duten funtsezko paperarekin. Izan ere, CDn, T linfozito CD4+ak gai dira zeliakoaren hestean agertzen diren glutenetik eratorritako peptidoak

ezagutzeko, ez ordea indibiduo osasuntsuen kasuan. Zelula antigeno aurkezleen (APC) gainazaleko HLA-DQ2/HLA-DQ8 molekulek glutenetik eratorritako epitopoak aurkezten dituzte, glutenarekiko espezifikokoak diren T zelula CD4+en proliferazioa estimulatuz²⁸.

HLA eta CDren arteko asoziazioa ulertzeko bidean mugarririk aipagarrienetako bat HLA-DQ2 edo/eta -DQ8 eta gliadina peptidoen arteko lotura afinitateari dagokio; izan ere, 2 motako transglutaminasa (TG2) entzimak eraldatutako peptidoekiko altuagoa baita. Entzimak glutenetik eratorritako peptidoen karga negatiboek areagotzen dituen erreakzio bat katalizatzen du, epitopo eraldatuen eta HLA molekulen (DQ2 eta DQ8) arteko lotura faboratuz eta T zelula CD4+kopurua emendaraziz.

Glutenarekiko espezifikokoak diren T zelula autorreaktiboan aktibazioan HLA molekulek duten garrantzia ikusirik, azken hauen sekuentzia kodetzaileetako edozein aldaketak prozesuaren puntu ezberdinetan alterazioak ekar litzakeela ondoriozta genezake. Horrela, antigenoaren lotura guneak kodetzen dituen sekuentziaren polimorfismoek afinitatea aldarazi lezakete, glutenetik eratorritako peptidoen ezagutza faboratuz edo oztopatuz⁷⁰. Bestetik, eremu erregulatzailerako zenbait aldaerek HLA molekulen adierazpena eralda lezake, glutenarekiko erantzun immunea txikiagotuz edo areagotuz.

3.2. Genoma osoko asoziazio azterketak CDn

2006tik hona, lehenengo GWA azterketa publikatu zenetik⁷¹, aldagai eta eritasun konplexuen azterketa genetikoak planteatzeko modua aldatu da. Azken 10 urteetan izandako eboluzioari esker, gaur egun GWA azterketak markatzaile genetiko asko eta DNA lagin multzo handietan aztertzeko tresna oso baliagarriak dira, batez ere gaixotasun zehatzekin asoziatutako aldaera genetikoak aurkitu eta aztertzeko helburuarekin. Ikerketa hauek bereziki erabilgarriak dira populazioan arruntak diren eritasun genetiko konplexuetan eragina duten aldaerak aztertzen ditugunean, CDren kasuan bezala.

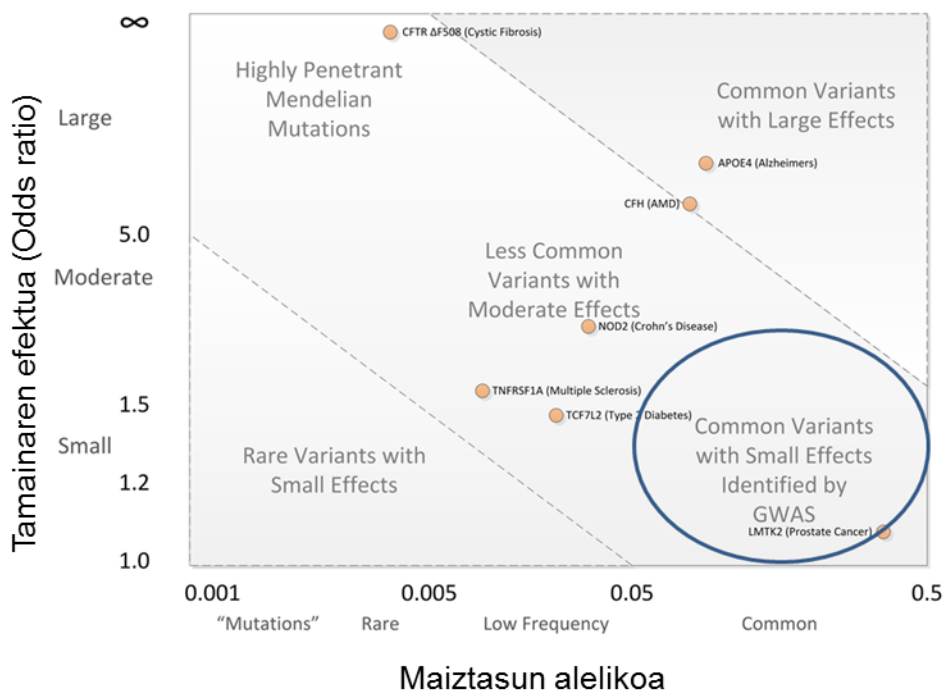
GWASen aldaera genetikoaren unitatea SNPa da. SNP gehienek eragin minimoa daukate sistema bilogikoetan, baina gutxi batzuk eragin funtzionalak izan ditzazkete, hala nola aminoazido aldaketak, transkribaturiko mRNAren egonkortasunean aldaketak, edota TFekiko lotura afinitatean aldaketak⁷².

GWAS azterketa bat burutzeko bi partaide mota erabiltzen dira: aztergai den eritasuna pairatzen duten indibiduoak (kasuak) eta antzeko ezaugarriak (sexua, adina...) dituzten

indibiduo ez gaixoak (kontrolak). Kasu zein kontrol laginen aukeraketa optimoa erabakigarria da GWA azterketen diseinuan. Kasuak aukeratzeko orduan heterogeneitate fenotipikoa ekiditeak garrantzia handia dauka, eta geroz eta interes handiago dago GWA metodologia populazioetan oinarrituriko kohorteetan erabiltzearen inguruan, nahiz eta gaur egun arte publikaturiko azterketa gehienetan kasu-kontrol hurbilketa erabili den ⁷³. Jatorri demografiko eta aneztral desberdina duten indibiduen presentziak populazioaren estratifikazioa eragin dezake. Alde honetatik kasu eta kontrolen arteko desberdintasunak egonez gero, laginaren jatorriarekin erlazionatuta dauden markaitzaileak eritasunaren presentziarekin lotu daitezke, asoziazio faltsua ondorioztatuz. Aukeratutako kontrolak populazio orokorraren islada ez izatea ohiko akatsa izaten da ere ⁷³. Laginaren tamaina izan ohi da GWA azterketetan kontutan edukitzeko beste arazo bat, eta kasu honetan konklusioa garbi dago: zenbat eta gehiago hobeto ⁷⁴.

Asoziazioatutako aldaerek, hau da, kasuetan kontroletan baino maiztasun altuagoa duten SNPek, eritasuna sortu edo eragina izan dezakeen aldaerak egon daitezkeen eremu genomikoak markatzen dituzte. Asoziaturiko SNPak ez dira zertan aldaera kausalak izan behar, horregatik, SNP hauek seinalatzen dituzten eremuen azterketa sakona burutzea guztiz beharrezkoa da. Gehienetan azterketak aurrera doaz asoziatu-eremuen sekuentziazio sakonarekin, aldaera genetiko zehatzak identifikatzeko helburuarekin eta, analisi funtzionalen bitartez aldaeren eta geneen adierazpen mailaren arteko korrelazioak aurkitzeko asmoarekin.

Beraz, GWA azterketak aldaera genetiko berriak identifikatzeko metodo egokia dira. Pedigriak erabiltzen dituzten familietan oinarrituriko asoziazio azterketak efektu fenotipiko handia duten aldaera arraroak identifikatzeko erabilgarriak diren bitartean, populazio orokorrean oinarritzen diren genoma osoko hurbilketak erabilgarriak dira efektu txikia duten aldaera arruntak identifikatzeko (4. irudia).



4. irudia: Gaixotasunetan eragina duten aleoen espektroa. Gaixotasunekiko asoziazioak 2 dimentsiotan kontzeptualizatu ohi dira: aleoen frekuentziak eta efektuaren tamaina. Eritasun Mendeliarren aleoek sarkortasun eta efektu handia daukaten aleo arraroak dira (goian ezkerrean). GWAS aurkikuntzak aldiz, efektu txikia duten SNP arruntak izaten dira (behean eskuman). (Bush et al. (2012) doi:10.1371/journal.pcbi.1002822.g001 artikulatik moldatua ⁷⁵).

3.2.1. Lehenengo GWASaren emaitzak

CDn gauzaturiko lehen GWASean, 778 gaixo eta 1.422 indibiduo kontrol ikertu ziren. Asoziazio azterketak populazio mailan alelo txikiaren maiztasuna %1 baino handiagoa zuten 310.605 SNPetan burutu ziren ⁷¹. Asoziaziorik indartsuena HLA lokusean identifikatu zen, espero bezala eta rs2187668*A aleloa HLA-DQ2.5 *cis* eran kodetzen duen DR3-DQ2 haplotipoaren markatzaile efizientea dela ondorioztatu zen. Lehenago aipatu bezala, hau da CDrekin asoziatu HLA-DQ2 haplotiporik arruntena. Lehenengo GWAS honetan, Erresuma Batuko populazioan, pazienteen %89,2k gutxienez HLA-DQ2.5 haplotipoaren kopia bat aurkezten zuten, kontrolen artean indibiduoaren %25,5ean soilik agertzen zen bitartean.

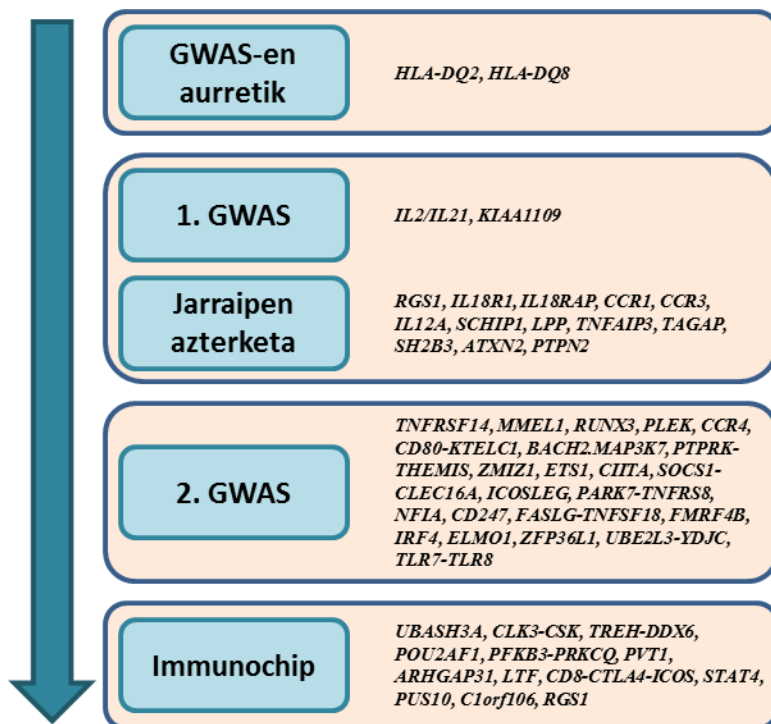
HLA eremutik kanpo, $p < 10^{-4}$ aurkezten zuten 56 SNP aurkitu ziren. Haietako batzuk elkarrengandik gertu zeuden, asoziazioa benetakoa zela eta polimorfismo horiek CDren osagai genetiko konplexuan parte hartzen zuten aldaera kausalekin lotura desorekan egon (LD) zitezkeela iradokiz.

Dena den, gaixotasunarekin asoziatuiko SNP bakarra ($p < 10^{-6}$) rs13119723 izan zen, lotura desoreka aurkezten duen eta *IL2* eta *IL21* hartzen dituen 4q27 eremuan kokatuta, hain zuzen ere. Emaiza hauek holandar eta irlandar kohorte independenteetan erreplikatu ziren. Ereku honek gaixotasuna pairatzeko arrisku genetikoaren %1 soilik azalduko zuela estimatu zen, oraindik ezezagunak ziren suszeptibilitate gene gehiagoren existentzia iradokiz. Asoziazio indartsuenak agertu zituzten 1.164 SNPak hiru kohorte europar independenteko 1.643 kasu eta 3.406 kontrol gehiagotan aztertu ziren⁷⁶. Azterketa honen ostean identifikaturiko eremuak sakonago ikertu ziren, CDren garapenean funtzionalki inplikaturiko hautagaiak aurkitzeko helburuz, eta bereziki sistema immunean parte hartzen duten geneak kontuan hartuz (5. irudia).

3.2.2. Bigarren GWASaren emaitzak

CDn egindako bigarren GWASa 2009. urtean gauzatu zen. HLA kanpoko 292.387 SNP analizatu ziren, jatorri europarreko 4.533 paziente zeliako eta 10.750 kontrol osasuntsutan. Gainera HLA eremuan ez zeuden beste 231.362 polimorfismo gehiago aztertu ziren 3.796 gaixoz eta 8.154 kontrol osaturiko kohorte independente batean⁷⁷.

Aurretik ezezagunak ziren asoziatuiko hamahiru arrisku eremu identifikatu ziren (5. irudia). Ereku hauek, sistema immunearekin erlazioatutako funtzioak dituzten hainbat gene daude: *BACH2*, *CCR4*, *CD80*, *CIITA-SOCS1-CLEC16A*, *ETS1*, *ICOSLG*, *RUNX3*, *THEMIS*, *TNFRSF14* eta *ZMIZ1*. Beste 13 lokusek ez zuten genoma osoko esangarritasun maila lortu, baina gaixotasunarekin nolabaiteko erlazioa zutela ikusi zen. Gainera, funtzio immunea zuten zenbait gene hartzen zuten eremu hauek, hala nola, *CD247*, *FASLG*, *TNFSF18*, *TNFSF4*, *IRF4*, *TLR7*, *TLR8*, *TNFRSF9* eta *YDJC*.



5. irudia: Aurrerapenak CDren genetikan. ImmunoChip ikerlanaren ostean, gaixotasuna garatzeko arriskuan parte hartzen duten 40 lokus identifikatu dira. Asoziazio hauen kausa diren aldaera espezifikoak ezezagunak izaten jarraitzen dute. Horrela, azterketa funtzionalak beharko dira asoziazio lan erraldoi hauen aplikazioak zeintzuk izan daitezkeen argitzeko. (Kumar V. et al, DOI 10.1007/s00281-012-0312-1 artikulutik moldatua ⁷⁸)

3.2.3. ImmunoChip proiektua

Asoziatu-rik aldaera berriak identifikatzeko xedez gauzaturiko azken proiektu erraldoia *ImmunoChip* izena jaso zuen. Lan honetan, sistema immunearekin erlazioaturiko bestelako gaixotasun batzuk pairatzen zituzten paziente askoren laginez gandi, CD zuten 12.000 lagun baino gehiago eta 12.000 kontrol inguru analizatu ziren, 200.000 aldaera baino gehiago kontuan harturik ⁷⁹.

Azterketa zabal honetan, HLAtik kanpoko 183 eremu aztertu ziren, sistema immunearekin erlazioatuta zeudenak. Hauetatik, 36 CDrekin asoziatuta zeudela ikusi zen (aurretiaz

GWASetan aurkitutako eremuak gehi beste 13 lokus berri) (5. irudia). Asoziaturiko polimorfismo guztiak arruntak ziren eta %5 baino maiztasun handiagoa aurkezten zuten alelo txikian. Frekuentzia baxuko aldaeren asoziazioa lau eremutan baino ez zen topatu. Lan honetan, eremu interesgarrien mapaketa zehatza egin zen, aldaera etiologikoak aurkitzeko asmoz.

Asoziaturiko SNPen anotazio funtzionalaren ostean, ondorio nagusietako bat eremu kodetzaillean kokatzen diren aldaerak oso urriak direla izan zen, nahiz eta haietako asko zenbait generen transkripzio hasiera guneeetatik gertu zein 3'UTR eremuetan egoten diren. Asoziazio seinaleen eta eremu erregulatzailleen arteko gertutasuna kontuan hartuta proposaturiko hautagai batzuk honakoak izan ziren: *THEMIS/PTPRK*, *TAGAP*, *ETS1*, *RUNX3* eta *RGS1*. Hauetako batzuk GWAS ostean aurkitutako multzoan zeuden.

3.2.4. GWAS azterketen jarraipena

CDren asoziazio genetikoak sendoak eta populazioen artean konparagarriak diren frogatzeko, GWAS azterketetan lortzen diren asoziazio emaitzak errepikatzea beharrezkoa da. Horregatik, CDrekin asoziatutako eremuak zenbait populazio independenteetan aztertu izan dira, emaitza kontraesanak lortuz.

CD GWAS emaitzen lehenengo jarraipen azterketa Italiar populazioan burutu zen 2009 urtean, non 538 CD gaixo eta 593 kontrol aztertu ziren. Lehenengo GWASean identifikaturiko 8 lokusetatik, italiar populazioan 4 lokus aurkeztu zuten CDrekiko suszeptibilitatea (*IL2/IL21*, *RGS1*, *IL12A/SCHIP1*, *SH2B3*), 2 eremuek asoziazio ahula (*LPP*, *TAGAP*) eta beste 2 eremu inolako asoziaziorik gabe (*CCR3*, *IL18RAP*)⁸⁰.

Honetaz gain, kohorte italiarrean erreplikatu ezin izan ziren 2 lokusak (rs917997-2q12 eta rs6441961-3p21) espainiar populazio batean aztertu ziren urte berean. Kasu-kontrol azterketa honetan 722 CD paziente eta 794 kontrol ez zeliako aztertu ziren, eta 3p21 eremu genetikoaren asoziazioa konfirmatu zen⁸¹.

Antzeko hurbilketarekin, HLAtik kanpoko GWASeko top 1.020 SNPak genotipatu eta aztertu ziren EEBBetako 906 CD gaixo eta 3.819 kontrolarekin. GWASean identifikaturiko 8 eremuetatik 5 CDrekin asoziazio aurkeztu zuten, 1q31, 3q25, 3q28, 4q27 eta 12q24. Asoziaziorik sendoena 4q27 (*IL2/IL21*) eremuan aurkitu zen, GWAS eta jarraipen azterketetan ere asoziazio altueneko eremua zelarik. Azterketa honetako bigarren

asoziazio seinalerik sendoena 3q28 eremuan aurkitu zen, *LPP* genetik gertu. Honetaz gain, ikerketa honek aurretik proposatu ez zen beste asoziazio eremu batentzako ebidentziak aurkeztu zituen, 2q31 eremua, non dagoeneko gene hautagaia den *ITGA4* genea aurkitzen den ⁸².

2010 urtean, aldeztu aurretik *IL2/IL21* eremuan asoziazioa aurkeztu zuen Suedia/Norvegiako populazio bat aztertu zen, baina oraingo honetan familia kohorte baten burutu zen transmisio desoreka testa, populazio estratifikazioak eragindako emaitza postibo faltsuak ekiditen dituen ⁸³. 325 familia zeliakotan genotipatu ziren asoziatuak 9 SNP. Hauetatik, 5 SNP (rs2816316-1q31, rs6441961-3p21, rs17810564/rs9811792-3q25-26, rs1464510-3q128) asoziazio esanguratsua aurkeztu zuten, SNP batek (rs917997-2q11-12) esangarritasunaren mugan geratu zen, eta gainontzeko 2 SNPetan (rs13015714-2q11-12, rs1738074-6q25) ez zen asoziaziorik aurkitu. 12q24 eremuan dagoen rs3184504 (*SH2B3*) ez zen aztertu genotipaketak ez zuelako ondo funtzionatu.

Kontraesanezko emaitza guzti hauek, erreplikatu daitezkeen emaitza sendoak lortzeko lagin multzo handiak aztertearen garrantzia eta beharra indartzen dute. Hala ere, egindako azterketa guztien emaitza ikusita argi dago seinalen indarra erlatiboki txikia dela, eta gehienetan, asoziazio esanguratsuak detektatzeko ahalmena populazio batetik bestera oso gutxi aldatzen diren aleoen frekuentzien menpe dago, eta honek, lortutako emaitza negatiboen azalpena izan zitekeen.

3.2.5. GWAS emaitzen azterketa funtzionalak

Asoziazio azterketa genetikoen beste helburu nagusia, gaixotasunen patogenesia ulertzen lagunduko diguten erantzule funtzionalak identifikatzea da. Gene hautagaien ustezko rola aztertzeko, GWAS osteko adierazpen azterketak oso erabilgarriak izan ohi dira.

Gure taldean burututako lan batean, lokus berean kokaturiko bi generen (*PTPRK* eta *THEMIS*) arteko koadierazpena ikusi zen, gaixo aktiboetan zein GFD trataturiko pazienteen biopsiak gliadinarekin in vitro kitzikatu ondoren. Lan honetan ondorioztatu zen, eremu honetan asoziatuak SNPek inguruko geneen adierazpenean eragina izan dezaketela, baina eragin hori konstitutiboa izan beharrean, badirudi *PTPRK* eta *THEMIS*en erregulazioa dietaren glutenak sorturiko erantzun immunearen ondorio dela. Aurkikuntza honek adierazten digu, DNA sekuentzian kodetuta dauden mekanismo erregulatuak

arruntek gene desberdinen adierazpen patroiak kontrolatu ditzazketela, eta erregulazio hori estimulu immunogenikoen ondorioz piztu edo itzali daitekeela ⁸⁴.

Beste aldetik, gaixotasun konplexuetan azalpenik gabe jarraitzen duen herentziaren zati handi bat argitzen lagunduko duen hipotesi berria postulatu izan da orain dela gutxi. Hipotesi hau, "GWAS aldaera sintetiko arraro" bezala ezagutzen da, eta aldaera arrunten asoziazioaren atzean egongo ziren aldaera kausal arraroetan oinarrituta dago. Hala ere, orain dela gutxi burututako lan batean non GWASeko 25 arrisku generen exoi kodetzailleak sekuentziatu eta genotipatu diren jatorri kaukasiarreko EBko 41.911 biztanleetan (24.892 indibiduo 6 gaixotasun autoimmunerekin eta 17.019 kontrol), frogatu izan da lokus ezagunetan dauden eremu kodetzailleetan aurkitzen diren aldaera arraroek efektu mespretxagarria dutela eritasun autimmuneen suszeptibilitatean, CD barne ⁸⁵.

Azterketa guzti hauetatik, ikerketa funtzionalak burutzearen hala nola suszeptibilitate gene hautagaien aukeraketa arbitrarioa baztertzearen garrantzia ondorioztatu daiteke. Honetaz gain, genomaren erregulazio sistema konplexuaren atzean dauden elementuak identifikatzeko egin gabe dagoen lan guztia agerian usten dute, eta atea irekitzen diete etorkizunean burutuko diren beste hainbat ikerketei, non ikertzaileek azterketa exhaustiboak burutu beharko dituzten aldaera mota desberdinak (hala nola aldaera estrukturalak eta epigenetikoak) zein genoma ez kodetzaille zabala aztertzeko. Ikerketa hauen helburu nagusia, eritasun arrunten genetika konplexua eta GWAS azterketetan identifikatu diren gaixotasunekin asoziatuiko aldaeren efektua ulertzea izango da.

Helburuak

Lan honek, eritasun zeliakoaren patogenesia ulertzen laguntzeko asmoa duten bi helburu nagusi ditu:

- I. Genoma osoko asoziazio azterketetan CDrekin asoziatuta agertu diren aldaera genetikoak aztertzea espainiar populazioan.
 - a. GWASeko asoziazio emaitzak espainiar populazioan errepikatzea.
 - b. Aurretik deskribatu ez diren asoziazio seinale eta eremu berriak bilatzea espainiar populazioan.

- II. Proposatutako gene hautagaiak CDren garapenean duten inplikazioa aztertzea.
 - a. CD gene hautagaien adierazpena aztertzea gaixo zeliako eta kontrolen heste ehunean.
 - b. CD gene hautagaien adierazpena aztertzea gaixo zeliako eta kontrolen heste epitelioko zelula populazioetan, eta biopsia osoetan lortutako emaitzekin konparatzea.
 - c. Eritasunarekin asoziatutako aldaera genetikoek gene hautagaien gain izan dezaketen eragina zehaztea.
 - d. Koadierazpen azterketak burutzea CDn aldatuta dauden elementu erregulatuak amankomunak identifikatzeko asmoarekin.

Materialak eta Metodoak

1. Pazienteak

Eritasun zeliakoa erreklutamenduaren garaian indarrean zegoen ESPGHAN (Gastroenterologia, Hepatologia eta Elikadura Pediatrikoaren Elkarte Europarra) irizpideen arabera diagnostikatu zen, horretarako anti-gliadina eta andi-endomizio (EMA) edo anti-transglutaminasa (TGA) antigorputzen determinazioa eta heste meharreko biopsiaz baliatuz.

Gliadinarekiko esposizio kroniko edo epe luzekoak gaixo zeliakoetan daukan eragina aztertzeko, 3 indibiduo talde aztertu ziren:

- CD gaixo aktiboak: diagnostikatu berri diren CD pazienteak, klinikoki gaixotasun aktiboa (CDrekin asoziaturiko antigorputzekiko positibo, hesteetako biloxken atrofia eta kripten hiperplasiarekin) eta dietan glutena hartzen dutenak.
- CD gaixo tratatuak: gutxienez 2 urtetan glutenik gabeko dieta (GFD) jarraitu duten CD paziente normalizatuak (sintomarik gabe, antigorputz negatiboak eta heste epitelio osasuntsua).
- Kontrol taldea: indibiduo ez zeliakoak eta endoskopiaren momentuan hesteetako inflamaziorik aurkezten ez zutenak erabili dira kontrol moduan.

2. Etikari dagokion adierazpena

Burutu diren azterketa guztiak, Gurutzeta Unibertsitate Ospitaleko Etika Batzordearen onespina lortu duten ikerketa proiektuen parte dira. Lagin guztiak 2003-2015 urteen artean jaso dira errutinazko diagnosi endoskopiaren bidez, emaleen edota beren gurasoen baimen informatua lortu ostean.

3. SNPen genotipaketa

3.1. DNA laginak

3.1.1. GWAS erreplikazio azterketa

Azterketa honetan 1.094 CD paziente eta 540 kontrol osasuntsuen DNAk erabili dira. Laginak CEGEC (Consortio Español para el estudio de la Genetica de la Enfermedad Celíaca) bildumatik jaso izan dira. 950 DNA lagin, 475 CD eta 475 kontrol, Gurutzeta Unibertsitate ospitaletik eta Arabako Unibertsitate ospitaletik (Euskal Herria) jaso dira; 95 CD lagin Hospital San Juan de Dios-etik (Catalunya), 190 CD lagin Hospital Universitario de Valladolid-etik (Castilla Leon); 190 lagin, 95 CD eta 95 kontrol, Hospital Virgen del Camino-tik (Nafarroa); 190 CD lagin Hospital General de Asturias-etik (Asturias) eta 60 CD lagin CATLAB-etik (Catalunya) (1. taula).

1 taula: DNA laginak GWAS erreplikazio azterketan.

CEGEC zentruak	Erkidegoa	CD paziente	Kontrol
Gurutzeta Unibertsitate ospitalea	Euskal Herria	475	475
Araba Unibertsitate ospitalea			
Hospital San Juan de Dios	Catalunya	95	-
Hospital Universitario de Valladolid	Castilla León	190	-
Hospital Virgen del Camino	Nafarroa	95	95
Hospital General de Asturias	Asturias	190	-
CATLAB	Catalunya	60	-

3.1.2. Immunochipa

Immunochip proiektuan, 12.041 CD gaixo eta 12.228 kontrolen DNA genotipatu zen. Lagin hauetatik, 545 CD paziente eta 308 odol emaile osasuntsu CEGEC bildumari dagozkio (2. taula). Tesi honetan CEGEC bildumaren emaitzak azalduko dira.

2. taula: ImmunoChip lagin bilduma.

Populazioa	CD paziente	Kontrol
Erresuma Batua	7.728	8.274
Holanda	1.123	1.147
Polonia	505	533
Espainia-CEGEC	545	308
Espainia-Madrid	537	320
Italia	1.374	1.255
India	229	391
TOTALA	12.041	12.228

3.2. DNA erauzketa

3.2.1. GWAS erreplikazio azterketa

Gurutzeta eta Arabako Unibertsitate Ospitaletako DNA genomikoa izozturiko 150 µl odoletik erazi zen ABI PRISM™ 6100 Nucleic Acid Prep Station (Applied Biosystems, Foster City, CA, USA) erabiliz, erauzketa kitak dakarren eluzio soluzioan bersuspenditu zelarik. Odolaren liseriketa, zelulen lisia, purifikazioa eta garbiketa pausuak fabrikatzailearen protokoloa jarraituz burutu ziren. Beste zentruetako laginak bakoitzaren prozedura estandarrak jarraituz burutu ziren.

