

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

**ERREGULATZAILE GENETIKO ETA
EPIGENETIKOEN IDENTIFIKAZIOA
ERITASUN ZELIAKOAN
HURBILKETA KONPUTAZIONAL
ETA ESPERIMENTALEN BIDEZ**

Doktorego tesia

Irati Romero Garmendia

Jose Ramon Bilbaoren eta Nora Fernandez-Jimenezen
zuzendaritzapean

Leioa, 2019

Lan hau Irati Romerok jasotako Euskal Herriko Unibertsitateko UPV/EHU Beka Predoktoralaren (PIF2014/408), PI13/01201 eta PI16/00258 (Espainiako ekonomia eta Lehiako Ministerioak eta Europako Eskualde Garapenerako Funtsak ko-finantziatutako Zientzia-ikerketa, Garapen eta Teknologia Berrikuntzarako 2013-2016 Plan Nazionalaren barruan), eta Eusko Jaurlaritzako Osasun Saileko 2011/111034 ikerketa proiektuek jasotako diru laguntzei esker burutu da. SGIkerrek (UPV/EHU, MICINN, GV/EJ, EGEF eta EGIF) emandako laguntza teknikoa eta lana eskertzen da.

*“Hobe egarri izan eta urik ez
Ura izan eta egarririk ez baino.”*

Anari.

Lehenengo eta behin, eskerrak eman nahi dizkiet urte hauetan laborategian topatu ditudan lankide eta lagunei, denak izan baitzarete modu batean edo bestean lan honen parte.

Eskerrik asko nire zuzendari izan diren Buli eta Norari. Buli mila esker Immunogenetics taldean sartzeko eta tesia egiteko aukera emateagatik. Zuri esker ikertzaile karrera hasteko aukera izan dut eta zu gabe guzti hau ez zen posible izango. Nora, mila esker laborategian master ikasle nintzela egin zenidan harrera goxoagatik eta urte hauetan nire lan eta etorkizunarekiko izan duzun arretagatik. Mila esker nire tesiko atzerriko estantzia posible egiteagatik eta bertan eskaini zenidan esperientzia pertsonal zein profesional aberasgarriagatik.

Ainarari, zure alaitasuna eta laguntza ezinbestekoak izan dira tesiko bigarren erdi honetan. Hamar mila gauzetara zabiltzan harren beti zara hamar mila eta bat gauzetan aritzeko kapaza. Eskerrik asko eskaini didazun denboragatik. Izortzeri, eskerrik asko zure laguntza eta babesagatik. Asko ikasi dut zure lan egiteko moduagatik, beti txukun.

Duela gutxi hasi diren Ane eta Maialeni ere eskerrak eman nahi dizkizuet. Aldrebesturik nuen egun bat baino gehiagori buelta eman diozue zuen gogo eta energiekin. Animo bioi, paregabeak zarete!

Eskerrik asko ere laborategia utzi zenuten arren gertu zaudetenoi. Leti, oso pertsona alaia zara, beti zerbait berria ikasteko eta irakasteko prest dagoena. Koldo, mila esker ezagutzen ez nuen mundu bat aurkezteagatik. Lan bikaina egiten duzu, jartoetan jartoena. Eskerrik asko mundu kaotiko horretarako sarrera erakusteagatik eta beti laguntzeko prest azaltzeagatik. Teresa, laborategian zein laborategitik kanpo pasa ditugun momentu on guztiengatik, beti gogoratuko ditugu par artean. Donatella grazie mille per la tua gioia e amicizia, buona fortuna con tutto.

Eskerrak gurutzetako jendeari eta kolaboratzaileak izan diren medikuei. Baita, paziente zein haien familiei beraien eskuzabaltasunagatik. Gure Bunkerreko kideei, bereziki Angela, Idoia eta Nereari, zuen hurbiltasunagatik eta pasatako momentuengatik. Mila esker, Lyonen ezagututako Dr. Herceg eta Epigenetics-eko talde osoari zuen harreragatik eta lehenengo egunetik taldeko parte egiteagatik.

Doktoretzako urte hauetan zehar pertsona asko sentitu ditut ondoan, ezin izango nukeen lan hau gauzatu familia eta lagunak babesik gabe. Eskerrak eman nahi dizkizuet guztioi, aipatuko ditudanei eta aipatu gabe geratuko direnei ere.

Lehenik eta behin eskerrak eman nahi dizkiot nire kuadrilari. Ane, Claudia, Goizane eta Maialen, mila esker azken aldian gutxiago ikusi garen arren beti presente egoteagatik. Zuen animoak oso garrantzitsuak izan dira niretzat. Mila esker azken hilabeteetan urrun egon naizen arren nire lekutxoak gordetzeagatik eta maitatua sentiarazteagatik.