3.2.2. ImmunoChips

Gurutzeta eta Arabako Unibertsitate Ospitaletako DNA genomikoa izozturiko 200 µl odoletik erazi zen NucleoSpin Genomic DNA Blood kita (Macherey-Nagel, Düren, Germany) erabiliz eta ddH₂Otan bereseki zen fabrikatzailearen instrukzioak jarraituz. DNA Quant-it PicoGreen dsDNA erreaktiboaren bidez kuantifikatu zen (Invitrogen, Carlsbad, CA) eta kontzentrazioak 50 ng/ µl-tara doitu ziren, Biomek NXP Laboratory Automation Workstation sistema automatikoaz (Beckman Coulter, Fullerton, CA).

3.3. Genoma osoaren amplifikazioa

Gurutzeta Unibertsitate ospitaleko 480 paziente eta 384 kontrolen DNA laginetan, zeinetan DNA kantitate gutxi zegoen eskuragarri, genoma osoko amplifikazioa (WGA) burutu zen GenomiPhi V2 DNA Amplification Kita (QIAGEN GmbH, Hilden, Germany) erabiliz. Horretarako, DNA bero bidez desnaturalizatu eta DNArekin modu inespezifikoan lotzen diren random hexameroak dituen lagin bufferrean hoztu zen. DNA polimerasa, random hexamero gehiago, nukleotidoak eta gatzak dituen master-mixa gehitu eta gero amplifikazio isotermikoa burutu zen, 30°C-tan 1,5 orduz. Amplifikazioaren ostean entzimak inaktibatu ziren 10 minututan 65°Ctan inkubatuz.

3.4. SNPen aukeraketa

3.4.1. GWAS erreplikazio azterketa

Hunt et al.⁸⁶ identifikaturiko zazpi eremuetan kokatzen diren 10 SNP aukeratu ziren laginetan genotipatzeko: rs2816316 (1q31, *RGSI*), rs917997 eta rs13015714 (2q11-12, *IL18RAP/IL18RI*), rs6441961 (3p21, *CCR1/CCR3/CCR2*), rs17810546 eta rs9811792 (3q25-26, *IL12A/SCHIP1*), rs1464510 (3q28, *LPP*), rs6822844 eta rs13119723 (4q27, *IL2/IL21* eta *KIAA1109*), rs1738074 (6q25, *TAGAP*) eta rs3184504 (12q24, *SH2B3*) (3. taula).

3.4.2. Immunochipa

Immunochip proiektuko markatzaileen aukeraketa Trynka et al.-ren artikuluan dago azalduta⁷⁹. Kongsortzioak, 12 gaixotasun autoimmunitarioetan (tiroide eritasun autoimmunitarioa, espondilitis ankilosantea, Crohn-en gaixotasuna, eritasun zeliakoa, IgA defizientzia, esklerosi multiplea, zirrosi biliar primarioa, psoriasis, artritis erreumatoidea, lupus eritematoso sistemikoa, 1 motako diabetes eta colitis ultzeraduna) GWAS esangarritasuna ($p < 5 \times 10^{-8}$) lortu zuten 186 lokusetan dauden markatzaileak aukeratu zituen. Laburki, 1000 Genomes proiektuko fase pilotuan, CEU populazioa gaixotasun bakoitzarekiko asoziazioa aurkeztu zuten SNP esanguratsuetatik 0,1 cM-eko distantziara zeuden aldaera guztiak aukeratu ziren. Honetaz gain, CD, Crohn eta T1D-ean sekuentzaturiko eremu genomiko batzuk ere aztertu ziren Immunochip proiektuan.

3.5. SNPen genotipaketa

3.5.1. GWAS erreplikazio azterketa

SNPak genotipatzeko, TaqMan allelic discrimination assayak (Applied Biosystems, Foster City, California, USA) erabili ziren. Assay hauek, SNP espezifikoak detektatzeko PCR hasle bikoteaz gain bi MGB zonda alelo-espezikoz osatuta daude, bakoitzak markatzaile fluoreszente desberdin batekin markatua (3. taula). Genotipaketa fabrikatzailearen instrukzioak jarraituz burutu zen. Anplifikazioa PCR termozikladore arruntan egin zen, pre-eta post-anplifikazio fluoresentziak ABIPrism 7900HT sequence detection system (Applied Biosystems, Foster City, CA, USA) baten neurtu zirelarik. DNA laginak, Biomek NXp automated liquid handler (Beckman Coulter, Fullerton, CA, USA) robotarekin prestatu ziren. 384 putxutako plaketan prozesatu ziren eta plaka bakoitzean 4 kontrol negatibo kokatu ziren.

3. taula: TaqMan Genotyping assay zerrenda.

Gene hautagaia	Eremua	SNP	Assay ID
<i>RGS1</i>	1q31	rs2816316	C_15810686_10
<i>IL18R1</i>	2q11-12	rs13015714	C_31439507_10
<i>IL18RAP</i>	2q11-12	rs917997	C_345197_1
<i>CCR3</i>	3p21	rs6441961	C_26450554_10
<i>SCHIP</i>	3q25-26	rs17810546	C_32594555_10
<i>IL12A</i>	3q25-26	rs9811792	C_2936004_10
<i>LPP</i>	3q28	rs1464510	C_8239299_10
<i>KIAA1109</i>	4q27	rs13119723	C_26404981_10
<i>IL21</i>	4q27	rs6822844	C_28983601_10
<i>TAGAP</i>	6q25	rs1738074	C_2966098_10
<i>SH2B3</i>	12q24	rs3184504	C_2981072_10

3.5.2. **Immunochipa**

Immunochipeko laginak Illumina-ren protokoloa jarraituz genotipatu ziren laborategi desberdinetan (London, UK; Hinxton, UK; Groningen, The Netherlands; and Charlottesville, Virginia, USA). NCBI build 36 (hg18) mapeaketa erabili zen (Illumina manifest file Immuno_BeadChip_11419691_B.bpm).

Honetaz gain, Immunochip proiektuko 44 SNP esanguratsuenak Fluidigm Biomark dynamic array (48.48) eta SNPtype assayekin genotipatu ziren (Fluidigm Corp.) adierazpen emaitzak eta DNA eskuragarri zuten beste 26 laginetan. Hauetariko 8 lagin dagoeneko genotipatuta zeuden Immunochipeko lagin multzoan eta horri esker, genotipaketaren kalitate kontrol modura erabiltzea posible izan zen. Genotipaketa arazoak eman zituzten 3 lagin ikerketatik kendu behar izan ziren, amaierako lagin multzoa 23 laginetan utziz, 14 kontrol eta 9 paziente zeliako. Assayen diseinua Fluidigm Assay Design taldearen bitartez burutu zen. Zazpi itu SNPk ez zituzten assay diseinurako eskakizun teknikoak bete, itu SNPetik 20-30 baseetara zeuden alboko SNPen presentzia, %65 baino gehiagoko GC edukia edota SNP trialelikoak izanagatik. Zazpi SNP hauek sakonki aztertu eta gero, itu SNPen alboko SNPen maiztasun alelikoak kontutan edukita eta SNP trialeliko (rs61907765) bakarraren alelo bakoitzaren maiztasuna Ensembl-en aztertuta, 6 SNPren diseinu oztupoak alde batera ustea erabaki genuen. rs60215663 SNParen kasuan aldiz, bere alelo txikiaren maiztasuna alboko SNPen maiztasuna baino txikiagoa izanda, analisitik kanpo ustea erabaki zen (1. taula gehigarria).

3.6. **Datuen analisisa**

3.6.1. **SNP genotipoak**

Genotipoak SDS 2.3 software bertsioarekin zehaztu ziren eta alelo txikien maiztasunak 2x2 kontingentzia taulen bidez konparatu ziren. Kasu eta kontrolen frekuentzia aleliko eta genotipikoak Fisher's exact testa eta χ^2 testa erabiliz konparatu ziren (test genotipikoak modelo dominante eta errezesiboetarako) EPI_INFO v.6.0 programan (Centers for Disease Control and Prevention, Atlanta, Georgia, USA).

3.6.2. Immunochiparen analisi estatistikoak

Immunochipeko lagin osoaren kasu-kontrol asoziazio analisiak PLINK v1.07 programaren bidez burutu ziren. Grafikoak LocusZoom39⁸⁷ erabiliz eraiki ziren.

Honetaz gain, CEGEC lagin multzoa independenteki aztertu zen kasu-kontrol asoziazioa aztertuz, populazioarekiko espezifikoa izan daitezkeen asoziazio seinaleak identifikatzeko helburuarekin.

4. Gene hautagaien analisi funtzionalak

4.1. Biopsiak

Pazienteen duodenoko biopsiak pediatria gastroenterologoek burututako prozedura kliniko estandarrak jarraituz lortu ziren; biopsiaren zati bat diagnostikoan erabili zen eta beste zatia ikerketa honetan. Ume zeliakoen diagnostiko biopsiak ume berdinak 2 urte baino denbora luzeagoz GFD tratamendua egin eta gero lortutako biopsiekin konparatu ziren, baita inflamaziorik gabeko kontrol ez-zeliakoen biopsiekin ere. Aztertutako talde bakoitzaren deskribapen sakonagoa pazienteen atalean aurki daiteke.

4.1.1. GWAS erreplikazio azterketak

29 ume zeliakoetan diagnostiko garaian (18 neska/11 mutil) eta paziente berdinetan 2 urte baino denbora luzeagoz GFD egin eta gero lortutako biopsiak konparatu ziren. kontrol modura 8 ume ez zeliakoen biopsiak erabili ziren.

4.1.2. Immunochipa

15 umeren hesteetako biopsia bikoteak (diagnostiko eta GFD >2 urte) eta kontrol ez-zeliakoen veste 15 biopsia lagin aztertu ziren.

4.2. Biopsietako zelula populazioak

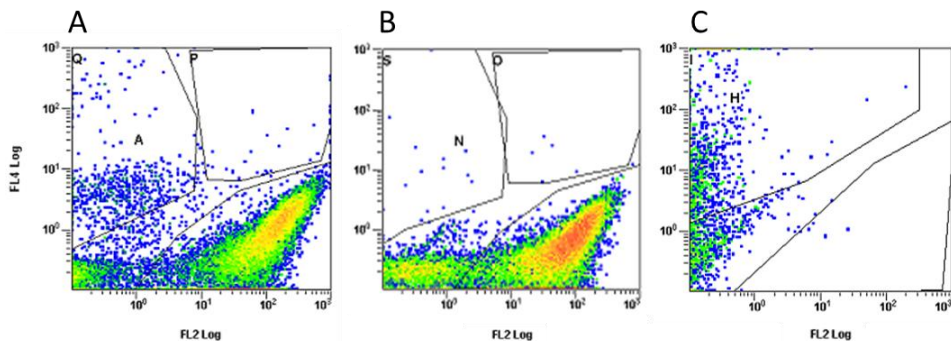
Hesteetako biopsietan agertzen diren bi zelula populazio nagusiak enterozitoak eta immunitate zelulak dira. Enterozitoak zelula CD326 positiboak dira, honek esan nahi du zelula hauek haien azalera zelula adhesio epitelialerako molekula (Epithelial cell adhesion molecule (EPCAM) adierazten dutela, zelula seinalizazioan, migrazioan, proliferazioan eta diferentziazioan parte hartzen duen molekula. Beste aldetik, immunitate zelulek CD45 antigenoa adierazten dute haien gainazalean, proteina tirosina fosfataseen familiako *PTPRC* geneak kodetzen duen proteina.

Ezaugarri hauek, biopsia laginetatik bi zelula populazioak banatzea ahalbidetzen dute, eta honi esker bi zelula moten azterketa independenteak egin daitezke hurrengo analisisietan.

Helburu horrekin, MACS zelula banaketa magnetikoa erabili izan da (Miltenyi Biotec), CD45 MicroBeads eta MS banaketa zutabeentzako fabrikatzailearen protokoloa jarraituz. Laburki, biopsiak jaso eta jarraian prozesatu ziren. Zelulak modu mekanikoan banatu ziren medio osoan (RPMI + antibiotikoak + FBS + DTT + EDTA) agitazio bidez ordu baten. Mediora askatutako zelulak zentrifugazio bidez jaso ziren medioa aurre-banaketa filtroetatik (20 µm) pasatu eta gero. Desegin gabe geratzen zen lamina propria RLT bufferrean gorde zen -80 °C-tan.

Zelula bizien suspentzio homogeneoa lortzeko, zelula hilak kentzeko kita (Dead cell removal kit, Miltenyi Biotec) erabili zen lehenengo purifikazio pausu batean. Bigarren purifikazio pausuan, zelulak CD45 microbead magnetikoekin markatu ziren zelula nahastea bi frakzioetan banatu ahal izateko, CD45 positiboak diren immunitate zelulak (zutabeen atxikituta geratuko direnak) eta zelula CD45 negatiboak edo enterozitoak (zutabea zeharkatuko dutenak). Lortutako bi frakzioak RLT bufferrean gorde ziren -80 °C-tan azido nukleikoak erauzi ziren arte.

Bi zelula frakzioen purutasuna konfirmatzeko helburuarekin, zitometroaren bidez aztertu ziren (6. irudia).



6. irudia: Biopsia osoetatik separatutako zelula populazioen zitometria emaitzak. X ardatzean dagoen FL2 anti-CD326 antigorputzari dagokio eta Y ardatzean dagoen FL4 anti-CD45 antigorputzari. A: Zelula nahasketa separazioa baino lehen. B: CD45- eta CD326+ zelula frakzioa. C: CD45+ eta CD326- zelula frakzioa.

4.3. RNA erauzketa

4.3.1. GWAS erreplikazioa azterketa

Biopsiak erabili arte nitronego likidoan gorde ziren. Izozturiko laginak plastikozko kirtenekin (Kontes, Vineland, NJ, USA) apurtu ziren 1,5 ml-ko mikrozentrifuga saioidietan. Tesi honen hasieran, biopsiak QIAshredder zutabetan (QIAGEN GmbH, Hilden, Germany) homogeneizatu ziren eta RNA isolatzeko RNeasy Micro Kit-a (QIAGEN) erabili zen, RNA DNase I-ekin tratatu eta -80°C -tan gorde zen erabili arte. RNA, UV espektrofotometriaz kuantifikatu zen Nanodrop espektrofotometroaren bidez (Thermo Scientific, Waltham, MA) eta RNA lagin guztien kontzentrazioa $8\text{ ng}/\mu\text{l}$ -tara doitu zen.

4.3.2. Immunochipa

microRNA aztertzeko esperimentu berriak diseinatzen hastean, Immunochipeko gene hautagaiak aztertzeko erabili zen RNA NucleoSpin microRNA kit-arekin (Macherey-Nagel) purifikatu zen fabrikatzailearen intrukzioak jarraituz. RNA txikia eta handia frakzio bakar batean jaso ziren (RNA totala) eta -80°C -tan gorde zen erabili arte. RNA Nanodrop espektrofotometroaren bidez (Thermo Scientific, Waltham, MA) kuantifikatu zen eta RNA lagin guztien kontzentrazioa $8\text{ ng}/\mu\text{l}$ -tara doitu zen.

4.4. Gene adierazpena: RT-PCR

4.4.1. GWAS erreplikazioa azterketa

GWASeko gene hautagaien adierazpena RT-PCR bidez kuantifikatu zen. Markatu gabeko bi hasle eta FAM bidez markaturiko MGB zunda batez osaturiko adierazpen assay komertzialak erabili ziren (Applied Biosystems, Foster City, CA, USA) (4. taula). *RPLPO* (large ribosomal protein) gene endogenoaren adierazpena aldi berean kuantifikatu zen esperimentu bakoitzean eta jarritako RNA kantitatearen kontrol modura erabili zen (VIC-rekin markatutako MGB zunda). RT-PCR erreakzio hirukoitzak burutu ziren 7900 Real Time PCR System (Applied Biosystems, Foster City, CA, USA) batean, 8 ng RNA eta erreakzio bakarreko entzimen nahastura (QuantiTect Probe RT-PCR kit - QIAGEN) erabilita. Gene bakoitzaren adierazpen erlatiboa, Ct metodo zehatza erabilita kalkulatu zen ³⁴.

4. taula: TaqMan Gene Expression assay-en zerrenda GWASeko gene hautagaien adierazpen analisirako.

Gene sinboloa	Assay ID
<i>RGS1</i>	HS00175260
<i>IL18R1</i>	HS00175381
<i>IL18RAP</i>	HS00187256
<i>SCHIP</i>	HS00205829
<i>IL12A</i>	HS00168405
<i>LPP</i>	HS00194400
<i>KIAA1109</i>	HS00361070
<i>IL21</i>	HS00222327
<i>TAGAP</i>	HS00611823
<i>SH2B3</i>	HS00193878

4.4.2. Immunochipa

RNA totala 8 ng/ μ l-tan normalizatu eta cDNA-ra bihurtu zen AffinityScript cDNA Synthesis kit-a (Agilent Technologies, Santa Clara, CA, USA) erabiliz, fabrikatzailearen protokoloa jarraituz. Geneen adierazpen analisia Fluidigm Biomark 48.48 dynamic array-ak (Fluidigm Corp., South San Francisco, CA, USA) eta TaqMan Gene Expression assay komertzialak (5. taula) erabiliz burutu zen. *RPLPO* gene endogenoaren adierazpena aldi berean kuantifikatu zen esperimentu bakoitzean eta jarritako RNA kantitatearen kontrol modura erabili zen. Gene bakoitzaren adierazpen erlatiboa, Ct metodo zehatza erabilita kalkulatu zen ³⁴ eta adierazpen emaitzak 15 kontrol laginen medierekiko normalizatu ziren. Gene adierazpenaren emaitzak publikoak dira eta Gene Expression Omnibus data gordailuan (<http://www.ncbi.nlm.nih.gov/geo/>) eskuragarri daude GSE61849 erregistro zenbakiarekin.

5. taula: TaqMan Gene Expression assay zerrenda Immunochip-eko gene hautagaien adierazpen analisirako.

Gene Sinboloa	Assay ID	Gene Sinboloa	Assay ID
<i>ADAD1</i>	Hs00330122_m1	<i>PVT1</i>	Hs01069044_m1
<i>ARHGAP31</i>	Hs00393361_m1	<i>TMEM187</i>	Hs01920894_s1
<i>ATXN2</i>	Hs00268077_m1	<i>TNFSF18</i>	Hs00183225_m1
<i>BACH2</i>	Hs00222364_m1	<i>TREH</i>	Hs00389383_m1
<i>CCR1</i>	Hs00928897_s1	<i>TTC34</i>	Hs01128292_m1
<i>CCR2</i>	Hs00704702_s1	<i>UBASH3A</i>	Hs00957643_m1
<i>CD28</i>	Hs01007422_m1	<i>UBE2E2</i>	Hs00994287_m1
<i>CIITA</i>	Hs00172094_m1	<i>UBE2L3</i>	Hs00748530_s1
<i>CLK3</i>	Hs00999875_m1	<i>YDJC</i>	Hs00419214_g1
<i>CSK</i>	Hs01062585_m1	<i>ZFP36L1</i>	Hs00245183_m1
<i>DDX6</i>	Hs00898915_g1	<i>ZMIZ1</i>	Hs00393480_m1
<i>ELMO1</i>	Hs00404994_m1	<i>SOCS1</i>	Hs00705164_s1
<i>GLB1</i>	Hs01035168_m1	<i>FASLG</i>	Hs00181225_m1
<i>HCFC1</i>	Hs00232039_m1	<i>ITGA4</i>	Hs00168433_m1

<i>MMEL1</i>	Hs00364353_m1	<i>ICOS</i>	Hs00359999_m1
<i>OLIG3</i>	Hs00703087_s1	<i>CTLA4</i>	Hs03044418_m1
<i>PFKFB3</i>	Hs00998700_m1	<i>IRF4</i>	Hs01056533_m1
<i>PLEK</i>	Hs00950975_m1	<i>IRAK1</i>	Hs01018347_m1
<i>POU2AF1</i>	Hs01573371_m1	<i>STAT4</i>	Hs01028017_m1
<i>PRM1</i>	Hs00358158_g1	<i>PRKCQ</i>	Hs00989970_m1
<i>PRM2</i>	Hs04187294_g1	<i>LTF</i>	Hs00914334_m1
<i>PTPN2</i>	Hs00959886_g1	<i>CCR4</i>	Hs00747615_s1
<i>PUS10</i>	Hs00328708_m1	<i>ICOSLG</i>	Hs00323621_m1

CEGEC lagin multzoan asoziazioa aurkeztu zuten gene hautagaiak era berean aztertu ziren lagin multzo berean. TaqMan gene expression assay komertzialak 6. taulan adierazita daude.

6. taula: TaqMan Gene Expression assay zerrenda CEGECeko gene hautagaien adierazpen analisirako.

Gene Sinboloa	Assay ID
<i>BLK</i>	Hs01017452_m1
<i>CAPSL</i>	Hs00921468_m1
<i>FAM167A</i>	Hs00697562_m1
<i>GALC</i>	Hs01012300_m1
<i>GPR65</i>	Hs00269247_s1
<i>IL12RB2</i>	Hs00155486_m1
<i>IL23R</i>	Hs00332759_m1
<i>IL7R</i>	Hs00902334_m1
<i>PSMG2</i>	Hs00220315_m1
<i>SERBP1</i>	Hs00967385_g1
<i>UGT3A1</i>	Hs01014799_m1
<i>UGT3A2</i>	Hs04177793_m1

4.5. Datuen analisia eta estatistika

4.5.1. GWAS erreplikazioa azterketa

Geneen adierazpen mailen arteko desberdintasunak modu ez-parametrikotan aztertu ziren Wilcoxon matched-pairs rank testa (diagnostiko vs. tratatu) eta Mann Whitney testa (ez-zeliako vs. gaixoen bi taldeak) erabiliz InStat v.3.05 (GraphPad Software, Inc. La Jolla, CA, USA) software-an. 0,05 azpitik zeuden aldeko p balioak kontsideratu ziren esanguratsu.

4.5.2. Immunochipa

Geneen adierazpen mailen arteko desberdintasunak modu ez-parametrikotan aztertu ziren Wilcoxon matched-pairs rank testa (diagnostiko vs. tratatu) eta Mann Whitney testa (ez-zeliako vs. gaixoen bi taldeak). Kalkulu estatistiko guztiak GraphPad Prism 5 (GraphPad Software, La Jolla, CA, USA) programaren bidez burutu ziren.

4.6. Koadierazpenaren analisia

Immunochipeko gene hautagaien arteko koadierazpena Pearson-en korrelazioaren bidez kalkulatu zen. SNPen genotipoa eta geneen adierazpenaren arteko asoziazioa aztertzeko Merlin 1.1.2 software-a erabili zen⁸⁸. Asoziazioa modu independentean aztertu zen lagin talde bakoitzean, CD lagin bikoteetan duplikatuta agertuko liratekeen genotipoetatik eratorritako asoziazio faltsuak ekiditeko.

4.7. Genotipo-fenotipo korrelazioa

Geneen adierazpena eta SNPen genotipoen arteko erlazioa kalkulatzeko R korrelazio koefizientea (http://www.fon.hum.uva.nl/Service/Statistics/Correlation_coefficient.html) erabili zen adierazpen balioak eta SNPen genotipoak eskuragarria zituzten laginetan.

1. kapituluua:

Genoma osoko asoziazio azterketak CDn: erreplikazio azterketa
espainiako populazioan eta gene hautagaien adierazpen analisia.

1. Aurrekariak

CDa arrisku genetikoa duten pertsonetan glutenarekiko sentikortasunaren ondorioz garatzen den gaixotasun kroniko eta inflamatorioa da, biztanle kaukasiarren %1 inguruk jasaten dutela kalkulatu delarik ^{32,89}. Arrisku genetikoaren zatirik handiena 6p21 eremu kromosomikoan kokatuta dagoen HLA-ri dagokio. Gaixoen %90 baino gehiagok HLA-DQ2 heterodimeroa adierazten dute eta gainontzekoek HLA-DQ8 molekula. Hala ere, HLA-DQ2 aldaera ohikoa da populazio orokorrean, gutxi gorabehera kaukasiarren %30ean agertuz. HLAk CDren herentziaren %40 azaldu dezakeela kalkulatu da, eta esfortzu handia egin izan da CD garatzeko arriskuan eragina izan ditzaketen bestelako lokusen bilaketan ^{37,90,91}.

Bestelako gaixotasun konplexuetan bezala, CDn ere GWAS azterketak burutu dira. Mota honetako lehenengo azterketan, non Erresuma Batuko 778 CD gaixo eta 1422 kontrol aztertu ziren, 4q27 eremu kromosomikoan asoziazio indartsua zeukan 500kb-ko lotura desoreka (LD) blokea identifikatu zen ⁷¹. Eremu honetan, T zelulen desberdintzapen eta proliferazio prozesuetan parte hartzen duten *IL2* eta *IL21* zitokinak kodetzen dituzten geneak daude, CDren patogenesia interesgarriak izan daitezkeen hautagai funtzionalak. Gainera, eremu honek 1 motako diabetesa eta artritis erreumatoidearekiko asoziazioa agertu du ere, autoimmunitatearekiko arrisku lokus orokorra izan daitekeela adieraziz ⁹².

Lehenengo azterketa horretan lortutako emaitzetatik abiatuta, identifikaturiko top SNPen jarraipen azterketa bat burutu zen europar jatorriko lagin multzo batean (Holanda, Erresuma Batua eta Irlanda) eta 7 arrisku eremu berri identifikatu ziren ⁷⁶. Hauetariko zenbait lokusetan erantzun immunitarioarekin erlazioaturiko geneak daude, *RGS1* (1q31), *IL18R1/IL18RAP* (2q11e2q12), *CCR3* (3p21), *IL12A/SCHIP1* (3q25e3q26), *TAGAP* (6q25), eta *SH2B3* (12q24) besteak beste. Kasu horretan ere, gene batzuk T1D eta Crohn gaitzarekin asoziazioa erakutsi dute, gene hauek gaixotasun immunitarioetan izan dezaketen inplikazioa indartuz ^{93,94}.

Orain dela gutxi, bigarren GWAS eta erreplikazio azterketak burutu dira 9000 kasu eta 15000 kontrolatan, eta genoma osoko asoziazio esanguratsua 13 lokus berrietara zabaldu da, gehienetan erantzun immunitarioarekiko (*BACH2*, *CCR4*, *CD80*, *CIITA-SOCS1-CLEC16A*, *ICOSLG*, and *ZMIZ1*) edo T zelulen desberdintzapen prozesuekin (*ETS1*, *RUNX3*, *THEMIS*, and *TNFRSF14*) erlazioaturiko geneak agertzen direlarik ⁷⁷. Kasu batzuetan,

identifikaturiko SNPek geneen adierazpenean duten eragina aztertu izan da odolean, eta zenbait korrelazio aurkitu da *cis* geneen adierazpenarekin.

GWAS azterketek CDren osagarri genetikoaren ikuspuntua zabaldu dute, eta erakutsi digute sarkortasun txikiko aldaera genetiko askok izan daitezkeela CD garatzeko arrisku genetikoaren erantzule. Hala ere, polimorfismo bakoitzaren ekarpen indibiduala txikia da (OR < 1.5), kasu eta kontrolen arteko mazitasun alelikoen diferentziak gutxitan iristen direlako %10era. Honetaz gain, GWASek lagin multzo ez uniforme eta tamaina desberdinekoak erabiltzearen ondorioz, populazioen arteko diferentziak diluituta geratzen dira askotan, lagin talde bakoitzak asoziazio seinale orokorrean ekarpen desberdina duelako. Gainera, garrantzitsua da gogoratzea asoziazio azterketek seinalearen lokalizazioa ez besterik emateko gaitasuna dutela, hortik aurrera burutzen den gene hautagaien aukeraketa subjektiboa izan ohi da eta gehienetan aurretik daukagun gaixotasunaren patogenesiaren informazioa kontuan edukita hartzen dira erabakiak, gene horien partaidetza funtzionalak indartu ditzakeen inolako emaitza esperimentalik gabe.

Gai hau ikertzeko asmoarekin, CDn egindako lehenengo GWASean identifikaturiko 8 lokusen erreplikazio azterketa egin dugu espainiar populazioan. Honetaz gain, proposaturiko 10 gene hautagaien adierazpena aztertu dugu, CDren garapenean gene hauek izan dezaketen partaidetza ikertzeko, baita SNPen genotipoek adierazpenean duten eragina aztertzeke ere.

2. Metodoak

Eritasun zeliakoa, erreklutamenduaren garaian indarrean zegoen ESPGHAN irizpideen arabera diagnostikatu zen, horretarako anti-gliadina eta anti-transglutaminasa antigorputzen determinazioa eta heste meharreko biopsiaz baliatuz. Ikerketa honek, CEGECen partaide diren instituzio guztien etika batzordeen onarpena zeukan.

DNA laginak odoletik lortu ziren erauzketa metodo arrunten bidez. Lehengo GWASean deskribatutako 8 eremu asoziatuak aztertzeko 11 SNP genotipatu ziren 1094 kasu eta 540 kontrol osansutsuetan (CEGEC bilduma) TaqMan allelic discrimination assay komertzialak erabiliz (Applied Biosystems, Foster City, California, USA) ABI7900HT sekuentzia detekzio sisteman (Applied Biosystems).

Genotipoak SDS 2.3 software bertsioarekin zehaztu ziren eta alelo txikien maiztasunak 2x2 kontingentzia taulen bidez konparatu ziren, c2 testa burutuz EPI_INFO v.6.0 programan (Centers for Disease Control and Prevention, Atlanta, Georgia, USA). Azterketa honetan lortutako emaitzak aurretik egindako beste azterketa baten emaitzekin elkartu ziren asoziazioen metanalisi burutzeko (558 gaixo zeliako eta 465 kontrol, espainiar jatorrikoak eta GWASaren erreplikazio azterketan erabiliak ⁷⁷).

RGS1, *IL18RAP*, *IL18R1*, *IL12A*, *SCHIP1*, *LPP*, *IL21*, *KIAA1109*, *TAGAP* eta *SH2B3* gene hautagaien adierazpena alderantzizko transkripzio PCR (RT-PCR) bitartez neurtu zen TaqMan gene expression assay komertzialak (Applied Biosystems) eta erreakzio bakarreko entzima nahastea erabiliz (QuantiTect Probe RT-PCR kit, Qiagen, Hilden, Germany). RNA laginak hesteetako biopsietatik abiatuz erauzi ziren, 29 ume zeliakoetan diagnostiko garaian (18 neska/11 mutil, glutena daukan dieta jarraituz, CDari asoziatuako antigorputz positiboak, heste-biloen atrofia eta kripten hiperplasiarekin) eta paziente berdinetan 2 urte baino denbora luzeagoz GFD egin eta gero (asintomatiko, antigorputz negatiboak, heste hepitelio normalizatu). 8 ume ez zeliakoen biopsiak erabili ziren kontrol modura. Esperimetuak triplikatuan burutu ziren ABI7900HT sisteman eta *RPLPO* genea neurtu zen beste geneekin batera hasierako RNA kantitatearen kontrol endogeno modura. Lagin bakoitzaren adierazpen erlatiboa kalkulatzeko Ct metodo zehatza erabili zen eta adierazpen balioa kontrol lagin batekiko normalizatu zen aurretik deskribatu bezala ³⁴.

Lagin taldeen arteko adierazpen desberdintasunak Wilcoxonon zeinu frogaren ez-parametrik (diagnostiko vs tratatu) eta Mann Whitney U frogarekin (ez-zeliakoak vs gaixoen bi taldeak) analizatu ziren InStat v.3.05 (GraphPad Inc, La Jolla, California, USA)

erabiliz. Taldearen batz besterik $>3SD$ aldentzen ziren balioak akats metodologikotzat hartu ziren eta konparaketa estatistikoetatik kanpo utzi (lagin bat tratatuen taldean *SCHIP1* generako eta diagnosi taldeko paziente bat *KIAA1109*ren kasuan). Biopsia bakoitzetik lortu daitekeen RNA kantitate mugatua dela eta, gene bakoitzeko 15 biopsia bikote aztertu ziren, SNP bakoitzeko 3 genotipo posibleak uniformeki irudikatuta agertzea saiatur.

3. Emaitzak eta eztabaida

Azken aldian eritasun zeliakoan burututako GWASak, CDraiko arrisku genetikoaren eman dezaketen asoziatu-eremu genomikoen bilduma zabala eskeini du^{76,77,91}. Hala ere, CDren osagarri genetikoaren sendoa eta populazioen artean konparagarria den zehazteko asoziazio emaitzak errepikatzea beharrezkoa da. Honetaz gain, proposatutako gene eta aldaera genetikoaren partehartzea esperimentalki frogatu behar da ehun gaixoa edota zelula ereduak aztertuz. Lan honetan, lehenengo GWASean identifikaturiko 8 eremu genomikoen asoziazio azterketa egin dugu espainiar populazioan eta proposaturiko gene hautagaien azterketa funtzionalak burutu ditugu^{76,91}.

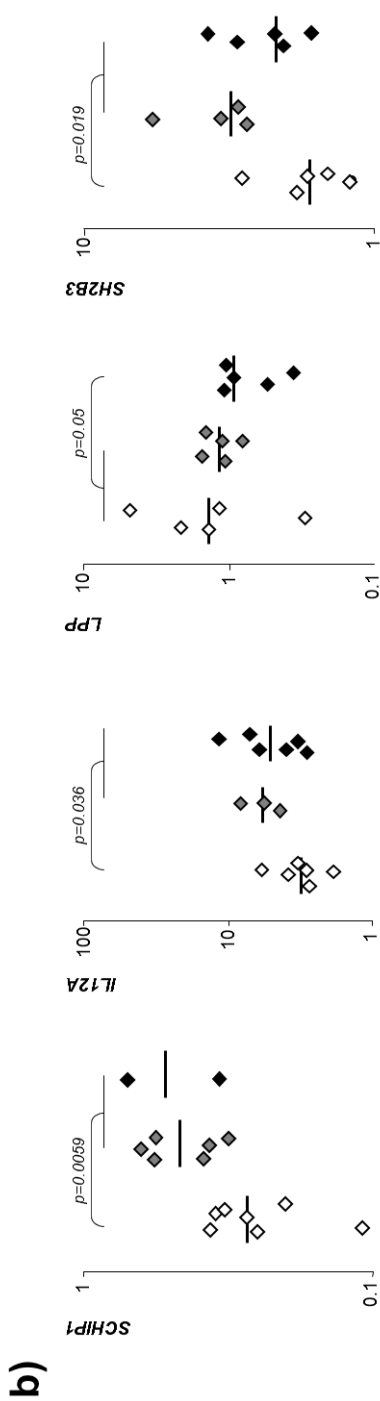
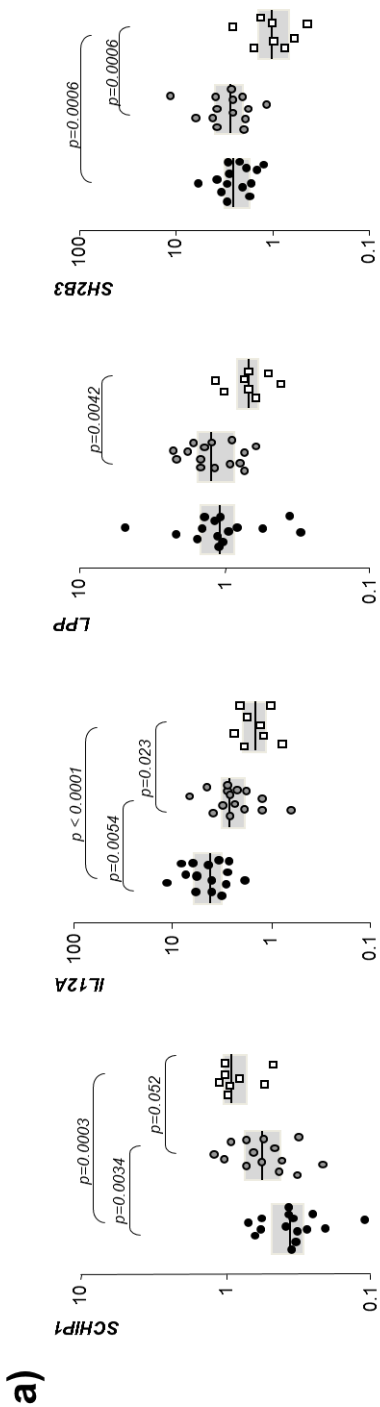
Zortzi eremu asoziatuetan dauden 11 SNP genotipatu eta gero, aztertu dugun espainiar populazioan CDrekiko asoziazio esanguratsua aurkitu dugu 3 eremuetan 1q31, 2q11-2q12 eta 3q25. Eremu hauetan agertzen diren gene hautagaiak, *RG51*, *IL18RAP/IL18R1* eta *SCHIP1* dira hurrenez hurren, eta erantzun immunitarioarekin erlazionatuta daude batez ere (7. taula). Gure genotipaketaren emaitzak aurretik aztertutako espainiar kohortekin konbinatuta, guk topaturiko 3 asoziazioak konfirmatzen dira, eta beste bi eremu berri esangura estatistikora heltzen dira: 3q28 (*LPP*), aurreko azterketan esanguratsu zena baina ez gure ikerketaren kasuan; 4q27 (*KIAA1109*), ez lehenengo azterketan ezta gurean ere esanguratsu ez zena. Beste aldetik, aurreko GWAS errepikapen azterketan espainiar kohortean 3p21 (*CCR3*) eremuan detektaturiko seinalea desagertzen da bi lagin sortak batuta.

Bere osotasunean, gure emaitzek gainontzeko azterketetan lortutako emaitzen antzekoak dira, eta orain arte eritasun zeliakoaren genetikari buruz dakiguna konfirmatuko lukete, errepikatu daitezkeen asoziazio emaitza indartsuak lortzeko lagin talde handiak aztertzearen beharri garrantzia emanik. Hala ere, bestelako azterketetan gertatzen den bezala, argi dago lortutako seinaleen indarra erlatiboki txikia dela eta gehienetan, asoziazio esanguratsuak detektatzeko potentzia populazioen arteko maiztasun aleliko desberdintasun txikien menpe dago eta lortutako zenbait emaitza negatiboen azalpena izan daiteke^{71,76,77,80}.

7. taula: GWAS azterketetan identifikaturiko SNPen asoziazio azterketa espainiar jatorriko kohortean eta, GWASaren erreplikapen azterketan publikatutako beste espainiar kohortean emaitzak⁷⁷.

Genea	SNP	CHR	BP*	Aleloak**	Espainako CEDEC			Espainako GWAs			Meta-analisia						
					MAF-A	MAF-U	p balioa	OR	MAF-A	MAF-U	p balioa	OR	MAF-A	MAF-U	p balioa	OR	KT
<i>RGS1</i>	rs2816316	1	190803436	A:C	0.1518	0.1994	0.0008	0.72	0.1536	0.1882	0.0421	0.78	0.1525	0.1944	0.0001	0.75	0.64-0.87
<i>IL18RAP</i>	rs917997	2	102437000	C:T	0.2806	0.2407	0.0159	1.23	0.2791	0.2691	0.6206	1.05	0.2801	0.2534	0.0357	1.15	1.01-1.31
<i>IL18R1</i>	rs13015714	2	102338297	G:T	0.2797	0.2273	0.0032	1.32	NA	NA	NA	NA	0.2797	0.2273	0.0032	1.32	1.09-1.60
<i>CCR3</i>	rs6441961	3	46327388	C:T	0.3307	0.3439	0.4538	0.94	0.3582	0.2875	0.0009	1.38	0.3400	0.3187	0.1152	1.1	0.97-1.24
<i>SCHIP1</i>	rs17810546	3	161147744	A:G	0.1197	0.0939	0.0288	1.31	0.1066	0.09469	0.3868	1.14	0.1153	0.0942	0.0358	1.25	1.03-1.52
<i>IL12A</i>	rs9811792	3	161179692	C:T	0.4936	0.4971	0.8511	0.99	NA	NA	NA	NA	0.4936	0.4971	0.8511	0.99	0.85-1.15
<i>LPP</i>	rs1464510	3	189595248	A:C	0.4495	0.4190	0.1108	1.13	0.4618	0.403	0.0090	1.27	0.4539	0.4116	0.0036	1.19	1.06-1.34
<i>IL2/IL21</i>	rs6822844	4	123728871	G:T	0.0950	0.1101	0.1786	0.85	NA	NA	NA	NA	0.0950	0.1101	0.1786	0.85	0.66-1.09
<i>KIAA1109</i>	rs13119723	4	123437763	A:G	0.0941	0.1122	0.1175	0.82	0.1082	0.1339	0.0804	0.78	0.0991	0.1222	0.0108	0.79	0.66-0.95
<i>TAGAP</i>	rs1738074	6	159385965	C:T	0.4229	0.4065	0.3812	1.07	0.4427	0.3995	0.0544	1.19	0.4297	0.4033	0.0646	1.11	0.99-1.25
<i>SH2B3</i>	rs3184504	12	110368991	C:T	0.4551	0.4341	0.2674	1.09	NA	NA	NA	NA	0.4551	0.4341	0.2674	1.09	0.93-1.27

Beste aldetik, asoziazio genetikoak bilatzen duten azterketen helburu garrantzitsua, gaixotasunaren garapenean parte hartzen duten errudun funtzionalen identifikazioa da, patogenesisia hobeto ulertzeko xede zuzenarekin. Gaixotasunaren bilakaeran izan dezaketen ustezko rola ikertzeko, GWASean proposaturiko gene hautagaien adierazpena aztertu zen gaixoen (diagnostikoan eta GFD tratamendu ostean) eta kontrol ez zeliakoen heste biopsietan. Aztertutako 10 geneetatik, lauk esanguratsuki gainadierazita zeuden gaixotasun aktiboan, kontrolekiko konparatuta: *IL18RAP*, *IL21*, *SH2B3* eta *IL12A*, T zelulen seinalizazioan eta Th1 edota Th17 erantzunen aktibazioan parte hartzen dute, CD aktiboan mukosaren inflamazioaren erantzuleak direlarik⁴⁷. Hala ere, CD aktiboan gene adierazpenaren alterazioa, neurri batean gutxienez, eritasunaren prozesuaren ondorio zuzena izan daiteke (adib. ingurune proinflamatorioa), eta ez jatorri genetikoarena. Hau horrela izanda, ezin izan dugu asoziatutako SNPen genotipoek gaixo aktiboan mukosan gainadierazita dauden *IL18RAP* eta *IL21* geneen adierazpenaren gain duten efektua detektatu. Bestalde, *SH2B3* konstitutiboki gainadierazita dago gaixoen mukosan, eritasunaren egoerarekiko modu independentean (gaixo aktibo zein tratatuetan), aldaera genetiko batek sortutako gaixotasunaren garapenean aurretiko akatsa izan daitekeela iradokiz (7.A irudia). Gainera, rs3184504*T aleloaren presentziak *SH2B3*ren adierazpen altuagoarekin asoziatuta dago CD gaixo aktiboan heste mukosan (7.B irudia). Polimorfismo hau, proteinaren domeinu garrantzitsu eta selekzio positiboaren menpe egon den R262W aminoazido aldaketa eragiten duen SNP ez-sinonimoa da, bakterioen aurkako erantzun eraginkorragoa ematen duelarik⁹⁵. Gure taldeak ez du aldaera genetiko honen asoziaziorik topatu espainiar kohortean, naiz eta bestelako talde europearretan topatu den^{76,91}. Antzeko egoera ematen da *IL2A*ren kasuan, GFD taldean gainadierazita agertzen dena, rs9811792*C aleloaren kopia bakarrak mRNA-ren ekoizpena igotzeko gai delarik. Honek, aldaera horren funtzio erregulatuak iradokiko luke, nahiz eta gure azterketan SNP horren asoziazio esanguratsurik topatu ez dugun (7. irudia).



7. irudia: GWAS azterketetan proposaturiko gene hautagaien adierazpen analisia heste-mukosan. Bakarrik erakusten dira GFD eta kontrol taldeen arteko ezberdintasun esanguratsuak aurkezten dituzten geneak. (A) CD gaixo aktibo (borobil beltzak), gaixo tratatu (borobil grisak) eta kontrolen (lauki zuriak) arteko konparaketa. (B) SNPen genotipoak geneen adierazpen erlatiboan duen eragina CD gaixo aktiboetan, babes alelorako homozigotoak (erronbo zuria), heterozigotoak (erronbo grisak) eta arrisku alelorako homozigotoak (erronbo beltzak). Kontuan eduki, arrisku aleloa SNP bakoitzerako CD gaixoetan maiztasun handiagoz agertzen den aleloa dela, ez du zertan alelo txikia izan behar.

Beste bi gene, *LPP* eta *SCHIP1*, biak gaixotasunaren prozesuan bete dezaketen funtzio argirik gabe, gaixo eta kontrolen arteko adierazpen aldaketak erakusten dituzte. Heste meharrean adierazpen altua aurkezten duen *LPP* genea zelulen arteko loturekin erlazionatu da, epitelioaren mantentzean erantzukizun estrukturala izan dezakeelarik ⁹⁶. Badirudi pazienteetan *LPP* aktibo dagoela kontrolekin konparatuta, ezberdintasun esanguratsuak GFD eta kontrol taldeen artean bakarrik aurkitzen badira ere, genetikak eragindako alterazio konstitutiboa izan daitekeenaren ideia indartuz (7.A irudia). Espainiar populazioan arrisku altuarekin asoziatuta dagoen rs1464510*A aleloaren presentziak espero zitekeenaren kontrako efektua dauka, rs1464510*AA genotipodun gaixo aktiboen biopsiek rs1464510*AC eta rs1464510*CC diren pazienteek baino *LPP* genearen mRNA maila baxuagoak baitituzte (7.B irudia). *SCHIP1* genearen kasuan, gaixoen biopsietan azpiadierazita badago ere, diferentzia handienak CD aktiboan ikusten dira. Hala ere, diferentzia hauek joera berdina aurkezten dute GFD taldea kontrolekin konparatzen dugunean. *LPP*ren kasuan bezala, rs17810546*G aleloak *SCHIP1*en adierazpenean duen efektua espero genezakeenaren kontrakoa da, hau da, CDrekin asoziatutako aleloaren presentzian mRNA kantitate baxuagoa espero den arren, arrisku aleloa duten gaixoen biopsietan *SCHIP1*en adierazpen altuagoa ematen da. *KIAA1109* genearen kasuan adierazpenaren jeitsiera ikusten da gaixo aktiboak kontrolekiko konparatzen ditugunean, baina ez dago asoziatutako rs13119723*A arrisku aleloak genearen erregulazioan parte hartzen duen ebidentziarik. *RGS1*, *IL18R1* eta *TAGAP* geneen kasuan ez da aztertutako biopsia taldeen arteko adierazpen aldaketarik aurkitu.

Lortutako emaitzetatik, asoziatutako eremuetan dauden SNPen genotipo eta gene hautagaien arteko erlazio funtzional ezberdinak daudela ondoriozta dezakegu, aldaera genetiko eta gaixotasunaren patogenesiaren arteko lotura ez dela zuzena frogatuz. Aldaera genetikoek inflamaziorik gabeko CD pazienteen adierazpen maila basalean eragina izan dezaketela kontuan hartuta, interesgarria da ikustea nola GFD eta kontrolen

arteko adierazpen mailen desberdintasunak mantentzen diren lau geneetan (*SCHIP1*, *IL12A*, *LPP* eta *SH2B3*) GWASean asoziatuiko SNParen efektua ikusi daitekeen. Guzti honek, gene hauek eta haiekin asoziatuiko aldaerek CDren patogenesiaren benetako partaide etiologikoak izan daitezkeela adieraz dezake. Hala ere, eritasun konplexuen biologia SNP batek eragindako funtzio aldaketa eta gaixotasunaren arteko erlazio zuzena baino askoz konplexuagoa da, eta kontuz ibili behar da soilik asoziazio eremuetan oinarrituta gene etiologikoak eta mekanismo patogenikoak proposatzen direnean. Azterketa funtzional zabalagoak (bai lagin tamainan baita asoziatuiko eremu bakoitzean aztertzen diren gene kopuruan ere) eta zelula edota animali ereduetan burututako ikerketa esperimentalak beharrezkoak dira GWAS emaitzen erabilgarritasuna argitzeko, azterketa hauek aldaera etiologiko garrantzitsuak identifikatzeko gai diren frogatu ahal izateko.

2. kapitulua:

Adierazpen azterketak heste mukosan argira ekartzen ditu CDn asoziatuturiko eremuen azpian dauden geneen arteko erlazio konplexuak

1. Aurrekariak

Eritasun zeliakoa (CD) suszeptibilitate genetikoa duten pertsonetan elikaduraren glutenarekiko intolerantziak eragindako ohiko gaixotasun immunitarioa da (1:110 prebalentzia). Heste meharreko mikrobilosken atrofia, kripten hiperplasia eta infiltrazio linfositikoa dira ezaugarri histologiko tipikoak, eta tratamendu eraginkor bakarra bizitza osorako glutenik gabeko dieta (GFD) egitea da ⁶². CD garatzeko arrisku gehien ematen duen lokusak MHC eremuan dago, 6p21 kromosoman, eta %40ko ekarpen genetikoa daukala kalkulatu da. Birtualki, CD paziente guztiak HLA-DQ2 edo HLA-DQ8 positibo dira ⁶⁵. Hala ere, HLA arrisku aldaerak CD garatzeko beharrezkoak badira ere, alelo hauek populazio orokorrean ohikoak direnez, CD garatzeko arriskua ematen duten bestelako lokus ezezagunak egon behar dira inplikaturak.

Egunerarte, bi Genoma osoko asoziazio azterketa (GWAS) egin dira CDn eta gaixotasuna pairatzeko arriskua eman dezaketen 26 eremu genomiko identifikatu dira ^{71,76,77}. Honetaz gain, orain dela gutxi beste 13 suszeptibilitate lokus berri aurkitu dira ImmunoChip genotipaketa array-aren bidez, non erantzun immunitarioarekin erlazionatuta dauden 12 gaixotasunetan (tiroidearen eritasun autoimmunitarioa, espondilitis ankilosantea, Crhonen eritasuna, CD, IgA defizientzia, esklerosi anizkoitza, behazun-zirrosi primarioa, psoriasis, artritis erreumatoidea, lupus eritematoso sistemikoa, 1 motako diabetesa eta ultzeradun kolitisa) aurretik genoma mailako esangarritasuna lortu zuten lokusetan aurkitzen diren markaitzaileak sakonki genotipatu diren ⁷⁹. CDn identifikaturiko lokus batzuk gainontzeko gaixotasun autoimmunitarioa edota gaixotasun kroniko immunitarioekin asoziatuta daude, batez ere 1 motako diabetesa ⁹⁷ eta artritis erreumatoidearekin ⁹⁸.

Eremu hauetako zenbait gene hautagai etiologiko modura proposatu da, gehienak aurretik erantzun immunitario edota T zelulen desberdintzapen prozesuarekin erlazionatu direlarik, eta CDren patogenesisaren etapa ezberdinetan parte hartu dezaketela proposatu da. Hala ere, asoziazio ikerketak arrisku lokusen kokapena adierazteko baino ez dira gai eta hortik aurrera egiten den gene hautagaien aukeraketa aprioristikoa da gehienetan eta CDren patogenesisian daukagun aurreiritzian oinarrituta egoten da, gene hautagaien zeregina frogatuko luketen esperimentu eta emaitza funtzionalik gabe, are gutxiago CD gaixoetan kaltetutako heste epitelioan egindako esperimentuak. Orain arte, CDn burututako eskala handiko azterketek CDrekin asoziatu-eremuetatik 57 seinale independente aurkitu dituzte HLAtik kanpoko 39 lokusetan ⁷⁹. Identifikaturiko eremuetatik, 29tan proteina bakarrik kodetzen duten geneak daude, baina seinale gehienak eremu intergenikoetan

kokatzen dira, arrisku gene posible bat baino gehiago edo agian oraindik deskribatu ez diren genomako elementu erregulatu-zaileak iradokiz. Guztira, 66 gene hautagai proposatu dira haien kokapenean oinarrituz (asoziatu-eremu- gailurren azpian), baina gaixotasunean kaltetzen den ehunean azterketa funtzionalak aurrera eramatea guztiz beharrezkoa da gene hauen eta asoziazio seinaleen mekanismoak ulertu ahal izateko.

Gure taldeak aurretik burutu-ako lan batean, lehenengo GWASean proposatu-eremu-ko 10 gene hautagaien adierazpena aztertu zen CD paziente eta kontrolen heste-eko biopsietan⁷⁶, asoziatu-eremu-ko SNPen genotipoak gene hauen adierazpenean izan dezakeen eragina aztertu eta CDren garapenean izan dezaketen inplikazioa zehazteko. Gaixotasunaren egoera desberdinetan adierazpen aldaketa jasaten duten zenbait gene identifikatu genituen, bai eta gene hautagaien eta SNPen genotipoen arteko zenbait erlazio funtzional aurkitu ere⁹⁹.

Oraingo lan honetan, proposatu berri diren gene hautagaiek CDren garapenean izan dezaketen inplikazio funtzionala aztertu da. Gaixotasunaren prozesuan izan dezaketen partaidetza ikertzeko, aurretik aztertu gabe zeuden eta funtzio ezaguna daukaten 45 gene hautagai (8. taula) aukeratu ziren haien adierazpena neurtzeko CD gaixo (diagnostikoan eta tratamenduaren ostean) eta kontrolen heste biopsietan ere. Honetaz gain, CDrekin asoziatu-eremu-ko aldaerek geneen adierazpenean eraginik duten aztertu da, immuno- chip proiektuko SNP esanguratsuen genotipoak kontutan hartuta gene hautagai bakoitzarekiko. Gainera, koadierazpen analisiak burutu dira gaixo zeliakoetan inflamazioaren ondorioz edota eragile genetikoaren suszeptibilitatea dela eta aldatuta egon daitezkeen elementu erregulatu-zaileak detektatzeko helburuarekin.

8. taula: Aztertutako gene hautagaiak eta asoziatutako SNPak.

dbSNP ID (release 138)	HGVS izena (GRCh38)	Gene proteina-kodetzaile hautagaiak
rs4445406	1:g.2607961T>C	<i>MMEL1</i> <i>TTC34</i>
rs12068671	1:g.172711891T>C	<i>FASLG</i>
rs859637	1:g.172741860T>C	<i>TNSF18</i>
rs13003464	2:g.60959694A>G	<i>PUS10</i>
rs10167650	2:g.68418428T>G	<i>PLEK</i>
rs1018326	2:g.181143073T>C	<i>ITGA4</i>
rs6715106	2:g.191048308A>G	<i>STAT4</i>
rs6752770	2:g.191108837A>G	
rs12998748	2:g.191083911G>T	
rs1980422	2:g.203745673C>T	<i>CD28</i>
rs34037980	2:g.203905331A>G	<i>CTLA4</i>
rs10207814	2:g.203595238C>T	<i>ICOS</i>
rs4678523	3:g.32996229T>C	<i>CCR4</i> <i>GLB1</i>
rs2097282	3:g.46336534C>T	<i>CCR1</i>
rs7616215	3:g.46164194C>T	<i>CCR2</i> <i>LTF</i>
rs61579022	3:g.119404431G>A	<i>ARHGAP31</i>
rs1050976	6:g.408079C>T	<i>IRF4</i>
rs12203592	6:g.396321C>T	
rs7753008	6:g.90099920T>C	<i>BACH2</i>
rs17264332	6:g.137684378A>G	<i>OLIG3</i>
rs77027760	6:g.137680924G>A	
rs79758729	7:g.37378851A>G	<i>ELMO1</i>
rs10808568	8:g.128251814A>C	<i>PVT1</i>

rs2387397	10:g.6348230G>C	<i>PFKFB3</i> <i>PRKCQ</i>
rs1250552	10:g.79298270A>G	<i>ZMIZ1</i>
rs7104791	11:g.111326133T>C	<i>POU2AF1</i>
rs10892258	11:g.118709156G>A	<i>TREH</i> <i>DDX6</i>
rs61907765	11:g.128522042C>T	<i>ETS1</i>
rs3184504	12:g.111446804T>C	<i>ATXN2</i>
rs11851414	14:g.68792785T>C	<i>ZFP36L1</i>
rs1378938	15:g.74804102T>C	<i>CLK3</i> <i>CSK</i>
rs6498114	16:g.10870261G>T	<i>CIITA</i>
rs243323	16:g.11267345A>G	<i>SOCS1</i>
rs80073729	16:g.11279940G>A	<i>PRM1</i>
rs9673543	16:g.11291099A>G	<i>PRM2</i>
rs11875687	18:g.12843138T>C	<i>PTPN2</i>
rs62097857	18:g.12857759G>A	
rs1893592	21:g.42434957A>C	<i>UBASH3A</i>
rs58911644	21:g.44209238A>T	<i>ICOSLG</i>
rs4821124	22:g.21625000T>C	<i>UBE2L3</i> <i>YDJC</i>
rs13397	X:g.153982797G>A	<i>HCFC1</i> <i>TMEM187</i> <i>IRAK1</i>

2. Materiala eta metodoak

2.1. Pazienteak eta biopsiak

CD ESPGHAN elkartearen orduko irizpideen arabera diagnostikatu zen, horretarako anti-gliadina, anti-endomisiko eta anti-transglutaminasa antigorputzen determinazioak eta heste meharreko biopsia bidez. Ikerketa lan honek Gurutzeta Unibertsitate Ospitaleko Etika Batzordearen (CEIC-E09/10) eta Euskal ikerketa kliniko eta Etika Batzordearen (PI2013072) onospena lortu zuen. Azterketa guztiak, emaileen edota beren gurasoen baimen informatua lortu ostean burutu ziren. Paziente bakoitzaren duodeno biopsia diagnostikorako errutinazko endoskopia prozesuan lortu zen.

Ikerkera honetarako lagin multzoa osatzen dute: 15 ume zeliakoaren diagnostikoaren uneko biopsia (glutena daukan dieta jarraituz, CDari asoziatutako antigorputz positiboak, heste-biloen atrofia eta kripten hiperplasiarekin) eta ume berdinei 2 urte baino denbora luzeagoz GFD tratamendua egin eta gero hartutako biopsia (asintomatiko, antigorputz negatiboak, heste hepitelio normalizatua), gehi 15 ume ez-zeliakoaren biopsiak kontrol modura (biopsia momentuan inflamaziorik ez zutenak). Biopsia hauetatik RNA osoa erauzi zen NucleoSpin microRNA kit-a (Macherey-Nagel, Düren, Germany) erabiliz eta bertan emandako protokoloa jarraituz.

2.2. RNA laginak eta gene adierazpena

RNA 8 ng/ul-ko kontzentrazioa normalizatu eta cDNA bihurtu zen AffinityScript cDNA Synthesis kit-a (Agilent Technologies, Santa Clara, CA, USA) erabiliz fabrikatzailearen protokoloa jarraituz. Geneen adierazpen analisiak Fluidigm Biomark 48.48 plataforman (Fluidigm Corp., South San Francisco, CA, USA) egin ziren TaqMan Gene Expression assay komertzialak erabiliz. *RPLPO* genea neurtu zen aldi berean RNAREN hasierako kontzentrazioaren kontrol endogeno modura (Life Technologies, Thermo Fisher Scientific Inc., Waltham, MA, USA). Lagin bakoitzaren adierazpen erlatiboa Ct metodo zehatzaren³⁴ bidez kalkulatu zen eta 15 kontrolen batzuekiko adierazpenarekiko normalizatu zen. Adierazpen emaitza guztiak publikoki eskuragarri daude Gene Expression Omnibus gordailuan (<http://www.ncbi.nlm.nih.gov/geo/>) GSE61849 atxikitze zenbakiarekin.

Lagin taldeen arteko adierazpen desberdintasunak Wilcoxon matched pairs rank test ez-parametrikotik (diagnostiko vs tratatu) eta Mann-Whitney U testarekin (ez-zeliakoak vs gaixoen bi taldeak) analizatu ziren. Koadierazpena Pearson korrelazioaren bidez kalkulatu

zen. Estatistika kalkulua guztiak GraphPad Prism 5 softwarearekin burutu ziren (GraphPad Software, La Jolla, CA, USA). Taldearen batzuetan besteko balioak $>3SD$ aldentzen ziren balioak akats metodologiko gisa kontsideratu ziren eta konparaketa estatistikoetatik kanpo utzi ziren.

2.3. SNPen genotipaketa

ImmunoChip proiektuko 44 SNP esanguratsuenak Fluidigm Biomark dynamic array (48.48) eta SNPtype assayekin genotipatu ziren (Fluidigm Corp.), adierazpen emaitzak eta DNA eskuragarri zuten 26 laginetan. Hautariko 8 lagin dagoeneko genotipatuta zeuden ImmunoChipeko lagin multzoan eta, genotipaketaren kalitate kontrol modura erabiltzea posible izan zen. Genotipaketa arazoak eman zituzten 3 lagin ikerketatik kendu ziren, amaierako lagin multzoa 23 laginetan utziz, 14 kontrol eta 9 paziente zeliako. Assayen diseinua Fluidigm Assay Design taldearen bitartez burutu zen. Zazpi SNP ez zituzten assay diseinurako eskakizun teknikoak bete, iturri SNPetik 20-30 baseetara zeuden alboko SNPen presentzia, %65 baino gehiagoko GC edukia edota SNP trialelikoak izanagatik. Zazpi SNP hauek sakonki aztertu eta gero, iturri SNPen alboko SNPen maiztasun alelikoak kontutan edukita eta SNP trialeliko (rs61907765) bakarraren alelo bakoitzaren maiztasuna Ensembl-en aztertuta, 6 SNPren diseinu oztopoak alde batera uztea erabaki genuen. rs60215663 SNParen kasuan aldiz, bere alelo txikiaren maiztasuna alboko SNPen maiztasuna baino txikiagoa izanda, analitika kanpo uztea erabaki zen.

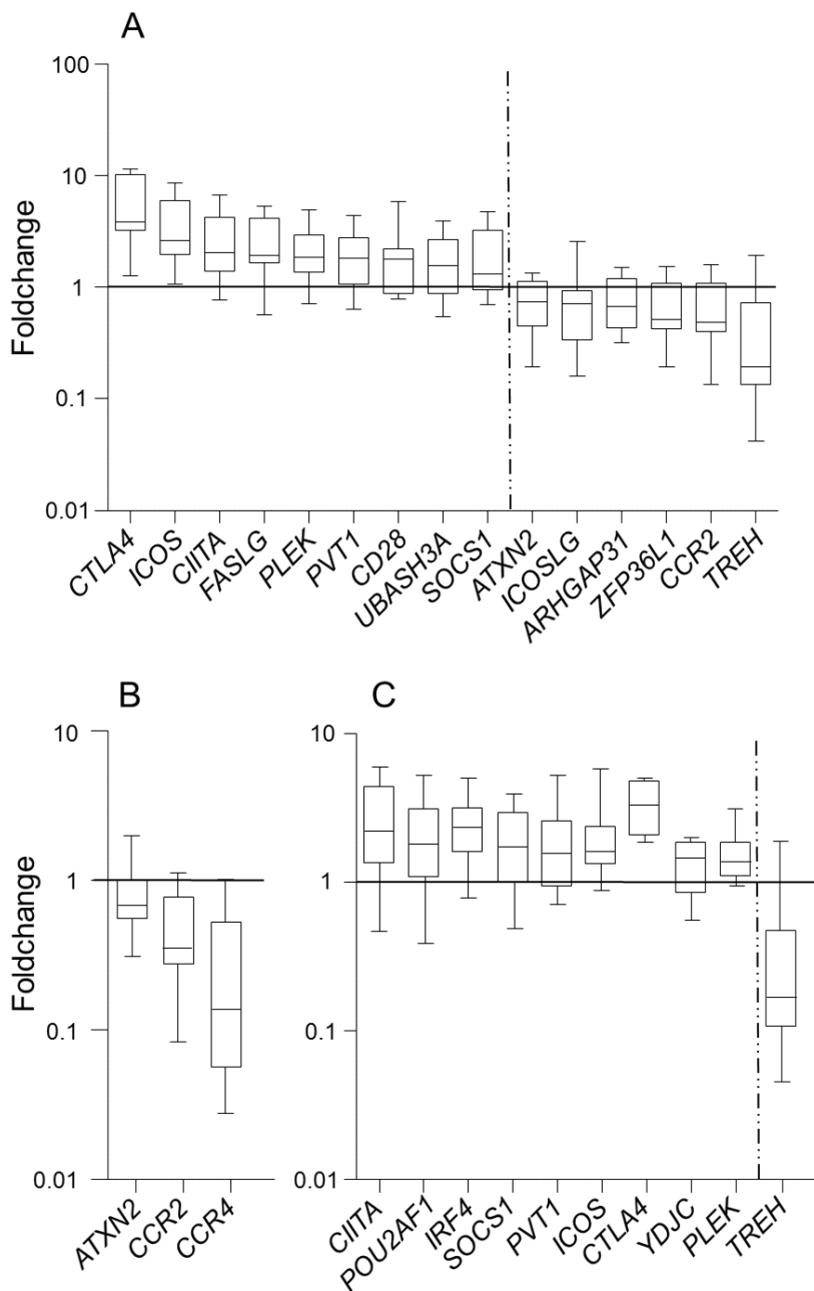
2.4. Koadierazpen analisiak

Aztertutako SNPen genotipo eta gene hautagaien adierazpenaren arteko asoziazioak ikertzeko Merlin 1.1.2 software-a erabili zen⁸⁸. Paziente taldean aztertutako biopsia bikoteen (debut vs tratatu) genotipo bikoiztuen ondorioz lortu zitezkeen asoziazio faltsuak ekiditeko, asoziazioa modu independentean aztertu zen ikerketa talde bakoitzean.

3. Emaitzak

3.1. CDn derberdin adierazten diren geneak

Gaixo zeliako aktibo eta kontrolen arteko arrazoen aldaketa konparatuta, aztertutako 45 geneetatik 15 genek adierazpen desberdina aurkezten dute. Bederatzi gene gaixotasun aktiboan modu esanguratsuan gainadierazita agertzen dira (*CTLA4*, *ICOS*, *CIITA*, *FASLG*, *PLEK*, *PVT1*, *CD28*, *UBASH3A* eta *SOCS1*) eta beste 6 azpiadierazita (*ATXN2*, *ICOSLG*, *ARHGAP31*, *ZFP36L1*, *CCR2* eta *TREH*) (8-A). Gene hautagaien aukeraketa aprioristikoaren ondorioz espero daitekeen bezala, aldatutako geneen GO-termino analisiak erantzun immunitarioarekin erlazionaturiko prozesuetan aberastuta dago, T zelulen erregulazioa, linfzito eta leukozitoen aktibazioa eta proliferazioa, eta linfzitoen koestimulazioa, besteak beste. GO-termino hauen atzean dauden generik esanguratsuenak honako hauek dira: *ICOSLG* (inducible T-cell costimulator ligand); *CCR2* (chemokine (C-C motif) receptor 2), espezifikoki monozitoen kemotaxian eragiten duen eta eritasun inflamatorioetan monozitoen infiltrazioan parte hartzen duen kemokinen hartzaila; *PLEK* (pleckstrin); *CTLA4* (cytotoxic T-lymphocyte-associated protein 4), T zelulen seinale inhibitzailea transmititzen duen proteina kodetzen duen immunoglobulinen superfamilaren kidea; *CD28*, T zelulen proliferazioa eta biziraupen, baita zitokina ekoizpen eta T-laguntzaile mota-2 zelulen garapenerako ezinbestekoa den proteina, eta *ICOS* (inducible T-cell costimulator), *CD28* eta *CTLA4* zelula gainazaleko hartzailen familiari dagokiona eta zelula-zelula seinalizazioan, erantzun immunitarioan eta zelulen proliferazioan erantzukizun garrantzitsua daukana. *ICOS*, *CD28* eta *CTLA4* geneak CELIAC3 lokusean daude, zenbait gaixotasun autoimmunitarioekin (CD barne) lotura daukan eremu ezaguna. Eremu hau Holopainen et al.¹⁰⁰ identifikatu zuten lehenengo aldiz, eta bestelako ikerketetan erreplikatu izan da behin baino gehiagotan. Paziente zeliako tratatuak eta kontrolak konparatuz, bakarrik 3 genek aurkezten dituzte adierazpen desberdintasun esanguratsuak (*ATXN2*, *CCR2* eta *CCR4*) eta hirurak daude gaixoen taldean konstitutiboki azpiadieraziak (8. irudia-B).



8. irudia: Adierazpen erlatiboaren arrazoiaren aldaketa (FC) adierazpen aldatua duten geneetan. (a) CD aktiboak vs kontrolak, (b) CD tratatua vs kontrolak eta (c) CD aktiboak vs CD tratatua.

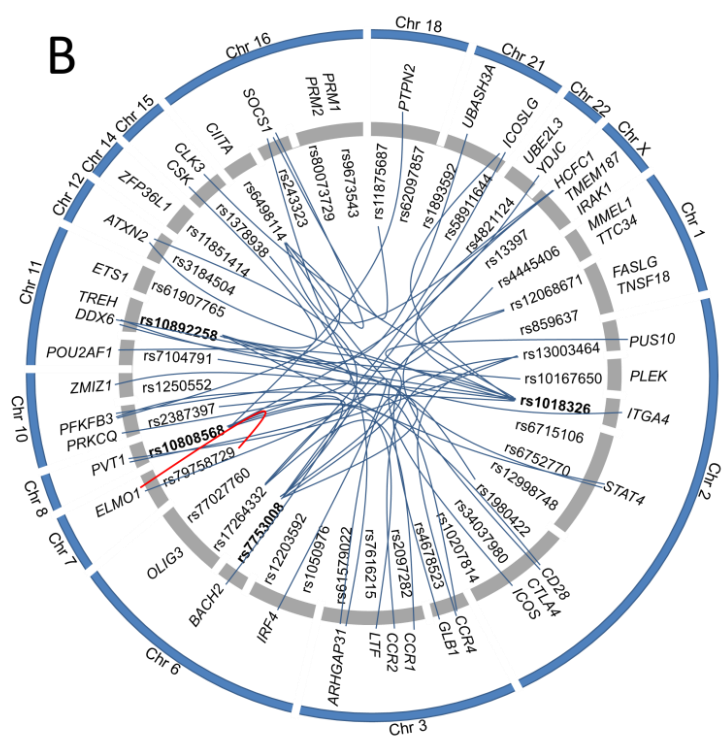
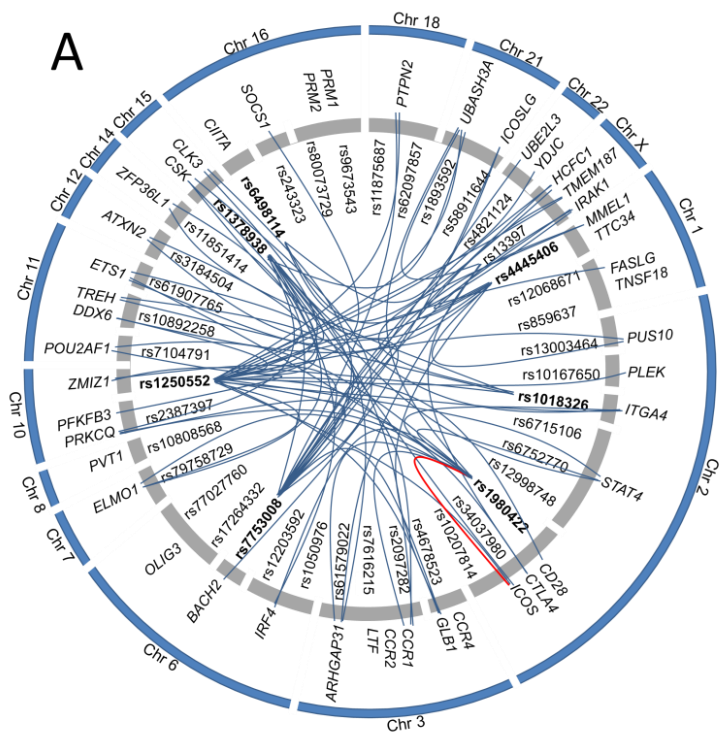
Gaixo aktibo eta tratatuen arteko konparaketak adierazpen aldatua duten 10 gene identifikatu zituen, 9 gene gaixotasun aktiboan gainadierazita (*CIITA*, *POU2AF1*, *IRF4*, *SOCS1*, *PVT1*, *ICOS*, *CTLA4*, *YDJC* eta *PLEK*) eta gene bakar bat azpiadierazita (*TREH*) (8-C irudia). Gaixo aktiboak eta kontrolak konparatzean lortutako emaitzekin bezala, emaitza hauetatik lortzen diren GO-termino guztiak immunitate zelulen aktibazioaren erregulazioarekin erlazionatuta daude *CTLA4*, *ICOS* eta *PLEK* geneen adierazpen aldatuaren ondorioz, gehi *IRF4* (interferon regulatory factor 4), birusen infekzioen aurrean interferonaren erregulazioan garrantzitsua den TLR seinalizazioaren erregulazio negatiboa duen transkripzio faktore linfuzito espezifikoa. Honetaz gain, kasu honetan gamma interferona erantzunarekin erlazionatuta dauden GO-terminoak agertzen dira ere, gaixotasun aktiboan ematen den prozesu inflamatorioaren ondorioz gainadierazita dauden 3 geneengatik, *CIITA* (class II MHC transactivator), *IRF4* eta *SOCS1* (suppressor of cytokine signaling 1).

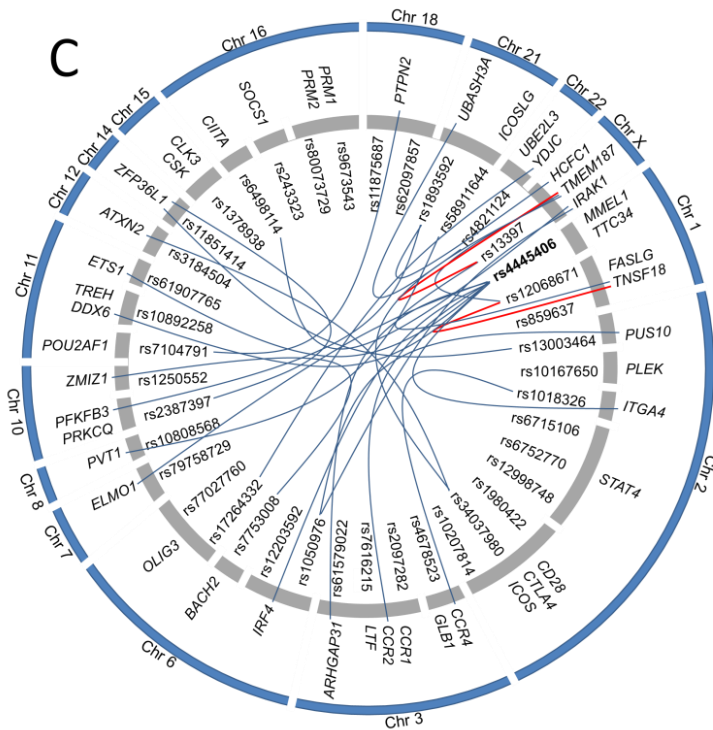
3.2. Genotipoaren efektua geneen adierazpenean

Nahiz eta ikerketa honetan erabilitako lagin biologikoen kopurua txikia izan, SNPen genotipoa eta geneen adierazpen mailaren arteko erlazioak aztertu dira. Bai genotipo zein adierazpen emaitzak eskuragarri zituzten 14 kontrol eta 9 gaixo bikote aztertu dira. Lagin kopuru txikiaren ondorioz kasu batzuetan ezinezkoa izan da talde guztietan 3 genotipoen presentzia lortzea; hau dela eta, potentzia estatistikoa emendatzeko helburuarekin lagin heterozigotoak eta alelo txikiarentzako homozigoto direnak elkartu dira.

Zenbait SNPen genotipoek hainbat generen adierazpenean eragina daukatela detektatu dugu, baina harrigarria bada ere, efektua estimuluaren menpekkoa dela dirudi, aztertutako taldeen arteko desberdintasunak ikusten baitira. Honetaz gain, eQTL gehienak *trans* egoeran ematen dira eta bakarrik 4 gene hautagai kokatzen dira haien gain eragina izan beharko lukeen balizko SNP erregulatuzailearen gailur berean; rs1980422-*ICOS* debutetan, rs79758729-*ELMO1* gaixo tratatuetan, rs12068671-*TNSF18* eta rs13397-*TMEM187* kontroletan (9. irudia). Emaitza hau azaltzen saiatzeko helburuarekin, asoziatu-eremu bakoitzaren eremu genomikoa sakonki aztertu genuen *trans* eQTLak azaldu ditzaketen elementu erregulatuzaileen bila. Horretarako, online dauden zenbait datu baseetan bilaketak burutu ziren, hala nola Haploreg (<http://www.broadinstitute.org/mammals/haploreg/haploreg.php>)¹⁰¹, Ensembl (<http://www.ensembl.org>)¹⁰² eta UCSC Genome browser (<http://genome.ucsc.edu>)¹⁰³. Espero genezakeen bezala, SNP erregulatuzaile potentzialen eraginpean egon daitezkeen elementuak kromatina irekiko eremuak, proteina-kodetzaile sekuentzia berriak, antisense

transkrito prozesatuak, pseudogeneak, microRNAk, lincRNA berriak eta eraldatutako proteina lotzeko itzak dira besteak beste. Aurkikuntza honek atea irekitzen die ikerkuntza berriei, sekuentzia hauek benetan geneen adierazpenean eta CDren garapenean eragin funtzionala duten ala ez frogatzeko.





9. irudia: SNPen genotipoaren eragina gene hautagaien adierazpen mailan, gaixotasunaren estatus desberdinetan. (a) Diagnostika, (b) >2 urte GFD eta (c) kontrolak, Anlisi estatistikoak Merlin 1.1.2 softwarearekin burutu dira; SNPen efektu esanguratsua adierazteko p balioa $< 0,005$ finkatu zen. Zenbait geneetan eragina duten SNPak beltzez adierazi dira. Lerro urdinak *trans*-eQTLak adierazten dituzte eta lerro gorriak *cis*-eQTLak.

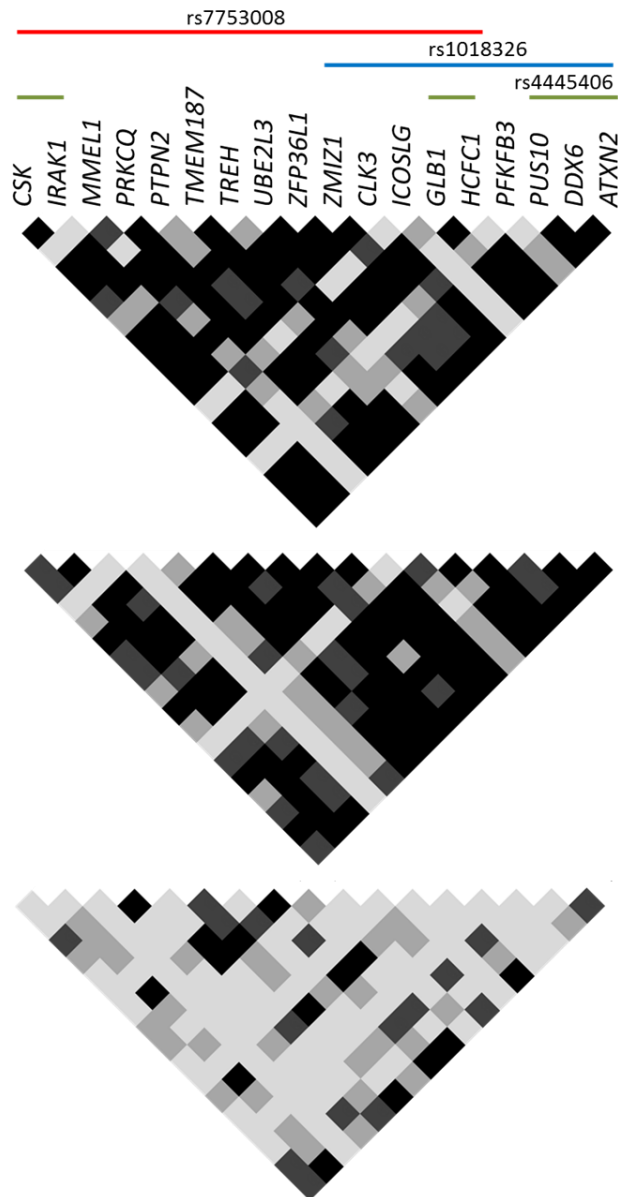
3.3. Koadierazpen gene-patroiak CDn

Gaixo zeliakoetan inflamazioaren ondorioz edota determinatzaile genetikoaren ondorioz aldatuta egon daitezkeen patroia erregulatuak orokorrak identifikatzeko helburuarekin, koadierazpen analisia burutu ziren. Koadierazpen patroia desberdinak ikusi ditugu aztertutako 3 taldeetan, koadierazpen maila handitik txikira glutena jaten duten gaixo zeliako, paziente tratatu eta kontrol ez zeliakoetan hurrenez hurren (10. irudia). Gaixoen bi taldeetan baina ez kontrolen koadierazten diren geneak bakarrik aztertuta, pazienteetan estuki koadierazten diren 18 gene osatutako azpimultzo bat identifikatu dugu. Honetaz gain, badirudi 18 gene hauek 3 SNPren kontrolpean egon daitezkeela (11. irudia)

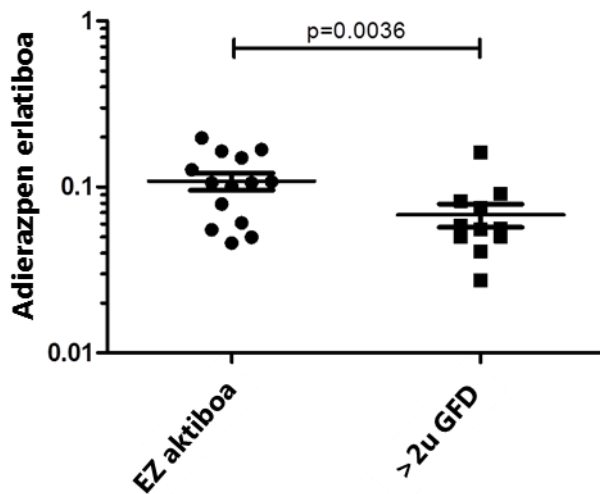


10. irudia: Gene pareen koadierazpen matrizeak gaixotasunaren estatus desberdinetan. Lauki txiki bakoitzak gene pare espezifiko baten adierazpen maila erlatiboaren arteko korrelazioaren p balorea adierazten du. Beltzak, gris ilunak, gris argiak eta zurian ondoko Pearsonen korrelazio p balioak adierazten dituzte, hurrenez hurren: $p < 0,0001$, $p < 0,001$, $p < 0,01$ eta $p > 0,01$.

Hauetariko SNP batek, rs1018326, *UBE2E3* eta *ITGA4* genean artean dagoen eremu intergenikoan dago kokatuta, bigarren kromosoman, eta momentuz funtzio ezezaguna daukan lincRNA (AC104820.2) baten gainean aurkitzen da. RNA gene honek 5 transkrito (*spliced variants*) ditu, 342-1771 base pare bitartekoak. AC104820.2 lincRNAren adierazpena modu esanguratsuan eraldatuta dago eritasunaren egoera desberdinean dauden biopsia bikoteak haien artean konparatzerakoan, biopsia aktiboetan gainadierazita dagoelarik (12. irudia). Parekatu gabeko biopsiak edo paziente aktibo eta tratatu independenteak konparatzean aldiz, ez da adierazpen diferentzia esangarririk ikusten. Hau horrela izanik, CD gaixoen arteko aldakortasun handia argi geratzen da berriro ere, eta konparaketa efektiboak lortzeko laginen arteko parekaketa zorrotzaren garrantzia ondorioztatzen da.



11. irudia: Gene pareen koadierazpen matrizeak gaixotasunaren estatus desberdinetan, pazienteetan baina ez kontrolatan korrelazioa aurkezten duten gene multzoan. (a) Diagnostika, (b) >2 urte GFD eta (c) kontrolak. Lauki txiki bakoitzak gene pare espezifiko baten adierazpen maila erlatiboaren arteko korrelazioaren p balorea adierazten du. Beltzak, gris ilunak, gris argiak eta zurian ondoko Pearsonen korrelazio p balioak adierazten dituzte, hurrenez hurren: $p < 0,0001$, $p < 0,001$, $p < 0,01$ eta $p > 0,01$.



12. irudia: AC104820.2 lincRNA-ren adierazpen erlatiboa 11 biopsia bikoteetan. Analisi estatistikoetan t-test parekatua aplikatu da.

4. Eztabaida

SNPen asoziazio azterketetan oinarrituta aukeratzen diren gene hautagaiak, ikertzen ari garen gaixotasunari buruzko alde aurretiko ezaguerengatik baldintzatuta daude gehienetan, eritasunean parte hartzen duten mekanismo patogeniko ezagunak kontutan hartzen direlarik. Azterketa funtzionalak aldiz, benetazko partaide funtzionalak identifikatzeko hurbilketa alboragabea eskeintzen dute. Orain arte, azterketa gutxi batzuk bakarrik burutu dute asoziatu-erregioen azterketa sakona suszeptibilitate gene hautagaiak proposatu baino lehen. *THEMIS* eta *PTPRK* geneen azterketa genetiko eta funtzionala egiterakoan, chr6: 127.99-128.38 Mb CD asoziazio eremuan aurkitzen diren gene hautagai hauen adierazpenak korrelazioa esanguratsua aurkeztu zuen CD gaixoetan baina ez kontrol taldean⁸⁴. Aurkikuntza honek, bi gene hauek CDrekin zerikusia izan dezaketela adierazteaz gain, eremu intergeniko ez kodetzailean egon daitekeen erregulazio sistema amankomunaren existentzia frogatuko luke. Hurbilketa desberdina erabiliz, aldaera funtzional posibleak identifikatzeko burututako *LPP* locusaren mapeketa sakonari esker erregulazio eremuetan aurkitzen diren sei SNP identifikatu ziren, rs4686484 izan zelarik CD gaixoetan *LPP* genearen adierazpenean eragin handiena izan dezakeena¹⁰⁴. Azkenik, Östensson M. *et al*¹⁰⁵ orain dela gutxi burutu izan dituzte bidezidor-analisiak eta bi-locus arteko elkarrekintzak ikertzeko azterketak, asoziazio seinaleak sakonki aztertzeko helburuarekin. CD pazienteen heste-mukosan adierazpen aldakorra aurkeztzen duten zenbait gene aurkitzeaz gain, asoziazio eremuetan kokatzen diren suszeptibilitate geneak kategoria desberdinetan elkartu daitezkeela ikusi zuten. Emaidza hauekin, geneak eta bidezidorrak batera aztertuta CD eta beste gaixotasun inflamatorio kronikoren sorreran eragina izan dezaketen mekanismo biologiko potentzial berriak aurkitu daitezkeela proposatzen dute.

Asoziatu-eremuko SNPek geneen adierazpenean izan dezaketen eragina aurretik argitaratutako lan askotan aztertu izan da, zenbait *cis*- eta *trans*-eQTL aurkitu direlarik, baina azterketa hauetan adierazpena beti neurtu izan da odol laginetan¹⁰⁵. Orain dela gutxi argitaratutako lan batean, monozitoen aktibazioak eragindako aldaera erregulatzaileen efektua aztertu da, eta ikusi dute aldaeren proportzio esanguratsu batek bakarrik aurkeztzen duela aktibitatea baldintza zehatz batzuetan. Hau dela eta, aldaera genetiko funtzionalak hobeto ulertu eta aztertzeko, egoera genetiko, zelularra eta ingurumena aztertu behar direla proposatzen dute¹⁰⁶. Hau izan da oraingo azterketa honetan aurkitu dugun egoera, non gliadinaren luzerako insultoak eta CD paziente aktiboetan ematen den inflamazioak eragindako eQTL desberdinak bereizteko gai izan garen azterketa talde bakoitzean. Honetaz gain, asoziatu-eremuko aldaera genetikoaren itzulpen funtzionalean parte

hartzen duten mekanismo konplexuetan murgilduta, badirudi rs1018326 SNPak lincRNA (AC104820.2) baten bidez burutzen duela bere efektua. LincRNA honen adierazpena aztertuz ikusi izan dugu, gaixotasunarekin erlazionaturiko adierazpen aldaketak bakarrik nabaritzen direla paziente berdinaren bi biopsiak konparatzen ditugunean, gaixotasunaren bi estadioak pertsona berean konparatzean hain zuzen ere.

Identifikaturiko koadierazpen ereduak dagokienez, gaixotetan ikusitako korrelazio altuenak aldaketa edo alterazio koordinatu baten existentzia adierazi dezake, berriro ere mekanismo erregulatu konplexuen ondorio izan daitekeena. Gaixotasun aktiboaren kasuan, kontutan hartuta proposaturiko gene hautagai guztiak erantzun immunitarioarekin erlazionatuta daudela, gliadinaren ingestioak eragindako ingurune inflamatoriak azal zezakeen korrelazio altuaren agerpena. Hala ere, trataturiko pazienteetan mantendu egiten da deskribatutako korrelazioa, nahiz eta gliadina gabeko dieta egin izan 2 urte baino denbora luzeagoan. Hau azaltzen saiatzeko 2 aukera ditugu: indibiduo zeliakoetan emte den eta gliadinaren ingestioarekiko independentea den alterazio intrinseko eta konstitutibo baten aurrean egotea, edo, gliadinak sortutako erantzun patroia izatea zeinetan 2 urteko GFD nahikoa ez den gene hauen adierazpen maila basalak lortzeko.

Kontrako koadierazpen ereduak ikusi izan da gure taldeko beste ikerketa lerro batean, non NFkB bidezorra aztertu izan den CDn¹⁰⁷. Kasu horretan korrelazio maila altuena kontrol taldean aurkitu izan da, honen arrazoia izan zitekeen heste osasuntsuan bidezidor horren kontrol erregulatuak estua ematen dela, gaixoen kasuan berriz aldatuta agertzen delarik. Bi azterketa hauetan lortutako kontrako emaitzek sententzia eduki dezakete hurrengo kontutan hartuz gero, NFkB bidezidorrean aurkitzen diren geneak mekanismo erregulatu berdinaren menpe egotea espero genezake, bidezidor beraren parte izanagatik. Gaixotasunarekin erlazionatutako locietan aldiz, erantzun immunitarioarekin zerikusia duten geneetan aberastuta egon arren, ezin dugu espero gene hauek modu koordinatuan erantzutea inguruneko estimulu baten aurrean aldaera erregulatu berarekin erlazionatuta ez badaude. Ikerketa honetan adierazpen mailan aztertu ditugun gene hautagaiek amankomunean dute haien partehartzea erantzun immunitarioan, CDren ondorioz aldatuta dagoena, eta arrazoi honengatik aztertutako geneen alterazio koordinatua ulergarria izan daiteke.

Ikerketa honetan proposatzen den ideia frogatu ahal izateko, berrespen experimental sendoak behar ditugu, eta eritasun genetiko arrunten suszeptibilitatearen puzzlean pieza asko enkatatu behar dira oraindik. Hala ere, argi dago asoziatutako aldaeren efektua,

ondoko lokusen kontrol transkripzionalaren ideia sinple baino baino sinpleagotik urrun dagoela. Ingurunekeo estimulu baten aurrean ematen den erantzun koordinatu eta osasuntsua ahalmentzen duten elkarrekintza konplexuak gure genoman idatzita daude, eta kontrol finaren mekanismoen haustura izango da amaieran gaixotasuna erango duten gertaera segidaren lehenengo pausua.

3. kapituluua:

Immunoehipeko gene hautagaien azterketa gaixo zeliakoen heste
epitelioko zelula populazioetan

1. Aurrekariak

Eritasun zeliakoa (CD) suszeptibilitate genetikoa duten pertsonetan elikaduraren glutenarekiko intolerantziak eragindako ohiko gaixotasun immunitarioa da. Heste meharreko mikrobilosken atrofia, kripten hiperplasia eta infiltrazio linfositikoa dira ezaugarri histologiko tipikoak, eta tratamendu eraginkor bakarria bizitza osorako glutenik gabeko dieta (GFD) egitea da ⁶². CD garatzeko arrisku handiena ematen duen lokusak MHC eremuan dago, 6p21 kromosoman, eta %40ko ekarpen genetikoa daukala kalkulatu da. Hala ere, HLAren arrisku aldaerak (HLA-DQ2/-DQ8) CD garatzeko beharrezkoak badira ere, populazio orokorrean alelo hauek ohikoak direnez, CD garatzeko arriskua ematen duten bestelako lokus ezezagunak egon behar dira inplikaturak.

Genoma osoko asoziazio azterketak (GWAS) izan dira CD pairatzeko arriskuaren atzean dauden eremu genomikoak aurkitzeko erabili den estrategia nagusia. GWAS emaitzak eta orain dela gutxi argitaratu den Immunochip proiektuak HLAtik kanpo dauden 39 eremu genomiko berri asoziatu dituzte CDrekin ^{71,79,86}.

Populazioen arteko desberdintasunak agerian geratzen dira asoziazio ikerketak populazio independenteetan erreplikatzeko bakoitzean ^{80-83,99}. Arazo honen atzean dagoen arrazoia, asoziazio ikerketa handietan erabilitako laginen jatorri genetikoa izan daiteke. Immunochip proiektuan adibidez, CD aztertzeko laginen erdia Erresuma Batuko bildumatik eratorriak ziren (7.728/12.041 kasu eta 8.274/12.228 kontrol) ⁷⁹. Hau dela eta, eskuragarri dauden genotipaketa sakoneko emaitzak beste modu batean aztertuta, zeinetan populazio-espezifikokoak diren diferentziak zuzentzen diren, asoziazio emaitza berriak sortuko lituzke, asoziazio seinale eta gene hautagai edota elementu erregulatzaile berriekin.

Ikerketa hauetatik abiatuta 66 gene hautagai proposatu dira asoziazio eremuetan kokatuta egoteagatik. Gene hauen inguruan zenbait ikerketa funtzional burutu dira, batez ere odol periferikoko zelula mononuklearretan (PBMC) eta gutxi batzuk gaixotasunaren itu ehunean, hesteetako epitelioa hain zuzen ere ^{84,99,108}.

Hesteetako biopsia laginak zelulen nahastura heterogeneoa dira, eta honek emaitza negatibo zein positibo faltsuak ekar ditzake zelula populazio desberdinetan gene hautagaien adierazpen desberdina egonez gero. Hesteetako biopsietan agertzen diren zelula populazio nagusiak enterozitoak eta sistema immunitarioko zelulak dira, batez ere linfzito intraepitelialak (IELs) ¹⁰⁹. Enterozitoen ezaugarri bereizgarria, gainazalean adierazten duten zelula epitelialen adhesio molekula (EpCAM) da, molekula hau zelula seinalizazioan, migrazioan, proliferazioan eta desberdintzapenean parte hartzen duelarik. Molekula hau edukitzeagatik, enterozitoak zelula CD326 porritiboak direla esaten da. Bestalde, immunitate zelulek CD45 antigenoa aurkezten dute haien gainazalean, *PTPRC* geneak kodetzen duen proteina tirosina fosfatasa. Ezaugarri hauei esker, bi zelula populazioak banandu daitezke hurrengo analisisietan modu independenteen azterzea ahalbidetuz, eta lortutako emaitzak biopsia osoetan egindako azterketa funtzionalekin konparatu daitezke ¹⁰⁸.

Ikerketa honetan, populazioen artean asoziazio emaitzak homeogeneoak ez direla kontutan edukita, lehenik eta behin immunohiperproiektutik eskuragarri dauden emaitza genotipikoak aztertu ditugu jatorri espainiarra duten CEGECeko laginetan, lagin osoan asoziaziorik aurkeztu ez zuten eremu genomiko berriak identifikatzeko helburuarekin. Gainera, proposaturiko gene hautagaiek eritasunaren garapenean izan dezaketen ardura aztertu dugu heste epitelioko zelula populazioak bananduta aztertuz, orain arte erabili ez den planteamendu metodologikoa erabiliz. Modu honetan, bi zelula populazioetan adierazpen maila diskordantea identifikatzeko gai izan gara.

2. Materiala eta metodoak

2.1. Pazienteak eta laginak

Eritasun zeliakoa erreklutamenduaren garaian indarrean zegoen ESPGHAN (European Society for Pediatric Gastroenterology) irizpideen arabera diagnostikatu zen, horretarako anti-gliadina eta andi-edomisio (EMA) edo anti-transglutaminasa (TGA) antigorputzen determinazioa eta heste meharreko biopsiaz baliatuz. Ikerketa guztiak etika batzordeen onespina lortu zuten (Gurutzetako unibertsitate ospitalea kodea CEIC-E09/10) eta analisi guztiak emailen edota beren gurasoen baimen informatua lortu ostean burutu dira. Biopsia laginak errutinazko diagnosi endoskopien bidez lortu dira, lagin bat diagnostikorako eta bestea ikerketa honetan erabili direlarik. Diagnostiko garaian zeuden 8 gaixoen heste biopsiak, zeliakoak ez diren eta endoskopia garaian inflamaziorik aurkezten ez zuten kontrolen biopsiekin konparatu dira. Honetaz gain, 9 gaixo zeliakoen odola jaso zen PBMC-ak isolatu eta azterketa funtzionalak burutzeko.

2.2. Zelula populazioen bereizketa hesteetako biopsietatik

Biopsietako epitelio zelulak eta immunitate zelulak bereizteko, MACS zelula banaketa magnetikoa erabili izan da (Miltenyi Biotec), CD45 MicroBeads eta MS banaketa zutabeentzako fabrikatzailearen protokoloa jarraituz. Laburki, biopsiak jaso eta jarraian prozesatu ziren. Zelulak modu mekanikoan banatu ziren medio osoan (RPMI + antibiotikoak + FBS + DTT + EDTA) agitazio bidez ordu baten. Mediora askatutako zelulak zentrifugazio bidez jaso ziren medioa aurre-banaketa filtroetatik (20 µm) pasatu eta gero. Desegin gabe geratzen zen lamina propria RLT bufferrean gorde zen -80 °C-tan.

Zelula bizien suspentzio homogenea lortzeko, zelula hilak kentzeko kita (Dead cell removal kit, Miltenyi Biotec) erabili zen lehenego purifikazio pausu batean. Bigarren purifikazio pausuan, zelulak CD45 microbead magnetikoekin markatu ziren zelula nahastea bi frakzioetan banatu ahal izateko, CD45 positiboak diren immunitate zelulak (zutabeen atxikituta geratuko direnak) eta zelula CD45 negatiboak edo enterozitoak (zutabea zeharkatuko dutenak). Lortutako bi frakzioak RLT buferrean gorde ziren -80 °C-tan azido nukleikoak erauzi ziren arte.

Laginaren portzio txiki bat zitometria bidez aztertu zen zelula populazioen banaketa modu egokian egin zela ziurtatzeko.

2.3. **Odol-periferikoko zelula mononuklearren isolamendua**

Gaixo zeliakoen odol laginak bildu ziren eta momentuan PBMCak banandu ziren Ficolla erabiliz, dentsitate gradiente zentrifugazioa burutuz. Zentrifugazioaren ostean, Ficollak odola geruzatan banatzen du, linfuzitoak eta monozitoak plasma geruzaren azpian isolatuz. PBMC geruza pasteur pipeta batekin jaso zen eta RPMI medioan 2 garbiketa burutu ostean, PBMC-ak -80 °C-tan kriopreserbatu ziren izozteko medio berezian, azido nukleikoak erauzi arte.

2.4. **RNA laginak eta gene adierazpena**

RNA NucleoSpin microRNA kit-arekin (Macherey-Nagel) purifikatu zen fabrikatzailearen instrukzioak jarraituz. RNA totala 8 ng/μl-tan normalizatu eta cDNA bihurtu zen AffinityScript cDNA Synthesis kit-a (Agilent Technologies, Santa Clara, CA, USA) erabiliz, fabrikatzailearen protokoloaren arabera. Geneen adierazpen analisia Fluidigm Biomark 48.48 dynamic array-tan (Fluidigm Corp., South San Francisco, CA, USA) eta TaqMan Gene Expression entsegu komertzialekin burutu zen. *RPLPO* gene endogenoaren adierazpena esperimendu berean kuantifikatu zen eta erabilitako RNA kantitatearen kontrol modura erabili zen. Gene bakoitzaren adierazpen erlatiboa, Ct metodo zehatza erabilita kalkulatu zen³⁴ eta adierazpen emaitzak lagin guztien mediarekiko normalizatu ziren.

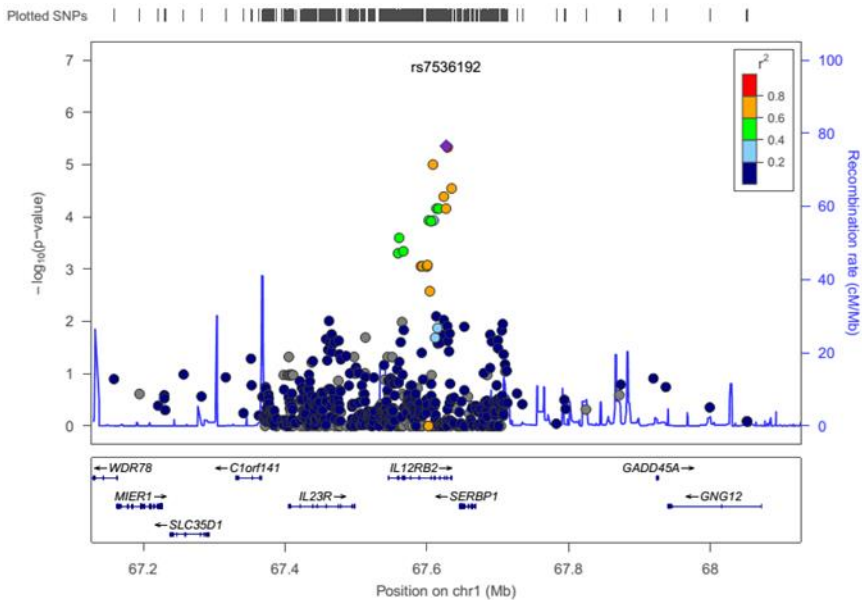
2.5. **Asoziazio azterketa**

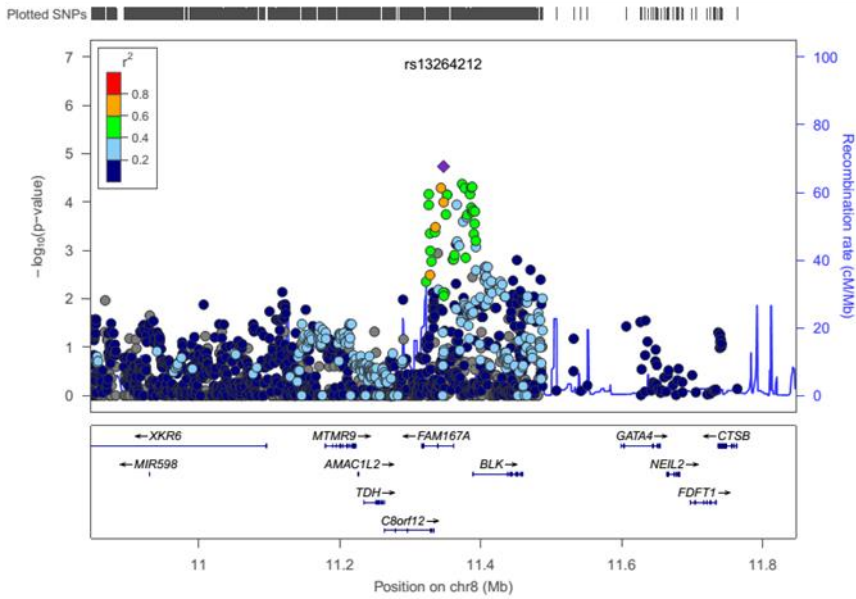
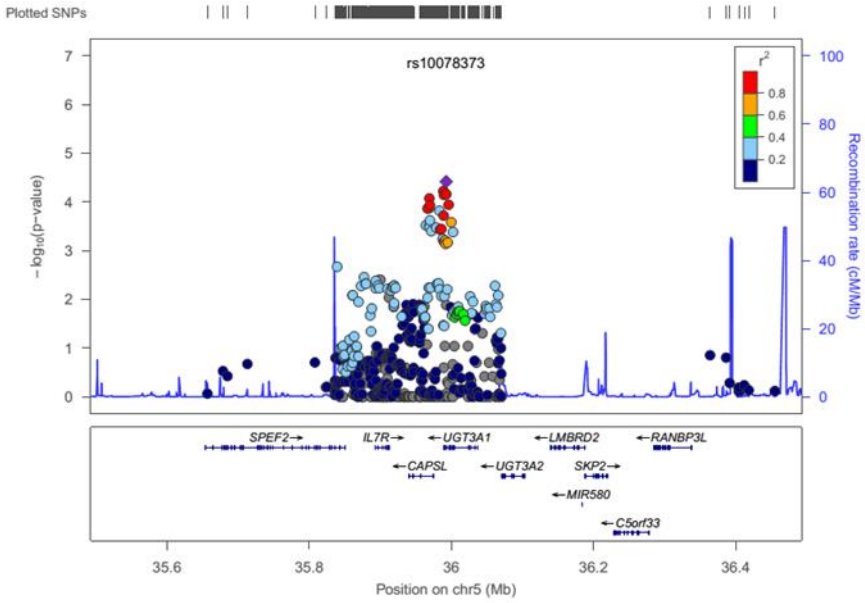
ImmunoChip proiektuan genotipatu zen CEGEC lagin multzoan, 545 gaixo zeliako eta 308 kontrol ez-zeliako zeuden⁷⁹. Lagin hauen asoziazio analisia PLINK v1.07¹¹⁰ programarekin egin zen, asoziaziatutako 50 SNP gorenak aukeratu ziren. Asoziatu 4 SNP edo gehiago zituzten LD-blokeak LocusZoom-arekin⁸⁷ kokatu ziren gene hautagai berriak aukeratzeko.

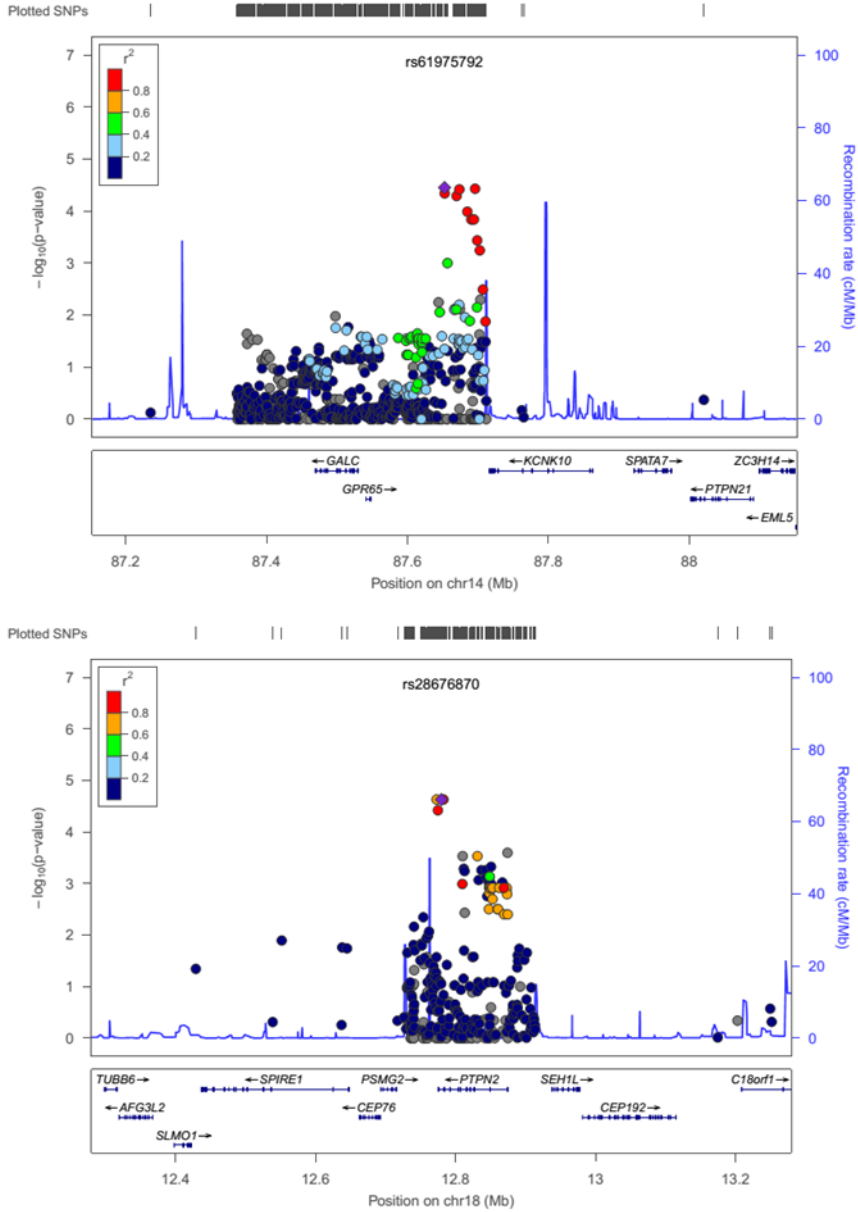
3. Emaitzak

ImmunoChip proiektuko CEGECeko genotipoak bakarrik aztertu ziren, populazioen arteko desberdintasunek izkutuan utzitako gene hautagaiak identifikatzeko helburuarekin. CEGECeko espainiar populazioko 50 SNP esanguratsuenak hartu ziren, p balioak 7.90×10^{-5} eta 1.53×10^{-12} artean zeudelarik (2. taula gehigarria). Seigarren kromosoma azterketa honetatik kanpo utzi zen.

Asoziatutako SNP bat baino gehiago duten 5 eremu genomiko berri identifikatu dira. SNP bat baino gehiago asoziatuta edukitzeak, seinalea benetakoa izateari indarra ematen dio. Sakon genotipaturiko eremuetan dauden seinaleak 13. irudian aurkezten dira





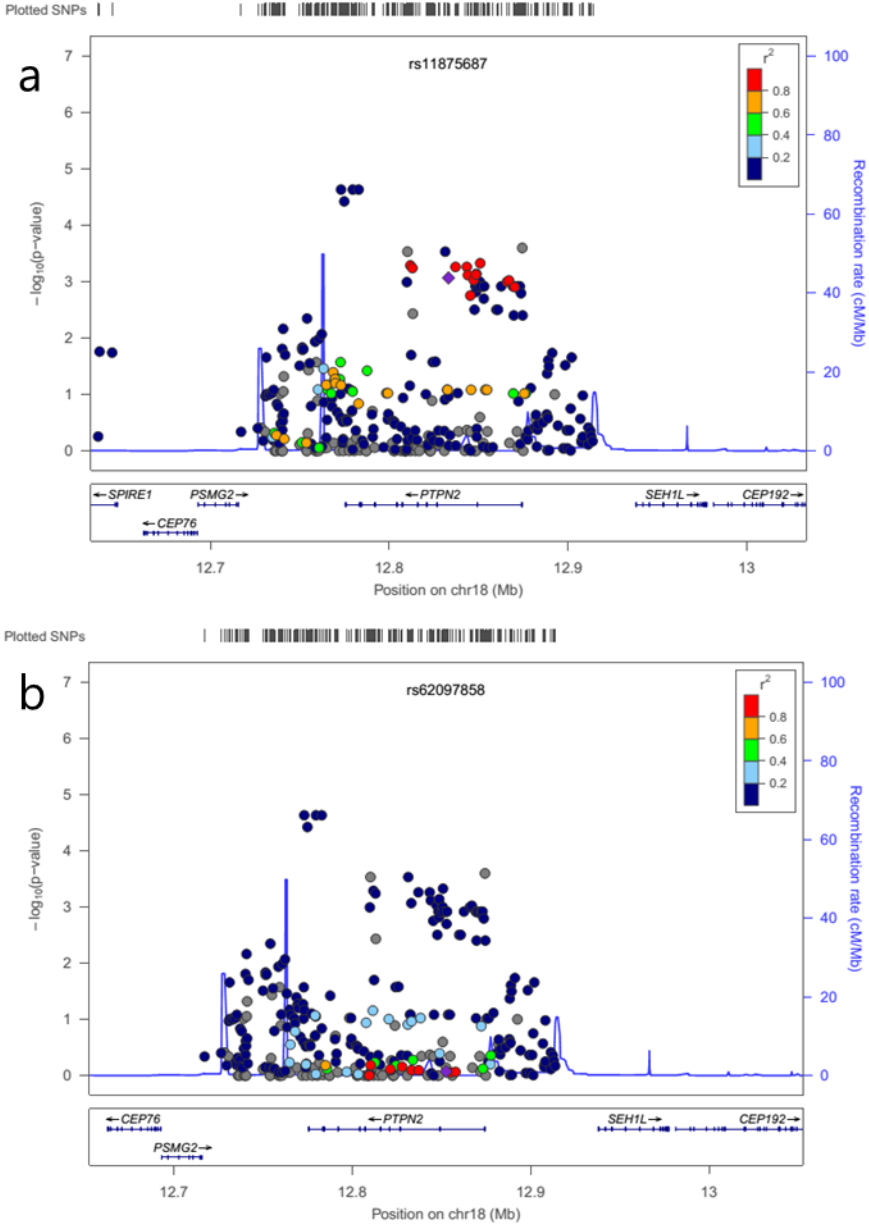


13. irudia: CEGEC laginean asoziatuiko 5 eremuak Locus Zoom bidez irudikatuta. Lokus bakoitzean, lehenengo seinalea diamante more bidez adierazten da eta r^2 balioaren arabera LD-n dauden SNP-ak.

Asoziazio seinaleek eremu horietan dauden gene hautagai berriak proposatzen dituzte. *IL23R*, *IL12RB2* eta *SERBP1* 1. kromosoman (67353404-67390948, rs7536192 $p_{\min} = 4.38 \times 10^{-06}$); *IL7R*, *CAPSL*, *UGT3A1* eta *UGT3A2* 5. kromosoman (35930601-35962030, rs10078373 $p_{\min} = 3.78 \times 10^{-05}$); *FAM167A*, *C8orf12* eta *BLK* 8. kromosoman (11433457-11457539, rs13264212 $p_{\min} = 1.82 \times 10^{-05}$); *GALC* eta *GPR65* 14. kromosoman (88099567-88154762, rs61975792 $p_{\min} = 3.53 \times 10^{-05}$), eta *PTPN2* eta *PSMG2* 18. kromosoman (12777265-12804934, rs28676870 $p_{\min} = 2.33 \times 10^{-05}$).

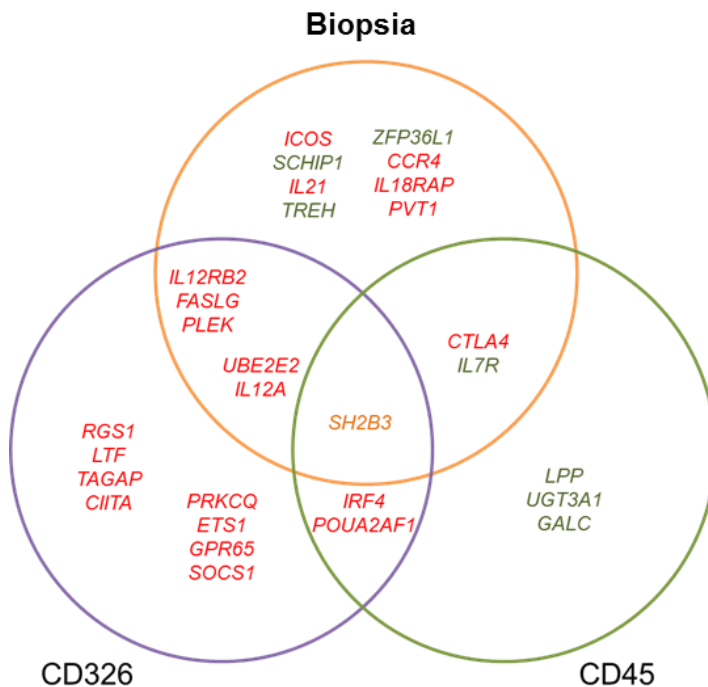
Gene hauetatik, *PTPN2* da aurretik immunoehipeko emaitzetan oinarrituta gene hautagai moduan proposatu den bakarra, 2 SNP desberdinek seinale independentea aurkezten zutelako, rs11875687 eta rs62097858. Harrigarria izan daiteke 2 SNP hauen p balioak CEGECeko laginean 0,000856 eta 0,999 izatea hurrenez hurren, gainera ez bata ez bestea ez daude LD-n gure populazioko SNP asoziazioatuenarekin (rs28676870) (14. irudia), populazio-espezifikokoak izan daitezkeen haplotipo desberdinak iradokiz. *IL23R* genearen CDrekiko asoziazioa aurkitu dute aurreko zenbait ikerketek, baina populazioen arteko desberdintasunak ikusi dira¹¹¹⁻¹¹³ eta immunoehipean ez da asoziaziorik ikusi.

Lan honetan buruturiko gainontzeko azterketa funtzionaletan, CEGEC laginean identifikatu diren gene hautagai berriak gehitu dira aurretik immunoehipean identifikaturiko gene zerrendara. Immunoehipeko emaitzetan oinarrituta proposaturiko gene hautagaiak gure taldearen publikazio batean aztertu ziren¹⁰⁸ eta gure azterketa funtzionalak gaixo zeliakoen heste biopsian (gaixotasun aktiboa eta tratatua) eta kontrolatan desberdin adierazten diren zenbait gene identifikatu zituen.



14. irudia: *PTPN2* eremua a) Immunochipean gehien asoziatuiko SNP-ak b) CEGECeko espainiar populazioan.

Zelula epitelialak aztertzerakoan (CD326+) 16 genetan (*IL12RB2*, *FASLG*, *RGS1*, *PLEK*, *UBE2E2*, *LTF*, *IL12A*, *IRF4*, *TAGAP*, *PRKCQ*, *POU2AF1*, *ETS1*, *SH2B3*, *GPR65*, *CIITA* eta *SOCS1*) ikusi ziren adierazpen desberdintasunak ($p < 0,01$) gaixo zeliako aktibo eta kontrolen artean. Gene guzti horiek gainadierazita daude gaixoen laginetan (15. irudia). Aldatuta dauden geneen GO-terminoen analisiak alfa-beta T zelulen aktibazioan parte hartzen duten geneen aberastea adierazi zuen, prozesu honetan garrantzi gehien duten geneak *PRKCQ*, *IL12A* eta *IRF4* izanik. Zazpi genez osatutako azpitalde batek (*IL12RB2*, *FASLG*, *LTF*, *IL12A*, *PRKCQ*, *ETS1* eta *SOCS1*) bidezidor elkarrekintza aurkeztu zuen GeneMANIA predikzio zerbitzariaren arabera ¹¹⁴. KEGG bidezidor azterketak, Jak-STAT seinalizazio bidezidorrean, zitokina-zitokinen hartzaille elkarrekintzan eta hesteetako eritasun inflamatorioan (IBD) kokatu zituen gene hauek. Azken hau, IBD, CDrekin asoziaturiko lokus asko partekatzen dituen gaixotasun immunologikoa da ^{115,116}.



15. irudia: Ikertutako taldeetan adierazpen desberdintasunak dituzten geneak. Gaixo zeliakoetan gainadierazita dauden geneak gorritz, azpiadierazitakoak berdez, eta lagin motaren arabera joera desberdina duten geneak laranja agertzen dira.

Beste aldetik, zelula immunitarioek (CD45+) adierazpen aldatuta dute aztertutako 8 genetan. Hauetariko 5 gene (*LPP*, *IL7R*, *UGT3A1*, *SH2B3* eta *GALC*) azpiadierazita daude gaioxtasun aktiboan kontrolerikiko konparatzen ditugunean, eta gainontzeko 3 geneak (*CTLA4*, *IRF4* eta *POU2AF1*) gainadierazita. GO terminoen analisiak ez zuen prozesu esanguratsurik aurkitu.

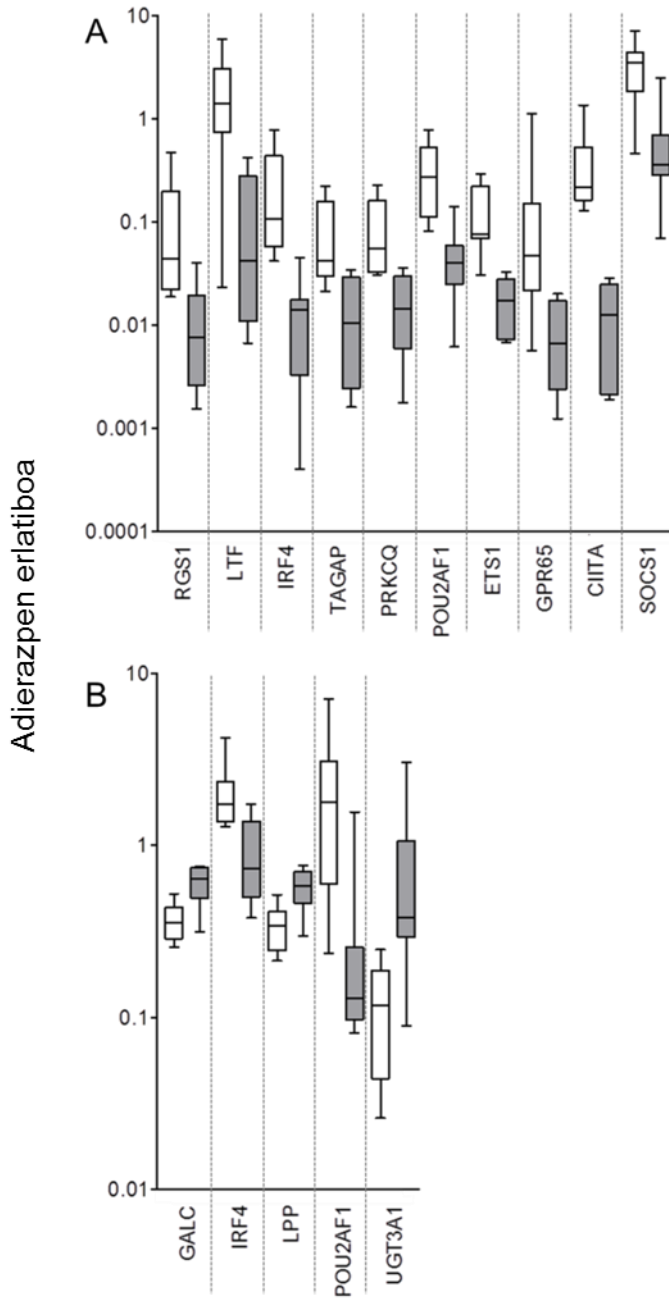
Interesgarria da ikustea nola 3 gene (*IRF4*, *POU2AF1* eta *SH2B3*) alterazio berdina aurkezten duten bi zelula populazioetan, baina *SH2B3* geneak kontrako efektua dauka CD326+ eta CD45+ zeluletan, zelula epitelialetan gainadierazita eta immunitate zeluletan azpiadierazita agertzen delarik gaixo zeliakoan laginetan.

Ikerketa honetan lortutako emaitza berriak aurretik biopsia osoetan lortutako emaitzekin^{99,108} konparatzean, 3 lagin moten arteko desberdintasunak agerian geratzen dira. *SH2B3* da 3 laginetan adierazpen desberdina aurkezten duen gene bakarra, 5 gene desberdintasunak aurkezten dituzte zelula epitelial zein biopsietan (*IL12RB2*, *FASLG*, *PLEK*, *UBE2E2* eta *IL12A*), eta bi (*CTLA4* eta *IL7R*) linfozito intraepitelial eta biopsietan. Beste 8 gene (*IL18RAP*, *ICOS*, *CCR4*, *SCHIP1*, *IL21*, *PVT1*, *TREH* eta *ZFP36L1*), soilik biopsia osoak aztertzen direnean aurkezten dituzte desberdintasunak adierazpen mailan.

Lagin mota bakoitzean lortutako emaitzak konparatuz, zelula motaren espezifikoak diren gene aldatuak identifikatzeko gai izan gara. Gene horietan ez da inolako adierazpen desberdintasunik antzematen biopsia osoetan aztertzerakoan. Zortzi generen kasuan (*RGS1*, *LTF*, *IRF4*, *TAGAP*, *PRKCQ*, *POU2AF1*, *ETS1* eta *GPR65*), adierazpen aldakortasun esanguratsua ($p < 0,01$) CD326 zelula populazioan baino ez da ikusten. Gene hauetatik, *GPR65* CEGECeko laginetan identifikaturiko gene hautagaia da. Modu esklusiboan CD326 zeluletan aldatutako gene guztiak gaioxtasun aktiboan gainadierazita daude kontrolerikiko konparatuta (16. irudia).

Beste aldetik, 5 gene CD45 zelula populazioaren esklusiboa den adierazpen desberdina ($p < 0,01$) aurkeztu zuten (*GALC*, *IRF4*, *LPP*, *POU2AF1* eta *UGT3A1*). Gene hauetatik, *GALC* eta *UGT3A1* CEGEC lagin multzoan identifikatu ziren. *GALC*, *LPP* eta *UGT3A1* gaioxtasuna aktiboan azpieadierazita daude kontrolerikiko; aldiz, *IRF4* eta *POU2AF1* gaixo zeliako aktiboetan gainadierazita daude (16. irudia).

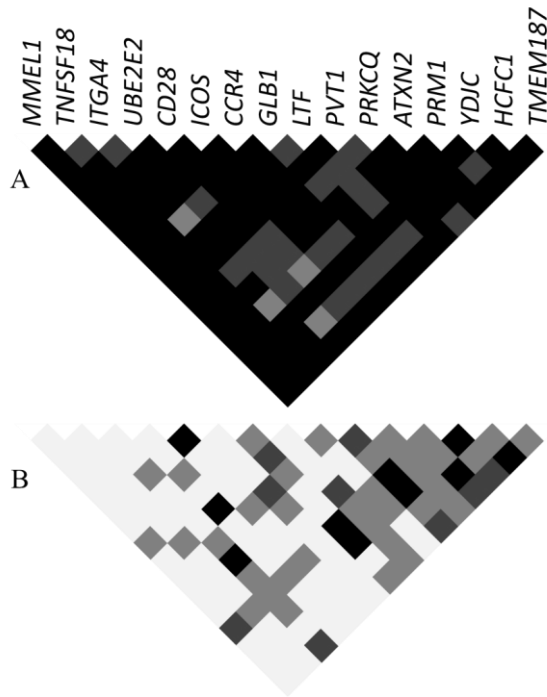
Interesgarria da nola adierazpen desberdina aurkezten duten bi gene zelula populazioak bereizteari esker identifikatu diren, bi zelula populazioetan baina ez biopsia osoetan baitute adierazpena aldatua. Bi geneek joera berdina aurkezten dute bi zelula populazioetan, CD326 zein CD45ean, gaioxtasun aktiboan gainadierazita daudelarik.



16. irudia: Soilik zelula populazioetan adierazpen aldatua duten geneen adierazpen erlatiboa. A: CD326+ zelula populazioa. B: CD45+ zelula populazioa. Grafikoak mediana eta 10-90 pertzentilak aurkezten dira.

Honetaz gain, geneen adierazpena aztertu dugu 9 gaixo zeliako aktiboen PBMCetan eta lortutako emaitzak debuten biopsietan lortutako gene adierazpen mailarekin konparatu ditugu. Aztertutako immunochipeko 52 gene hautagaietatik, 18k adierazpen desberdintasunak aurkeztu dituzte PBMC eta biopsia osoen artean. Hauetariko 8 genek (*FASLG*, *IL18RAP*, *CTLA4*, *CCR4*, *IL12A*, *SH2B3*, *SOCS1*, *ICOSLG*) aurretik identifikatu dira gure taldean gaixotasunak eraldatutako gene modura^{99,108}.

Koadierazpen analisiak burutu dira ere, gaixotasunaren egoera edota determinatzaile genetikoaren ondorioz eman daitezkeen zelula-espezifikoak diren erregulazio sinadura edo patrioiak identifikatzeko asmoarekin. Badirudi CD45+ zelula populazioak aztertutako gene batzuen koadierazpen maila altuagoa daukala gaixo zeliakoetan kontroletan baino. Gaixoen CD45+ zeluletan eta ez kontroletan oso modu estuan koadierazita dauden 16 genek osaturiko azpitalde bat identifikatu dugu (17. irudia).



17. irudia: Gaixo zeliakoen baina ez kontrolen CD45+ zelula populazioan koadierazpen estua aurkezten duten gene bikoteen matrizea. A: CD45+ zelulak paziente zeliakoetan B: CD45+ zelulak kontroletan. Lauki txiki bakoitzak gene pare espezifiko baten adierazpen maila erlatiboaren arteko korrelazioaren p balorea adierazten du. Beltzak, gris ilunak, gris argiak eta zurian ondoko Pearsonen korrelazio p balioak adierazten dituzte, hurrenez hurren: $p < 0,0001$, $p < 0,001$, $p < 0,01$ eta $p > 0,01$.

CD aktiboan, zeinetan inflamazio prozesu autoimmunitarioa martxan dagoen, CD45+ zeluletan koadierazpena aurkezten duten geneen GO terminoen analisiak zelula immunitario efektoreen erregulazio eta aktibazioarekin erlazonaturiko terminoak uzten ditu agerian. Eraitza honetan pisu gehien duten geneak *PRKCQ*, *CD28*, *ICOS* eta *TNFSF18* dira.

CD326+ zelula epitelialetan ez dugu gaixo zeliako eta kontrolen arteko desberdintasun esanguratsu aurkitu geneen koadierazpenari dagokionez.

4. Eztabaida

Dagoeneko ezaguna da GWAS azterketak egiterako orduan oso garrantzitsua dela laginaren tamaina, eta konklusioa argia dirudi: Zenbat eta lagin gehiago hobeto ⁷⁴. Hori da Immunochip proiektuan herrialde desberdinetako laginak erabiltzearen arrazoi nagusia, lagin tamaina ahalik eta handiena lortzea. Baina giza-populazioen arteko desberdintasunak behin eta berriro geratu dira agerian CDn egindako GWASetako emaitzak erreplikatzan saiatu diren bakoitzean ^{80-83,99}.

Oraingo ikerketa honetan ere, desberdintasun hauek agerian geratu dira eta asoziazio eremu berriak identifikatzeko gai izan gara CEGEC lagin multzoa bere aldetik aztertuta. Honek adierazten digu garrantzitsua dela laginen hondo-jatorri genetikoaren kontutan edukitzea asoziazio emaitzak aztertzen direnean. Orain dela gutxi, gure taldean indibiduoak beraien jatorri immunogenetikoaren arabera estratifikatzen dituen hurbilketa metodologiko berria erabili dugu immunochipeko emaitzak aztertzeko, CDrekin erlazionatuta dauden elementu genetiko berriak identifikatzeko helburuarekin (artikuluaren errebisioan).

Proposatu diren gene hautagaien inplikazio funtzionala frogatzeko nahitaezkoa da gene horien adierazpen analisiak gaixotasunaren iturriaren burutzea, batez ere aztertzen dugun eritasuna ehun-espezifikoa denean, CD kasu. Orain arte, CD aztertzeko burutu diren ikerketa funtzional gehienetan PBMCak aztertu dira, eta gutxi batzutan gaixo zeliako zein kontrolen heste biopsiak. Baina hurbilketa hauek ez dira nahiko izango kontuan edukita heste epiteliaren osaketa, non zelula populazio ugariena zelula epiteliaren zelula den, zelula immunitario infiltratuen aurretik, azken kasuan batez ere linfozitoak ¹⁰⁹.

Hau dela eta, gure taldeak azterketa funtzionalak egiteko hurbilketa metodologiko berri bat proposatu eta martxan jarri du, non heste biopsietako zelula epiteliak eta immunitate zelulak bereizten diren, zelula populazio bakoitzak CDren garapenean izan dezakeen ekarpena modu independentean aztertzeko.

Ikerketa honetan ikusi dugu adierazpen aldaketa gehienak (16 gene aztertutako 63tik) zelula epitelialetan gertatzen direla, eraldatuta dauden gene guztiak gaixo zeliako aktiboetan gainadierazita daudelarik. Gainera, 16 gene hauetatik 8 gene bakarrik daude desregulatuta CD326+ zeluletan, eta ondorioz ezinezkoa izango zen haien identifikazioa biopsia osoak aztertuz gero. Erantzun immunitarioarekin erlazionatutako geneak zelula epitelialetan gainadierazita egoteak, sistema immunitario innatoak CDn daukan garrantzia

agerian usten du. Kontuan eduki behar dugu gliadinaren peptidoek T-zelulen menpekoa ez den erantzuna martxan ipintzeko gai direla, eta erantzun honek T zelulen aktibazioa eta ehunaren suntsipenari hasiera emateko beharrezkoa den ingurune proinflamatorioa ezartzen duela ²⁹. CD45+ zelulen kasuan, aldatuta dauden 8 genetatik 3 gene zelula hauetan baino ez daude eraldatuta, eta badaude beste bi gene CD326+ zein CD45+ zeluletan aldatuta daudenak.

Honetaz gain, badirudi biopsietan lortzen diren emaitzak zelula epitelialetan lortzen direnen antz handiago daukatela immunitate zelulena baino, biopsia eta CD326+ zelulek aldaketak partekatzen dituzte 6 genetatan eta biopsiak eta CD45+ zelulak 2 gene baino ez. Horrelako emaitza espero genezake kontuan edukita biopsietan dauden zelula immunitarioen proportzioa zelula epitelialena baino txikiagoa dela kontutan hartuta. Gainera, aztertutako 3 lagin motetan adierazpen desberdina aurkezten duen gene bakarrak (*SH2B3*) kontrako joera erakusten du, erantzun zelula espezifikoak kontuan hartu behar direlaren ideai indartuz. Ondorioz, ikerketa honetan zelula populazioak bereiztu eta bakoitza modu independentean aztertzeko aukera ematen digun metodologia berria lagungarria izan daiteke CDn parte hartzen duten eta gene adierazpena biopsia osoetan aztertuta izkutuan mantentzen diren elementu funtzional berriak identifikatzeko.

Beste aldetik, gaixo zeliakoen PBMCetan lortutako emaitzek, biopsia eta PBMC artean adierazpen desberdina duten hainbat gene identifikatu dute eta gene hautagaien azterketa funtzionalak itu ehunean egitearen garrantzia azpimarratzen dute berriro ere. Gure ikuspuntutik, PBMC eta biopsien artean aurkitutako adierazpen desberdintasunek zaila egiten dute PBMCen azterketa CDren markatzaileak aztertzerako orduan.

Gaixo zeliakoen immunitate zeluletan baina ez kontroletan identifikaturiko koadierazpen patroiak, estuki korrelazionaturiko gene azpitaldearekin, gaixoen heste inflamatuaren gertatzen ari den erantzun immunitarioaren adierazgarri dira, gliadinaren aurrean ematen den gene askoren erantzun koordinatua. Zelula epitelialetan antzeko patroirik ez egoteak, zelula ez immunitario hauetan ematen den erantzun immunitario innatoa kaotikoagoa dela adieraz lezake.

Azken hausnarketak

Giza Genomaren sekuentziazio proiektuei esker milioika SNP identifikatu dira. Hauetariko SNP batzuk, tagSNP edo etiketa SNP deitutakoak, markatzaile genetiko gisa erabili dira GWAS azterketetan eta gaixotasun konplexuetan parte hartzen duten milaka suszeptibilitate bariante identifikatzea ahalbidetu dute. CDn burutu dituen 2 GWASak gehi follow-up ikerketak, 26 asoziazio eremu identifikatu dituzte HLAtik kanpo^{71,76,77}. CD eta beste eritasun autoimmunitarioekin asoziatuta dauden aldaerak identifikatzeko helburuarekin burutu den azken proiektu handia Immunochip proiektua izan da. Immunochipean, 12 eritasun immunitariotako GWASetan asoziazioa aurkeztu duten 186 lokus sakonki genotipatu dira, eta CDrekin asoziatuta dauden 13 eremu berri identifikatu dira⁷⁹.

Beraz, 39 eremu daude HLAtik kanpo CDrekin asoziatuta, 57 asoziazio seinale independentetarik eratorriak. Hauetatik, 19 eremuetan gene hautagai bakarrik aurkitzen dugu, baina asoziatutako SNPen artean bakarrik 3 daude lotuta eremu exonikotan kokatuta eta proteina eraldatzen duten aldaerekin. Hala ere, beste zenbait gene hautagai proposatu dira asoziazio seinaleak 5' edo 3' eremu erregulatuarekin gertu egonagatik.

SNP gehienak geneen arteko eremu ez-kodetzaile eta eremu intronikoetan kokatzen diren arren, CDrekin asoziatutako aldaerak eQTL eremuetan aurkitzen direla dirudi. eQTL hauek, mRNA desberdinen adierazpena edota proteinena erregula dezaketen eremu genomikoak dira. Genoma osoan eQTL emaitza multzo baten meta-analisia burutu ostean, non leukozito primarioen gene adierazpenaren adierazle suposatzen diren 1.469 gizakien odol lagin aztertu ziren, genoma osoan CDrekin asoziatuta eta HLAtik kanpo dauden 38 eremutan *cis* motako adierazpen-genotipo korrelazioa aurkitu da⁷⁷. eQTL hauetatik, 20 esanguratsua ziren, zoriz esperotakoak baino gehiago, CDrekin asoziatutako eremuak eQTLetan aberatsak direla adieraziz. Emaitza hauek, identifikatu diren arrisku aldaerek zenbait generen adierazpena eraldatuz CDn eragina izan dezaketela adierazten dute. Hala ere, ebidentzia askoren arabera *cis*-eQTLak asko aldatzen dira ehun batetik bestera, guztiz kontrako efektua izatera iritsi daitezkeelarik.

Hau dela eta, proposatutako gene hautagaien azterketa funtzionalak gaixotasunak kaltetutako ehunean egitea guztiz beharrezkoa da, eta hau izan da tesi honetan aurkeztutako lanaren helburu nagusia. Horretarako, CDn burututako lehenengo GWASean identifikatutako 8 eremuen asoziazio azterketa egin da espainiar populazioan. Eritasunaren garapenaren arabera heste mukosan adierazpen aldatua duten 4 gene identifikatu dira (*IL12A*, *LPP*, *SCHIP1* eta *SH2B3*) eta ikusi dugu asoziatutako aldaeren genotipoak eragina daukala. Gure emaitzen arabera, gene hauek modu konstitutiboan

alduta egon daitezke gaixo zeliakoetan, seguruenik eritasunaren sintomak agertu baino lehen, beraz, erantzukizun primarioa izan dezakete patogenesian.

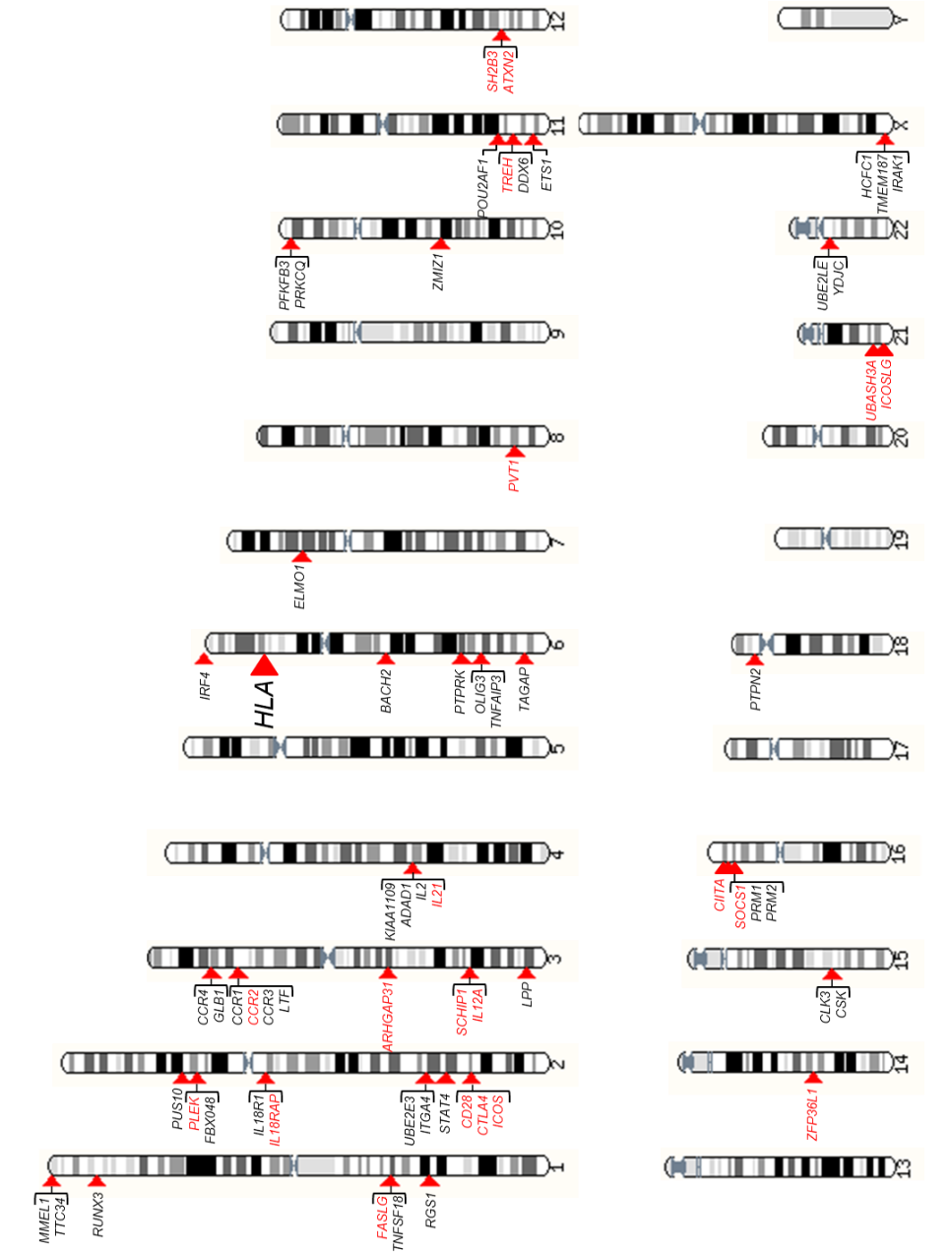


Figure 18: Celiac disease associated regions and proposed candidate genes. Genes highlighted in red showed differential expression in functional analysis.

ImmunoChip proiektuaren ostean, asoziatutako eremuetan aurkitzen diren gene hautagaien zerrenda asko luzatu da. Gene hautagai berri hauen adierazpena aztertu dugu gaixo zeliakoen heste epitelioko diagnostikoan eta tratamenduan ostean, eta emaitzak kontrol ez-zeliakoen adierazpenarekin konparatu ditugu. Honetaz gain, SNPen genotipoak adierazpenean daukan eragina aztertu, eta koadierazpen azterketak egin ditugu. Gene askok, adierazpen aldatua aurkezten dute aztertutako gaixo taldeen artean, gehienak erantzin immunitarioarekin erlazionatuta. *Trans*-eQTL nahiko eta bakarrik 4 *cis*-eQTL aurkitu dira, eta badirudi genotipoaren efektua estimuluaren dependentea dela, talde batetik bestera aldatzen baita. Koadierazpen balioak altutik baxura aldatzen dira gaixo aktibo, gaixo tratatu eta kontrol ez-zeliakoetan hurrenez hurren. Gaixo zeliakoen bi taldeetan koadierazpen maila altua aurkeztu duten eta kontrolen koadierazpenik aurkezten ez duten 18 genez osatutako azpitaldea identifikatu da. Gainera, gene azpitalde honek, 3 SNPen genotipoaren eragin handia pairatzen dute. Emaitza hauen arabera, gaixotasunarekin asoziatuta dauden SNPen eragina lokus hurbilean ematen den kontrol transkripzionalaren ideia sinpletik asko aldentzen da.

Zelula-espezifikoko diren gene adierazpen patroiak aztertzea da hurrengo pausua CDn gene hautagaiak ikertzeko. Heste epitelioko konplexutasun estrukturala dela eta, non zelula mota desberdinak agertzen diren proportzio desberdinetan, biopsia osoetan burutzen diren gene adierazpen azterketetan emaitza aldakorrak lortu daitezke. Heste biopsietako bi zelula populazio nagusiak bananduta, CD45+ immunitate zelulak eta CD326+ zelula epitelialak, eta geneen adierazpena modu independetean aztertuta zelula populazio bakoitzean, gaixo zeliakoen hestean gertatzen denaren irudi argiago lortu dezakegu.

CD gene hautagaiak aztertzeko metodologia berri honi esker, zelula populazio bakoitzaren espezifikoak diren aldaketak identifikatzeko gai izan gara, tesi honen 3. kapituluaz azaltzen den moduan. Laburtuz, 8 genek CD326+ zelula populazioaren espezifikoak diren adierazpen aldaketak aurkezten dituzte, 3 genek CD45+ zelulen espezifikoak direnak, beste 2 genek bi zelula populazioetan aurkezten dituzte desberdintasunak baina ez biopsietan, eta gene bakar bat dago aldatuta 3 lagin motetan.

Amaitzeko, egunerarte burututako azterketek, ikerketa funtzionalak garatzearen garrantzia isladatzen dute, hala nola, arrisku gene hautagaien proposamen arbitrarioa sahistearen garrantzia.

Ondorioak

1. GWAS azterketetan lortutako asoziazio genetikoak errepikatzea guztiz beharrezkoa da, CDren osaera genetikoa sendoa eta populazioen artean konparagarria den zehazteko.
 - a. Tesi honetan lortutako asoziazio emaitzek beste azterketetan lortutakoen joera berdina jarraitzen dute eta CDren genetikari buruz dakigunaren konfirmazio dira, emaitza sendo eta errepikakorrak lortzeko tamaina handiko laginak erabili behar direla indartuz.
 - b. Asoziatu-rikoin seinaleen indarra erlatiboki txikia da, eta askotan, asoziazio esanguratsuak detektatzeko ahalmena populazioen arteko alelo maiztasun txiki- en menpe dago. Honek, lortutako zenbait emaitza negatiboen atzean egon daiteke.

2. CDren partaide funtzionalak identifikatzeko, proposaturiko gene hautagaien eta aldaeren inplikazio funtzionala esperimentalki frogatu behar da kaltetutako ehuna edo modelo zelularrak erabiliz.
 - a. Gehienetan, SNPen asoziazio azterketa handien ostean egiten den gene hautagaien aukeraketa ez da guztiz objektiboa eta gaixotasunaren mekanismo patogenikoei buruz daukagun alde- z aurretiko ezagutzaren menpe egoten da. Benetako partaide funtzionalak identifikatzeko bide alboragabe bakar- ra azterketa funtzionalak dira.

3. CD gaixo aktiboen heste epitelioan ematen den geneen adierazpen aldaketa partzialki eritasunaren ondorioz gertatzen diren zenbait fenomenoren (ingurune proimmflamatorioa adibidez) ondorioa da, eta ez du zertan jatorri genetikoa duen gertaera primarioa isladatu behar.

4. Gene hautagaien eta asoziaturiko eremuetan kokatuta dauden SNPen genotipoen arteko erlazio funtzional desberdinak daude, eta eritasunaren patogenesiarekin duten lotura ez da beti zuzena.
 - a. Eritasun konplexuen biologia, SNP batek eragindako gene baten funtzioa eta gaixotasunaren arteko erlazio zuzena baino askoz konplexuagoa da. Hau dela eta, kontuz ibili behar gara bakarrik asoziatutako eremuetan oinarrituta gene hautagaiak eta mekanismo patogenikoak proposatzerako orduan.
 - b. Asoziaturiko aldaeren efektua, lokus hurbilean ematen den kontrol transkripzionalaren ideia sinpletik aldentzen da. Ingurune probokazio baten aurrean elkarrekintza konplexuak modu koordinatuan mantentzeko ematen den erantzun osasuntsua gure genomari idatzita dago, eta mekanismo hauen haustura izaten da gaixotasunaren garapenean ematen diren gertaera segidaren hasiera.

5. Proposatutako gene hautagaiak eritasunaren garapenean duten inplikazioa aztertzeko heste epitelioan dauden zelula populazioak kontutan edukitzea, orain arte erabili ez den planteamendu metodologiko berria da. Modu honetan, aztertutako bi zelula populazioetan (zelula epitelial CD326+ eta immunitate zelula CD45+) modu diskordantean adierazten diren geneak identifikatzeko gai gara.
 - a. Erantzun immunitarioarekin erlazionatuta dauden geneak zelula epitelialetan gainadierazita agertzeak, CDn sistema immunitario innatoak duen garrantzia adierazten du. Gliadinaren peptidoek, T zelulen menpekoa ez den erantzuna pizteko gai dira, eta honek, T zelulen aktibazioa eta ehunaren suntzipena emateko beharrezkoa den ingurune proinflatatorioa sortu dezake.
 - b. Azterketa funtzionalak burutzeko zelula populazioak banatzearen hurbilketa metodologiko berriak, geneen adierazpena biopsia osoetan aztertzean izkutuan geratzen diren CDren partaide funtzionalak agerian uzten lagundu dezake.

Material osagarria

1. taula osagarria: Immunochipean asoziatuiko 44 SNPen genotipazio zunden diseinua, Fluidigm Biomark dynamic arraiekin (48.48) eta SNPtype assay-ekin burutua.

SNP-ID	Sequence
rs4445406	GGCCTCCTCTGGGCTTGGCTAGGGTCTATGCTGGAGGAAGGGACAGACATCCAGGGGA GGCTGCCTGTGATGAGGCAGGCCCGGCTGTTAGGGGCCGGATGCCGTGGGGAGTGCCGT CTGACCTCTGGGCTGCCACTGGGGATAAAGGGCTGGGT[T/C]CCCGGCTGGGCTGYCGGT CTGAGTTGCGTTCAGTTTTCCACATGGTCTGGCCCTGCCGTTTGATATTCAGCCTCT CAGGCCAAAAATTAAGTAAAGGCGAGGACACAGGAAGTCATCTCCAGCTGGGRGTATTGC
rs12068671	CTGTTTCATGAATATTTGCATAACTGCTAATTAATGCTGGGAGAAGTGAATCACTGGAAT GTCATGGGGGAGATTATTCTACTAGGCTGTGGAAATCTCTTGTGGTGGGCATGGCC TGCTGGGCTTCTACAACAACTCCAGGAGCAGAAGCCC[T/C]TCCATTTCTAGCAGTTGAT GAGGGTTGTAGTGTGATGAACACTGATGACTTGGAAAGCTTTAGAACCTAGATAGGGG TCGGRRGGAAGTGCAGACCCTTGATTTCTTTCCAGGATGGGAAAGAGCTCATACGA
rs859637	CCTTCTGGATGGTATTCCAATGTTAGTTACATTTTCAAGGCCTTTGCTGTGCTGCTCT GGGTCTGYCTGTGCATGTGCTACTCAGAGATAACTCCAGTGTGCTTTAGYTACAGAATT AGAAGATTTCTTCTAGCTCTTTCCCTCAAGYATTC[T/-/C]TTCCCCAAATTC TCTGACTTCCAAGGGCTGTTTTCTAGTTATGTGGTCACTAGAAGATAAGATTTAACT KGAAGCTATAGTGACCATAATACCATAATTCTCCATGATCDGGGCTGCCTTTAGAGCAA
rs10800746	GGGCTTCGGACGAGCCCTGGACCCCGGGCACAGGTACGGCTGCTCTGGAGGGACCCCGG GGCCAGGCAGGAGGGCTTTTTGTTGGGGAGGGGCTGCAATCTAGTATGTCCAAGGGTA CAGAGGCATCARAGCAGCAACCTAGTGGCCGGTAATCCC[C/T]GTCTAGGAAGAACAGGCCA AAARATCATTTCTYTTGCTTTGGAAGGCACTCTTGGCTTCCCTGGGGCACCTCTAAT TTTGCTGCCATGTCTAGCTCAAGAGCATATACTGGTGGAGATGATCCCGAAGATAG
rs13003464	GCGCACCCACATGCCAGCTGATTTTTGATTTTTAGTAGACACAGGGTCTACCAVYGT GCCAGGCTGGTCTTGAATTCCTGGCTGCAAGTATTYCCCCCTCAGCCWCCAGAGT GCTGGGATTGCAGGTGCGCACCACTGYACTCGGCAGAGCA[A/G]TCACTTGGAGAGGACAG GCAGATATTCTCAGTGCCTTGGCTGAAATCCARCCAGGCCARCGGCACCCAKAACATG ATGCCATAGGCACAGGAAGCCTATGGCCAAGAGGCCAGACTTGGTAGTGATTAACAAAA
rs10167650	ATTCTGAATTTTCTACCACCTTACAGAAGGGTATAATATCTCATAATCTGTAGACAGG TTAAATACCAGGCCTGTTCAATTAGTTCTTAACTAACCTTTCTGAGCCTTGTATWTGCAT GTGTGATATGAGCATATTGATACTTAGCTTGCACTGTTG[A/T/G]TGAGGATTAATAAGATA ATGTAATAAAAATGACACTAATGTGGATATCAATATATGGTAGTTTAAACAAGATGTTAA AATGAGATCTTTTTGTCTGTATTCAGCTGAAGCTCTCAAAAAGAACAGAACAAATAGT
rs1018326	AGTAGAAGTATGAGAGGAACACTTATCTTTAAGAATAAGTGCAGACTTTTGTCTCTCTC AAAAATCATCTTCATTATAATTCCAAAAATATAATGAAAGAAAAATTTAAATGATAACC TATCTTATTCTTACAATTAACCTAAGATTTTTTTGAGTTA[T/C]TCCATCTATATACTGGAA TCTTAATACTGACCTTGGCATTAAAGTCCAGGTAGTATTAWTAATATGAATGTTACATTT GTGTTATTATCTACAGTTTAAAGTGCTTTTACTTATATAGTGGATTAATTTCTCTG
rs6715106	ACTTTACGCTTTATATAACTTTACTGAACTGAAATGYGGGTGGTGAGAGGGTGATTAT TTTTCAGTGATTCAAAATAATGCTAATGTTTACTTTAAATCTCTGATGCAATTGTGAG GCACATTTAGAACAAAATCCAGCAAGGAAGCTGACATACT[A/G]TACACAAGACTGGAGGAAA GACGCTCTGACCAATAAAGCATTATWAATTTTACAGCTTGCATGACTGTCATAGTCTT CATTATTTTCATCAGTGTAGTCTAAGCTAKAACACTATTCAGTCTCACATTCKCAGGTTCC
rs6752770	AAAACRGTGATGAATAGGTAAAAGGAGAAAAACAAGTGCATGCTCTGGGACCRGTATTCT GTGGTTGGGCTGAYGGATTTAAACCAGCTAAAAGGTCAGAGCCAAAATAAGTGTATTCT CCTAATGGAACCTCCCTAAAATTTCTAAAACACYGTGAC[A/G]TCCCTTCAAAATTTGATA AAGRITCACTAACAACAGTTTTTCTAATTTGAATTTTCCATGCTTACTACACTTTCTG AAAAATAAACCYGTTGTGTGACAAACCTTTTTCCACTTGGCAAAAACAGAAACCTCTT

rs12998748	TTCTCGCCAATGTCAAACACACTGGCCTGTGTGGGTATCAAACGACTTACTCATAAAT TAAGTAAATAAGCCCAAATGCCTTTCAAGTTCGCATGAKTTTAGTAATCTTTGATGAAAA AAAGACAGTTTTAAATATTGCTAAAATAAACAGCAAT[G/T]TCTTCAGAAATTTAGACATT TGGCCTAAATTAGTCCRCTTTGATAGGTGTTTTAAGATATAGAACTTTGTGGCTGGGTG TGGTGGCTCACGCCTGTAATTCYAGCACITTTGGGAGGCTGAGACAAGCAGATCATGAGGT
rs1980422	CTTCTTATACATTATTGAAGTCTTCATAGTGAACCTACATCTTGCATGTGATGGGTGCTCA GAATATCTTAATGAGTGAATCAATGGAATGGATGGWTAATGAATGACTATCTTTCATTG ATAAATATCCGCAAGCTATTTTGGTTTTGACAAAATTAGA[C/T]GAAACAGGTATTATGAAAA GACTTGGGAAAATTGAGACAATTAGTTAAC TAGATACTATGAAAAGACTTGGGAAAT CATATTTTTAAATATTGAAATGATTAATAATATGCATTTTAAAGACATTGTTTAGGC
rs34037980	TTTTCTTTTAGCAGGAGCAAAGATAATCTAWCAAATCAAAGATGGTCTTTTCAATAAAT AATGCTGGAAGACTGGACATCCACATGCAAAAACCAAGAACAAAAACAAAAACTGAA GCCAGAYACAGGTCTCCACATCACAAAAATTAACITCAA[A/G]TGATAATAGACTTAAATG CTAAACACAAAACATAAAAACCTTAGAAGATAACATTGGAGAAAATCTAGACGCCTTG GGTATGACTATGGCTTTTAGATATAACACCAAAGRCACAATCTATGAAAGAAAATAATTG
rs10207814	GTACCTTCAGTCAATTTGCTGTGAACCTGAAGCTGTTCATAAATGTAGTCTATTAATAA AAGGAAAAAAGGAGAGATATTAATGACCTTAACTTAATATATAGCAACAGTCTGAAA GGCTGTAGATGCACCAATGCAATGTAATGGAAGTGTAT[C/T]CTGTGCTCAGCTGGGGGC CAGGGAAAAGAACTCACTCTCTTCTCTTTCTAGTTTTACATTTCTGGGGTGTATTCTGT AGAACCAAGCTTAAAGAGACAAGACAAGTGAGAGGCTTGTAGATTAGACAATTAGCA
rs4678523	CCACATACATCTGTGAGATCTGGTAAACYGGCAGGATGACAGTATATTCCTCAGTGTTCY TGTGAGGGGCAAGTGAGTTAAACATAAACTCAGAAGTGYGCATGGTYGAAACAAGTCTT CAACAAATGAGAGCTACTGTCCATGTCAGGACAGGAATA[T/C]CAGAAACACCACATCTTT CCTATCTGTGCMCATACACAGCACACACACGTCATGCCTGACTACTGACATGTGTAT AGGCCTGTACTGTGTAATATGTAAGTTCTTATGTTGAACACACAAAACGGGATCCAA
rs2097282	TTTCAAAGCTATAATAATCAAGACAATGTGTTACTGACATAATGATAGACATATTGATC AATATAGACAATGGATGAATGGAGTAGAATTGACAGTCTAGAAAATAAACCTTATATATA TGGTCAATTTATTTCTTTATGTATTCTGTCAATCCAGAATA[C/T]ATAAAGAACCCTTCAACT CAATAACAACAATACCAACAAGGGTGCCAAGCCAATCAATGSAGGAAATAATAGTCTTT TCAAGAAATTTTGTGAGACAATAARTATTATATGCAAAAATAATGAGTTTGAATGGCT
rs7616215	GGGATCTCTGATCTGAGGGTKGCAAAAMTCTGTGGGAGAAGCATGGTTTCCYTGGGTCA CACAWTCACTCACGGCTTCTTGGCTGGGAGTGGGGTTCCCTSGGCTGTGTGTCTC CTGGGTGGGCTGACATCCTGCCCTGTTTTCTACATTCT[C/T]GTGGGTGAACAGTTTCTC YGATCAGTCTAATGTGAGCACCTGGATGTTTAGTTGAAGGTGCTGAATTTACTCACCC CTTTGTCCATCTCCGTGAGTGCCACACCCTGTAGCTGCTTCTAATCGGCCATCTGTGGCC
rs61579022	TTTTCCCCCTCTCTCTATAAATAAGAATACACATTTTGATTAAGTTGGTCTGGG AAAGACACTAATGGAGACTCCAGCTGCCTTCTACCTGTGGCCAATGGGCAGTCAGGGC AGAAAGAGGGGCTTGTGAAGGCAGCCATCCAGATCTGA[G/A]ACTGAGGGTGGAAGGGGGC TGCTGAGAGGCTCAGGGCAACAGGTCCCCTCACCTGCTCTTSAYGGGGTGAGGYAA CTTCAAGTCTTATCTTCTCCATGATTATCCATGTGGAGAAGTTCACCTTCATTAAT
rs1050976	ATTGCGGCAGACAAGCATGGAATAATCAGTGACATCTGATTGCGAGATGASCCTTAKTCA AARGGAAGGGKGGCTTGCATTTCTTGTGTTCTRTAGACTGCCATCATTGATGATCACTG TGAAAATTGACCAAGTATGTGTTACATTTACTGAAATG[C/T]GCTCTTAATTTGTGTAG ATTAGTCTTGTGGAAGACAGAGAAAACCTGCTTTTACAGATTGACACTGACTAGAGTG ATGACTGCTTGTAGGTATGTCTGTGCYATTTCTCAGGGAAGTAAGATGTAATTTGAAGAA
rs12203592	AATATGCTTCTCARGTCTTCTGGGAAACAGATGTTTTGTGGAAGTGAAGATTTTGAAG TASTGCCCTTATCATGTGAACCACAGGGCAGCTGATCTTCTACGGCTTTCYGTATGTGAA TGACAGCTTTGTTTCACTCTTGGTGGGTAAGAAGG[C/T]JAAATYYCCCTGTGGTACTT TTGGTGCCAGGTTTAGCCATATGACGAAGSTTTACATAAAACAGTACAAGTATCTCCATT GTCTTTATGRTCCCTCATGAGTGTITTTCACTTAGTCTGATGAAGGGTTCACTCCAGTCT

rs7753008	TTAACAGGTTTRTTGAGATATAATCAGCATATCATACAAYTCACACCTTTAAAATATACA ATTCAATGTTTTTTAGTATATTCACAGAGTTGTGYGACCATCACCACATCTAATATCAG AACATTTCCATCATCAAAAAAGAAAACCTACACCCATAG[C/T]RGTACTCTTCATTTCTT TCCCCTTCAACCCTTGGCAACCACCAATCTACTTTCTGCTCTGTGGATTTTCATTTCT ACACATTTTGTATAAATGGAATMATCTATTATGAGGCCTTTTGTCTAGCTCTTTCAC
rs55743914	TGCCCTGGTTTTCTGGTTTACTTTTCTGTGAGCATACTACACAGTGCTCCCAACCT GATCCACCATGACATACATATGTGTAACCACAGATATCTCTYGTAGAAGTAAAATG TGTTAGATAAGGTAACCTTTCTTACTCTGCATACAG[C/T]TTCACCTAGATGAAGTGT ACACACACACATACCTATCAGCTCTTATTTTGAAGCTATGTTCTGACCTTAYSAGCCTT GACAAGCTCAAGTAGCATGATGAAGTGCCTCTACAGTTGGTAAAGCTGAAAACTTA GCAAATATTTTTGGCTATAACAGAAAAAAATCTMAAAATTTCCAAATTAATGAAA ATAGTTTCATGATTAGCAGTTTATGCAATACRTGATTAAGTGTCTGTGGGTAGAGGGGA TGCTCAGCAATGTATTAATTAGGCAGGAAAGTCCCCACA[C/T]GTCTCATTTTCATCTGA TTCCCTATGTGTSATGAAGTCAAATCAATAAGTAGGTATATGACTAACTAATSTCTAAG ATGCCCTCCAAATCTAAGATTCTGTGACTACCCGGGGCTCYACCATGTTTGGCTTGCTG
rs72975916	TATTTGTTAAGCTGGTAAACATTTCTGACTTCTTGAGKGTCTAATCTAGGCTTGCMTAC TCTTTTAGGGGATSTTTTAAACTGTTTCTATGCTGGTATATAATCTCATATTCCTCCTA CTGTTATTTTAAAGTACCTYATTTTATTTTATTTTCT[A/G]CATGGTTCAGCCTAGTTGT TTCTATTAAGCCAAGATAACTTCAATTGCTCAACAACAAGGTGAACITGAAGCTGTGAG CATGATCAATCTCATCAAATGGCAAATGCTTATTGAGGGCTATCTATCACATGYGTG
rs17264332	GAAGTACAGCAGTACTGCCAAATAGTGTGACTTCATAAATAGTGAGCAATCATGAAA AAGCAGATTTTCTTTTGTAGAACAATAAACTCACACAAGAAAAGCTACAGTTTCAA TGGCTTTCTGGGAAACATAGTACAAATSTGAAATTCGAA[A/C]JGAATCAGGCTACTTGGTA AGTAGATACTGACTTAACTAAACSAAGTAAGATGCCATCTATTTCAAAGGTGAGAA TAAGAGYGC AAAGGACCAGTGCAGCAGATCCYCTGTATGGTGAGTCAAGTGGTGGT
rs10808568	AAATGAATGACAGAGAAAAARGAAAAAGCCACCAGSCTTGAAAAGTGGTAGCTTCTGC YGCCGACTTCTGCCRGTGCCAAGTCAGAKGGAGCTCTGCGTGTCTCAGTTCACCCCTCG CCAGGCCACACCGCATGCAAATAAGAAGCTGTTTCAGTKT[G/C]JGCCATCTGAGAYGCTGA CATAAAAAAGGAAGAAAGAAAGRAAAAAACCACCGTACCACCAACAAAAAACCCAAA ACTGATTAATCTMGTAGATTRCAGCAGAGCAGGGGGCCTGGTGTGACACATCTGTTT
rs2387397	TCYCAACCCCTGAGCCTCTGTGTCCCATTTATCACAGGGATGGGTAAGGGTCTCCACAG GCACCTGACACAGGAGCTCAGTGGTGAGACCCAGRTGAAGGGTGGGGAGGCCCTTCCAG GGCTGTCTYCTGAGCAGGACAGAGATCTGCGAGAGAGA[A/G]JGTTCCYTGAGGAGGCTC TCTCCCGAGGGGCATGGCTCCAGGCCCTGGGATGAGTGCRTATATGTCCATAGCCRT GGCCCCCTCTGTCTCYGTARTCCCACTAACYRTGCRGTGCTTTTCTTTCCCTCCAGCCTC
rs1250552	ATCCTCCTCTGTCYCTTGATTTTCTCTTTCATCTTCTCAACCTTTCTCTTTCTT TTTATACCCTAGTTTTATTCCTTACCTCTGACAATCTGCTCTTTGACAAAACATTA CTGAGATCTCAGACTGGCATGCAACCTGATTAGTAGAAA[C/T]GTCTTCTGGCTAGTG CAGTTTCATCATGCAGCCTCTTCTCTYCCCTGGCACCTCACAGCAGTCAAGCATCCAY GTGCTCTGCTAACAAGTGCAAGGCCAAATGAAGAAAACTAATAAATGATTACAAAA
rs7104791	ACCATATATRGCCACTGGTTAATGTACAAGATGTGTCTCCRTGCCAATGGTAGACAT CCTATAGCCAAGACTCCATCCAATTGCCTGACCTCCACATGTCCCTACTCTGACTGGAAC CYTGTGGGTAGTTAGGTCTCWGAGGATCAGTCAAGCTCA[G/A]JAGCAAGATTGAATAACCC TGGAAGGTGTGATATTCTATTTCCCAAGGATCAGAGGCATATTTACCATAAAGCTAAT GATGCTTATGCTTCAAGACTGCTCACTACTAAAGCCCTGGTGTAAACATGATCATATGC
rs10892258	TTGGGTGGGGCCCTCRGCCGTGCCACTACCCGGGGAGGGGAAAAAGCTCCAGATCGACT TTTTYGTCTTGATGATGGTGAGAGTCGGYTTGAGATCGACGGCCGCTTCATRGTTCCA GGAGTSGGGACGTACGGGATGGTAGCARGTTTTCAGTTA[C/G/T]YGTGTTTTTCTTTTAA GAGGATTAGTAACAGGGGGAGGGGACGGGGGAAATCCGACTTCTTCCAAAAATCTCA AATCCCGCTGCTTTCTTTCCCGCGCCGGACGGTGCAGCCCGGCACTCCAGGGGA
rs61907765	TTGGGTGGGGCCCTCRGCCGTGCCACTACCCGGGGAGGGGAAAAAGCTCCAGATCGACT TTTTYGTCTTGATGATGGTGAGAGTCGGYTTGAGATCGACGGCCGCTTCATRGTTCCA GGAGTSGGGACGTACGGGATGGTAGCARGTTTTCAGTTA[C/G/T]YGTGTTTTTCTTTTAA GAGGATTAGTAACAGGGGGAGGGGACGGGGGAAATCCGACTTCTTCCAAAAATCTCA AATCCCGCTGCTTTCTTTCCCGCGCCGGACGGTGCAGCCCGGCACTCCAGGGGA

rs3184504	ATACTCTCTCTAAAAGGGGGACTCTGGGGAGACTATAGACAAACTCAGGCCTGGCTGGA AGAAAGAGCMYACGAAACAAGCCTTGAGTACCCCAACYSTGTCGTAGAGYCAAGGCC CAAGCTACAAGCARCTTGCTCCAGCAYCCAGGAGGTCCGGT[C]JGGTGACACRRTTGAGAT GCCTGACAACCTTTACACCTTTGTGCTGAAGGTGAGTGACAAGGCTTTTCAMACCTGGG GCAATACAAATACMTACACATACAGCAGACCCAAACSTGTTCCCTTCCCTCCGCCAGGT
rs11851414	CCCTGCCAGGCAAACCTCGCCYCTCAAACCCTGGCCTCCAGATKACATGTAATCMCCG CCAGSAACTGTGAAACTCAAAGGGTGGGAAGGACGGGGCCAAATTCCTTCAAACCTGGG AGAAATGCCGGAGGAGAAAAGAATCATCYCGCTGCACCAC[T/C]TTSCCAATTGCCCTTCAAAG ACCCAAACTTTGGGGGTTCTTTCTTAAGGCAAAAGAAAAGACTTTTTGAAAAGCAAAT GCTCCGCCCCCTTTACCTTGCAATAARCTTCGCTCAAGTCGAAGATGGTGCCAGACACG
rs1378938	TGGGGTGTTCCTGTCCTGTTGTGGCCCGACCTTCTCTACCAGCCAACAGGACCTG AAATCCAGGRAGATCTGACTTSRAAGTCTACYTTGTGCTGTTCCTCAGCCCCACTSTAG TTCCATCYACCAGTCAACAGCCTCYTCGCCCTCCACT[T/C]CATCTTCTGCTACTGCTC CAGACTTAATTTTTYAAAACACACCACTTCAATCAACAACAAATATTGAAATCCCTGCT GTTTGTAGGACTGAGGTTCAATAAGCAAACACACCCTGTGGAGTCCACAGTTGGG
rs6498114	CCTGTGTCATGACCTGAGGAGGGAGACTATGGGGTYTCATGAGTCTGGCCAATTGCAGT ACTTCTGCMAAAAAAAAAAAAAAAAAAAAAAAAAAGTACTGCTTTAGCTTCTCTTCT RAACAACCAAGGGGCTACCGTAAATCTACCCAAATAGG[G/T]TGGCAGTCTCAAACGTAA ACCAGCCTGATGATTGAGAAGTCCCTCAGATATTGCAGTGCCTTAGGTGCAATTGTGC AGAAAGCCAGTTCTGGWGATGGGAGCTTAGGGGTTTCGCTCTGTGCTACACTGAGCAAC
rs243323	GGTGCCCCCAGCTATGTCCAGCACCTTCAGATTCTATGCCAGTCTCTCGGAACACAG GTTCTAACCCAGGCTCATTCCCTACGGGACATGACATSACCCTGCTCCCTTGATACAG CTCATTCTGACAAGTCTTGTTGTTGGACTGTAAGTCC[A/G]TAGGGCAGGGCCTTGTTG GTTCTAGTTCTGTCCAGAGCCMAGCACAGTCCCTGAACAGACTTTGGTCCCTTAGT AAAAGTTGTGAAATGGCTGCTTTGAGGAACGTTGAGGTTGGTACTATGTTACAGAATC
rs9673543	GGTCCAGAATTCAGACATCTCAGTAGAGGTGTCYAGTAASTAGGCAGCTGAGTATCTGAG TSTGACATTTAGCCTGGAAGTGATATTTGGKATTGTCAGTGTTACAATCAGGAGAC TGGCCAGGATCACTGATTTAGTTTTGCTGCTGCTGTAAC[A/G]AAGTACCATAAACTTTGTA ACTTAAACAGCAYGAGTTCAATTATGGTTTTTCGCYCCAACTCTGTCASCCAGGMTGGA ATGCAGTGGCATGWTACCGCTYATTGCACCCTTTACTTCTAGGCTCAAGTGATCCTAC
rs11875687	CGTAAACTAATTCRCGAGAAAGGAAAAGCGCAGAAACCTCAGGACTTCTTAGTACAGGT ACCTGGAAATTTGCGAYGCAGATGAAAAACATGGGTGAGGGGAGTTTGAGGGATCTAAAG ATGTGGATTCTGATCTCTCTTTGGCTCTACCAACCTACA[T/C]GATTCCAGAYGTAGTCATG ATTTCCAGAAAAGGGTACTACTTGTAGAAACACTATCAGCATCCAGTCCGCCGACA GTCTATGATAAGGGGCACTCAGGGGTCTGTTTTCTGTCTGCAAWATCAGTGGCTCRA
rs62097857	TAGAGAATTATCTGTAGCTCAGTGATTCTCAATYAGAGATTTGGAGGCAGGGGAAGG GCACCTAAGTACAGAATTAAGTCCACACCCRCTATTCCCTGAACGCACACCCACACTG GAACTGCCACCGTGATGAGCCACCTGTTTCAGAGGAAGTG[G/A]GYAAATCTCAAGTGTGC TGCAACTGAAGAAAATGGAACCACTTCTAGCTACTTGTCTTAAGAACACTTCTAA CTGGTAGTAAAGAAAAAAGCATTAAAGAAACATAGTATTTAATAGTAATTTATAACAAA
rs1893592	TGAAACTGAACGCTGATGCTTATTATCTCAGGGAAWTACTACAGGCCAGASCTGAA TTTCCCTGCACTGTCMASAYGGAGTCTGGGATCAAAGACTTTGAAAAYKATCCCCC ATTATCATCGTGTGGCATTTCAGTCCAGAWTTGCAGG[T/A/C]JTTTGGAGACTGTCTAGT AGGAAAGGTAACAATAACARCAACTGATTATGGCTAGCAGGCATCCAGCCTGAGCCCT AGCTACACACCCTTTGATACAGTGTAGCTTTGCTCCAGTCTGCTGAGGAGGCTGGGGC

rs58911644	TACAAAGGCCTACCTGCGTGAATGTCCACCTGCMTGGATGCCACCCRCACGGATGCCA CCTGCACAGTTATCCATCTGAACAGCCCTCTGGTGCCCTCTTTTGCTTGACATGTTCAA ATCTCCCAGTCTACAGTCGATTGTCCTAGTTGGGGTC[A/T]TCTGGGCATGTTATTCTA AAACTTGTTTTTCCCRCTCCACATCAAAGGAGAAAGGCTAGCTTGCTTCTTTGCTAT AGAGCAGGGCCTGAGTGAGGTCAGGACCACAGAGCAGACCCTATGAGTGTGGTCAGGAC
rs4821124	GCCCAGAAACAGCCAGTTACCACCCCTCCCAGAGCCAGAGACCTTAAGCCCTCTGTCA ACTTCTCTTCCCTACAGCCACCAGCCACCATTITGGCCTTGTTCACTATTAATTGTTT AGCATGAAATAAGCATCCATTTCACTACTYCTGTACCYK[T/C]GCAGCCTGACTTCTGTAC CTGCTGCCCTTGCTGCTKCTCCTTTCAATTTGGTGCCATCTAACCAGCCCTGAAAGAAG CCAGAGCTGCTCTCAGCATTCCACCCTGCTCTGCCTGGCTCTTCTTCTCTCTGCT
rs13397	CCTACTTAGCTTTGGGGGTGCTCTTGCCTGGGYTTGTGGTCTCAAGCTGTGTGACC ATCAGCTRCACGGTGGYGTCTTCCAGTGCCTCACAGGCCACTKCTGGTCCAAGGTCT GTGAYGTGCTCCAGTCCACTTTGCGTTTTTGTCTGAC[G/A]CATTCAACACTCACCCAA GATWCCATCCCTCTGGCGGGAAGACGCGTTGAACCCAGGGAAGAACCTGCTGAAAACYRA TGACCCCAAGCATTGAAATGGACTCTGAGATGGCAGCGTGGTGCAGTGCAGACATCCT

2. taula gehigarria: CEGEC laginean asoziazio gehien aurkeztu zuten 50 SNP. Asoazio eremu berean LDn dauden SNPak beltzez adierazita daude.

SNP	CHR	BP	p value	SNP	CHR	BP	p value
rs1105297	1	43184183	7.49E-05	rs13377037	10	1058639	6.08E-05
rs3790568	1	67608648	9.94E-06	rs11250242	10	1090064	4.28E-05
rs58438451	1	67613706	6.93E-05	rs7984030	13	43528420	7.26E-06
rs17129913	1	67616667	6.93E-05	rs11846679	14	21490247	1.34E-06
rs7515827	1	67624052	4.09E-05	rs61975792	14	87652580	3.53E-05
1kg_1_67626426	1	67626426	6.63E-05	rs7145673	14	87652974	4.55E-05
rs7513724	1	67627123	6.93E-05	rs17124095	14	87669880	5.11E-05
rs7536192	1	67627190	4.38E-06	rs1940550	14	87673873	3.82E-05
rs12409092	1	67629219	4.66E-06	rs1570194	14	87696186	3.72E-05
rs1874396	1	67635054	2.84E-05	imm_14_97450261	14	97450261	7.00E-05
rs13001423	2	100291772	4.21E-06	rs16940147	15	56463411	4.97E-05
seq-NOVEL-10879	2	162954368	7.90E-05	rs12593974	15	89952719	1.53E-12
rs11883509	2	181655408	2.86E-08	rs33973997	18	12773118	2.33E-05
imm_3_49278660	3	49278660	7.53E-05	rs45551338	18	12774908	3.79E-05
rs62259859	3	58471021	6.18E-05	rs28676870	18	12779681	2.33E-05
rs63332460	3	161215534	1.08E-06	rs2032174	18	12782954	2.33E-05
rs6451245	5	35988854	5.93E-05	rs2268278	21	35106809	4.54E-05
rs10045685	5	35989454	7.12E-05	rs17648208	22	28828978	1.43E-09
rs10078373	5	35992572	3.78E-05				
rs1478449	5	35992959	6.87E-05				
rs7785711	7	26828165	7.24E-08				
rs4841530	8	11326556	6.82E-05				
rs4840561	8	11343870	5.09E-05				
rs13264212	8	11347259	1.82E-05				
rs13253092	8	11352401	7.06E-05				
rs13251015	8	11352458	7.06E-05				
rs2244234	8	11373877	4.20E-05				
rs7812879	8	11377591	5.11E-05				
rs1478900	8	11385069	6.85E-05				
rs2736343	8	11386659	4.99E-05				
rs9694294	8	11388130	4.83E-05				
rs2553878	8	54218936	1.09E-05				

Bibliografia

1. WK D: Coeliakie. MD Thesis, 1950.
2. S. G: On the celiac disease. St Bart Hosp Rep., 1888, Vol 24, pp 17-20.
3. Feighery C: Fortnightly review: coeliac disease. *BMJ* 1999; **319**: 236-239.
4. Mäki M, Collin P: Coeliac disease. *Lancet* 1997; **349**: 1755-1759.
5. Marsh MN: Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* 1992; **102**: 330-354.
6. Stenman SM, Lindfors K, Korponay-Szabo IR *et al*: Secretion of celiac disease autoantibodies after in vitro gliadin challenge is dependent on small-bowel mucosal transglutaminase 2-specific IgA deposits. *BMC Immunol* 2008; **9**: 6.
7. Mäki M: The humoral immune system in coeliac disease. *Baillieres Clin Gastroenterol* 1995; **9**: 231-249.
8. Husby S, Koletzko S, Korponay-Szabó IR *et al*: European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr* 2012; **54**: 136-160.
9. Dubé C, Rostom A, Sy R *et al*: The prevalence of celiac disease in average-risk and at-risk Western European populations: a systematic review. *Gastroenterology* 2005; **128**: S57-67.
10. Ivarsson A, Persson LA, Nyström L *et al*: Epidemic of coeliac disease in Swedish children. *Acta Paediatr* 2000; **89**: 165-171.

11. Collin P, Reunala T, Rasmussen M *et al*: High incidence and prevalence of adult coeliac disease. Augmented diagnostic approach. *Scand J Gastroenterol* 1997; **32**: 1129-1133.
12. Bodé S, Gudmand-Høyer E: Incidence and prevalence of adult coeliac disease within a defined geographic area in Denmark. *Scand J Gastroenterol* 1996; **31**: 694-699.
13. Murray JA, Van Dyke C, Plevak MF, Dierkhising RA, Zinsmeister AR, Melton LJ: Trends in the identification and clinical features of celiac disease in a North American community, 1950-2001. *Clin Gastroenterol Hepatol* 2003; **1**: 19-27.
14. Plot L, Amital H: Infectious associations of Celiac disease. *Autoimmun Rev* 2009; **8**: 316-319.
15. Cataldo F, Montalto G: Celiac disease in the developing countries: a new and challenging public health problem. *World J Gastroenterol* 2007; **13**: 2153-2159.
16. de Kauwe AL, Chen Z, Anderson RP *et al*: Resistance to celiac disease in humanized HLA-DR3-DQ2-transgenic mice expressing specific anti-gliadin CD4+ T cells. *J Immunol* 2009; **182**: 7440-7450.
17. Rabassa EB, Sagaró E, Fragoso T, Castañeda C, Gra B: Coeliac disease in Cuban children. *Arch Dis Child* 1981; **56**: 128-131.
18. Sagaró E, Jimenez N: Family studies of coeliac disease in Cuba. *Arch Dis Child* 1981; **56**: 132-133.
19. Galvão LC, Gomes RC, Ramos AM: [Celiac disease: report of 20 cases in Rio Grande do Norte, Brazil]. *Arq Gastroenterol* 1992; **29**: 28-33.
20. al-Tawaty AI, Elbargathy SM: Coeliac disease in north-eastern Libya. *Ann Trop Paediatr* 1998; **18**: 27-30.

21. Suliman GI: Coeliac disease in Sudanese children. *Gut* 1978; **19**: 121-125.
22. Khuffash FA, Barakat MH, Shaltout AA, Farwana SS, Adnani MS, Tungekhar MF: Coeliac disease among children in Kuwait: difficulties in diagnosis and management. *Gut* 1987; **28**: 1595-1599.
23. al-Hassany M: Coeliac disease in Iraqi children. *J Trop Pediatr Environ Child Health* 1975; **21**: 178-179.
24. Fasano A, Catassi C: Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum. *Gastroenterology* 2001; **120**: 636-651.
25. Feighery C, Weir DG, Whelan A *et al*: Diagnosis of gluten-sensitive enteropathy: is exclusive reliance on histology appropriate? *Eur J Gastroenterol Hepatol* 1998; **10**: 919-925.
26. Di Sabatino A, Corazza GR: Coeliac disease. *Lancet* 2009; **373**: 1480-1493.
27. Lerner A: New therapeutic strategies for celiac disease. *Autoimmun Rev* 2010; **9**: 144-147.
28. Schuppan D, Junker Y, Barisani D: Celiac disease: from pathogenesis to novel therapies. *Gastroenterology* 2009; **137**: 1912-1933.
29. Maiuri L, Ciacci C, Ricciardelli I *et al*: Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *Lancet* 2003; **362**: 30-37.
30. Hue S, Mention JJ, Monteiro RC *et al*: A direct role for NKG2D/MICA interaction in villous atrophy during celiac disease. *Immunity* 2004; **21**: 367-377.

31. Castellanos-Rubio A, Santin I, Martin-Pagola A *et al*: Long-term and acute effects of gliadin on small intestine of patients on potentially pathogenic networks in celiac disease. *Autoimmunity* 2010; **43**: 131-139.
32. Jabri B, Sollid LM: Tissue-mediated control of immunopathology in coeliac disease. *NatRevImmunol* 2009; **9**: 858-870.
33. Maiuri L, Picarelli A, Boirivant M *et al*: Definition of the initial immunologic modifications upon in vitro gliadin challenge in the small intestine of celiac patients. *Gastroenterology* 1996; **110**: 1368-1378.
34. Martin-Pagola A, Perez-Nanclares G, Ortiz L *et al*: MICA response to gliadin in intestinal mucosa from celiac patients. *Immunogenetics* 2004; **56**: 549-554.
35. Folk JE, Cole PW: Transglutaminase: mechanistic features of the active site as determined by kinetic and inhibitor studies. *Biochim Biophys Acta* 1966; **122**: 244-264.
36. Folk JE, Chung SI: Transglutaminases. *Methods Enzymol* 1985; **113**: 358-375.
37. Heap GA, van Heel DA: Genetics and pathogenesis of coeliac disease. *SeminImmunol* 2009; **21**: 346-354.
38. Caputo I, Barone MV, Martucciello S, Lepretti M, Esposito C: Tissue transglutaminase in celiac disease: role of autoantibodies. *Amino Acids* 2009; **36**: 693-699.
39. Lindfors K, Kaukinen K, Mäki M: A role for anti-transglutaminase 2 autoantibodies in the pathogenesis of coeliac disease? *Amino Acids* 2009; **36**: 685-691.

40. Mazzarella G, Maglio M, Paparo F *et al*: An immunodominant DQ8 restricted gliadin peptide activates small intestinal immune response in in vitro cultured mucosa from HLA-DQ8 positive but not HLA-DQ8 negative coeliac patients. *Gut* 2003; **52**: 57-62.
41. Lundin KE, Sollid LM, Qvigstad E *et al*: T lymphocyte recognition of a celiac disease-associated cis- or trans-encoded HLA-DQ alpha/beta-heterodimer. *J Immunol* 1990; **145**: 136-139.
42. Nilsen EM, Lundin KE, Krajci P, Scott H, Sollid LM, Brandtzaeg P: Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma. *Gut* 1995; **37**: 766-776.
43. Troncone R, Gianfrani C, Mazzarella G *et al*: Majority of gliadin-specific T-cell clones from celiac small intestinal mucosa produce interferon-gamma and interleukin-4. *Dig Dis Sci* 1998; **43**: 156-161.
44. Monteleone G, Pender SL, Alstead E *et al*: Role of interferon alpha in promoting T helper cell type 1 responses in the small intestine in coeliac disease. *Gut* 2001; **48**: 425-429.
45. León AJ, Garrote JA, Blanco-Quirós A *et al*: Interleukin 18 maintains a long-standing inflammation in coeliac disease patients. *Clin Exp Immunol* 2006; **146**: 479-485.
46. Steinman L: A brief history of T(H)17, the first major revision in the T(H)1/T(H)2 hypothesis of T cell-mediated tissue damage. *Nat Med* 2007; **13**: 139-145.
47. Castellanos-Rubio A, Santin I, Irastorza I, Castaño L, Carlos Vitoria J, Ramon Bilbao J: TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. *Autoimmunity* 2009; **42**: 69-73.

48. Harris KM, Fasano A, Mann DL: Monocytes differentiated with IL-15 support Th17 and Th1 responses to wheat gliadin: implications for celiac disease. *Clin Immunol* 2010; **135**: 430-439.
49. Monteleone I, Sarra M, Del Vecchio Blanco G *et al*: Characterization of IL-17A-producing cells in celiac disease mucosa. *J Immunol* 2010; **184**: 2211-2218.
50. Sjöström H, Lundin KE, Molberg O *et al*: Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition. *Scand J Immunol* 1998; **48**: 111-115.
51. Sturgess R, Day P, Ellis HJ *et al*: Wheat peptide challenge in coeliac disease. *Lancet* 1994; **343**: 758-761.
52. Maiuri L, Troncone R, Mayer M *et al*: In vitro activities of A-gliadin-related synthetic peptides: damaging effect on the atrophic coeliac mucosa and activation of mucosal immune response in the treated coeliac mucosa. *Scand J Gastroenterol* 1996; **31**: 247-253.
53. Picarelli A, Di Tola M, Sabbatella L *et al*: 31-43 amino acid sequence of the alpha-gliadin induces anti-endomysial antibody production during in vitro challenge. *Scand J Gastroenterol* 1999; **34**: 1099-1102.
54. Arentz-Hansen H, Körner R, Molberg O *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* 2000; **191**: 603-612.
55. Anderson RP, Degano P, Godkin AJ, Jewell DP, Hill AV: In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat Med* 2000; **6**: 337-342.

56. Fina D, Sarra M, Caruso R *et al*: Interleukin 21 contributes to the mucosal T helper cell type 1 response in coeliac disease. *Gut* 2008; **57**: 887-892.
57. Sollid LM, Thorsby E: HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology* 1993; **105**: 910-922.
58. Greco L, Romino R, Coto I *et al*: The first large population based twin study of coeliac disease. *Gut* 2002; **50**: 624-628.
59. Gutierrez-Achury J, Zhernakova A, Pulit SL *et al*: Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat Genet* 2015; **47**: 577-578.
60. Horton R, Wilming L, Rand V *et al*: Gene map of the extended human MHC. *Nat Rev Genet* 2004; **5**: 889-899.
61. Ludwig H, Polymenidis Z, Granditsch G, Wick G: [Association of HL-A1 and HL-A8 with childhood celiac disease]. *Z Immunitatsforsch Exp Klin Immunol* 1973; **146**: 158-167.
62. Abadie V, Sollid LM, Barreiro LB, Jabri B: Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol* 2011; **29**: 493-525.
63. Karell K, Louka AS, Moodie SJ *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**: 469-477.
64. 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**: 188-192.

65. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E: Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J Exp Med* 1989; **169**: 345-350.
66. Sollid LM: Coeliac disease: dissecting a complex inflammatory disorder. *Nat Rev Immunol* 2002; **2**: 647-655.
67. 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**: 215-220.
68. Ploski R, Ek J, Thorsby E, Sollid LM: On the 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**: 173-177.
69. Lundin KE, Scott H, Hansen T *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**: 187-196.
70. Hovhannisyan Z, Weiss A, Martin A *et al*: The role of HLA-DQ8 beta57 polymorphism in the anti-gluten T-cell response in coeliac disease. *Nature* 2008; **456**: 534-538.
71. van Heel DA, Franke L, Hunt KA *et al*: A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *NatGenet* 2007; **39**: 827-829.
72. Griffith OL, Montgomery SB, Bernier B *et al*: ORegAnno: an open-access community-driven resource for regulatory annotation. *Nucleic Acids Res* 2008; **36**: D107-113.
73. McCarthy MI, Abecasis GR, Cardon LR *et al*: Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**: 356-369.

-
74. Consortium WTCC: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; **447**: 661-678.
 75. Bush WS, Moore JH: Chapter 11: Genome-wide association studies. *PLoS Comput Biol* 2012; **8**: e1002822.
 76. Hunt KA, Zhernakova A, Turner G *et al*: Newly identified genetic risk variants for celiac disease related to the immune response. *NatGenet* 2008; **40**: 395-402.
 77. Dubois PC, Trynka G, Franke L *et al*: Multiple common variants for celiac disease influencing immune gene expression. *NatGenet* 2010; **42**: 295-302.
 78. Kumar V, Wijmenga C, Withoff S: From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases. *Semin Immunopathol* 2012; **34**: 567-580.
 79. Trynka G, Hunt KA, Bockett NA *et al*: Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *NatGenet* 2011; **43**: 1193-1201.
 80. Romanos J, Barisani D, Trynka G, Zhernakova A, Bardella MT, Wijmenga C: Six new coeliac disease loci replicated in an Italian population confirm association with coeliac disease. *JMedGenet* 2009; **46**: 60-63.
 81. Dema B, Martínez A, Fernández-Arquero M *et al*: Association of IL18RAP and CCR3 with coeliac disease in the Spanish population. *J Med Genet* 2009; **46**: 617-619.
 82. Garner CP, Murray JA, Ding YC, Tien Z, van Heel DA, Neuhausen SL: Replication of celiac disease UK genome-wide association study results in a US population. *Hum Mol Genet* 2009; **18**: 4219-4225.

83. Amundsen SS, Rundberg J, Adamovic S *et al*: Four novel coeliac disease regions replicated in an association study of a Swedish-Norwegian family cohort. *Genes Immun* 2010; **11**: 79-86.

84. Bondar C, Plaza-Izurieta L, Fernandez-Jimenez N *et al*: THEMIS and PTPRK in celiac intestinal mucosa: coexpression in disease and after in vitro gliadin challenge. *Eur J Hum Genet* 2013.

85. Hunt KA, Mistry V, Bockett NA *et al*: Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. *Nature* 2013; **498**: 232-235.

86. Hunt KA, Zhernakova A, Turner G *et al*: Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet* 2008; **40**: 395-402.

87. Pruim RJ, Welch RP, Sanna S *et al*: LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010; **26**: 2336-2337.

88. Abecasis GR, Cherny SS, Cookson WO, Cardon LR: Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97-101.

89. Green PH, Cellier C: Celiac disease. *NEnglJMed* 2007; **357**: 1731-1743.

90. Nistico L, Fagnani C, Coto I *et al*: Concordance, disease progression, and heritability of coeliac disease in Italian twins. *Gut* 2006; **55**: 803-808.

91. van Heel DA, Hunt K, Greco L, Wijmenga C: Genetics in coeliac disease. *BestPractResClinGastroenterol* 2005; **19**: 323-339.

92. Zhernakova A, Alizadeh BZ, Bevova M *et al*: Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation

- of type 1 diabetes point to a general risk locus for autoimmune diseases. *AmJHumGenet* 2007; **81**: 1284-1288.
93. Todd JA, Walker NM, Cooper JD *et al*: Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *NatGenet* 2007; **39**: 857-864.
94. Zhernakova A, Festen EM, Franke L *et al*: Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. *AmJHumGenet* 2008; **82**: 1202-1210.
95. Zhernakova A, Elbers CC, Ferwerda B *et al*: Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. *AmJHumGenet* 2010; **86**: 970-977.
96. Trynka G, Wijmenga C, van Heel DA: A genetic perspective on coeliac disease. *Trends Mol Med* 2010; **16**: 537-550.
97. Smyth DJ, Plagnol V, Walker NM *et al*: Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med* 2008; **359**: 2767-2777.
98. Zhernakova A, Stahl EA, Trynka G *et al*: Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet* 2011; **7**: e1002004.
99. 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. *JMedGenet* 2011; **48**: 493-496.
100. Holopainen P, Naluai AT, Moodie S *et al*: Candidate gene region 2q33 in European families with coeliac disease. *Tissue Antigens* 2004; **63**: 212-222.

101. Ward LD, Kellis M: HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; **40**: D930-934.
102. Flicek P, Ahmed I, Amode MR *et al*: Ensembl 2013. *Nucleic Acids Res* 2013; **41**: D48-55.
103. Kent WJ, Sugnet CW, Furey TS *et al*: The human genome browser at UCSC. *Genome Res* 2002; **12**: 996-1006.
104. Almeida R, Ricaño-Ponce I, Kumar V *et al*: Fine mapping of the celiac disease-associated LPP locus reveals a potential functional variant. *Hum Mol Genet* 2013.
105. Östensson M, Montén C, Bacelis J *et al*: A possible mechanism behind autoimmune disorders discovered by genome-wide linkage and association analysis in celiac disease. *PLoS One* 2013; **8**: e70174.
106. Fairfax BP, Humburg P, Makino S *et al*: Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 2014; **343**: 1246949.
107. Fernandez-Jimenez N, Castellanos-Rubio A, Plaza-Izurieta L *et al*: Coregulation and modulation of NFκB-related genes in celiac disease: uncovered aspects of gut mucosal inflammation. *Hum Mol Genet* 2013.
108. Plaza-Izurieta L, Fernandez-Jimenez N, Irastorza I *et al*: Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in celiac disease. *Eur J Hum Genet* 2015; **23**: 1100-1105.
109. Mayhew TM, Myklebust R, Whybrow A, Jenkins R: Epithelial integrity, cell death and cell loss in mammalian small intestine. *Histol Histopathol* 1999; **14**: 257-267.

110. Purcell S, Neale B, Todd-Brown K *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559-575.
111. Núñez C, Dema B, Cénit MC *et al*: IL23R: a susceptibility locus for celiac disease and multiple sclerosis? *Genes Immun* 2008; **9**: 289-293.
112. Weersma RK, Zhernakova A, Nolte IM *et al*: ATG16L1 and IL23R are associated with inflammatory bowel diseases but not with celiac disease in the Netherlands. *Am J Gastroenterol* 2008; **103**: 621-627.
113. Einarsdottir E, Koskinen LL, Dukes E *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; **10**: 8.
114. Warde-Farley D, Donaldson SL, Comes O *et al*: The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 2010; **38**: W214-220.
115. Parmar AS, Lappalainen M, Paavola-Sakki P *et al*: Association of celiac disease genes with inflammatory bowel disease in Finnish and Swedish patients. *Genes Immun* 2012; **13**: 474-480.
116. Lawlor G, Peppercorn MA: New genetic data support an association between celiac disease and inflammatory bowel disease. *Inflamm Bowel Dis* 2011; **17**: E80-81.