Bilboko lagunei, batez ere azken urteetan nire bizitzan agertu den “Bilbo Peñari”. Talde heterogeneoa, bata besteengandik ikasteko prest dagoena. A Miguel el Murciano Cubano por alegrarnos con sus chistes e historias curiosas; Marc, zuen etxean beti ongietorria sentiarazteagatik; Itzi eta Aritz, asko baloratzen ditut zuen ontasuna eta giro ona jartzeko duzuen ahalmena; Aratz, 10 urte dira ezagutu ginela eta asko pozten nau zu ondoan izaten jarraitzeak, eskerrik asko mila momentutan indarra emateagatik; gracias Shukhrat por tu amabilidad y cariño recibidos durante este tiempo, siempre detallista. Mila esker Txekas zaren laguna izateagatik. Beti sentitu zaitut hurbileko eta momentu askotan izan zara euskarri, askotan konturatu ez zaren arren. Mila esker ostiraletako bisitengatik eta high-ko ekipoko buru izateagatik.

Bereziki eskertu nahiko nituzke Manuelitak, azken urteetan familia izan baikara. Araia, Urreta, Ainara eta Amaia mila esker elkarbizitzan pasa ditugun momentu

alaiengatik. Urreta, etxera etorri zinen lehenengoan argia joan zen arren kolorez bete zenuen etxea. Mila esker izan ditugun elkarrizketa eta barre algarengatik, ez daitezela bukatu! Araia beti eskuzabal, eskerrik asko zure tratu goxoagatik eta etxe-giroa sortzeagatik, kasu honetan ere ez daitezela barreak eta dantzak bukatu! Ainara, badira urte batzuk ezagutu ginela baina denbora pasa ahala geroz eta pertsona interesgarriago bat ezagutzen dut zurekin. Eskerrak eman nahi dizkizut izan zaren lagunagatik, beti ondoan egoteagatik, eta batez ere feminismoaren munduan izan zaren erreferentziagatik. Asko ikasi dut zurekin. Amaia, “siamesaatiempocompleto” azken hilabeteetan herrialde desberdinetan bizi garen arren bata bestearengandik beti bezain gertu jarraitu dugulako. Urte hauetan gertuen izan dudana pertsona izan zara, zentzu askotan, eta eskerrak eman nahi dizkizut nire bizitza aberastu duzun bezala aberasteagatik. Ez nintzateke gaur egun naizena izango zure ekarpenengatik izan ez balitz, eredu zara zentzu askotan. Azken urteetako malko eta algaren testigu izan zara, beti zintzo, entzuteko prest eta pertsonak handi bihurtzeko duzun don horrekin. Lan hau zurea ere bada.

Manu, karreran zehar abentura batzuk bizi izan ditugu eta oraindik ere ez gara nekatu! Eskerrik asko beti gertu egoteagatik. Animo tesiarekin, aurrera! Rodri, me ha encantado tenerte de flatemate y reírme tantas veces contigo, ve eligiendo SPA! Iosu eta Ainhoa, nirekin izan duzuen pazientziagatik eta eman didazuen maitasun eta alaitasunagatik. Guztiak aipatu ezin ditudan arren badira beste asko eskertu nahiko nituzkeenak.

A toda la familia de David por acogerme y hacerme una más de la familia. Sobre todo a Mari Carmen y a Antonio por todo el cariño y generosidad que me habéis ofrecido.

Azkenik nire familia eskertu nahi dut. Aitona eta amona pertsona eredugarriak zarete. Zuek zarete familia lotuta mantentzen duzuenak eta guztiok zaintzen gaituzuenak. Beti onak, gehien behar duenari eskua luzatzen eta jendearen onena

sustatzen. Aita eta ama, nigatik egin duzuen guztiagatik, gehien behar izan zaituztedanetan beti egon zarete ondoan, beti ulerkor eta nire ongizatean pentsatuz. Eskerrik asko pasatu dizkidazu balioengatik, pasio eta indarragatik. Harro nago zuetaz. Izaro, eskerrik asko nitaz arduratzeagatik eta ditugun jolas eta elkarrizketengatik, asteburuak poztu dizkidazu askotan. Zu ere pertsona indartsua zara, ideiak argi dituena eta horiek defendatzen dituena. Zu ere eredu izan zara askotan.

David, nire bizikide, familia eta lagun minenari. Egunak animatzen dituzu zure umore eta alaitasunarekin, mundua dantzan jartzen duzu eta horrela jarri ninduzun ni ere, dantzan. Eskerrik asko lerro zuzen horretatik kanpo geratzen diren burbuilatxoak sortzeagatik eta errespetuan oinarritzen den harreman polit honengatik. Zure indar, babes eta alaitasuna guztiz beharrezkoak izan dira askotan esan ez dizudan arren. Eskerrik asko nirekin izan duzun pazientziagatik, eta gehien behar izan dudanean eman didazun energia eta bultzadagatik. Beti eman didazu zure onena, bueltan ezer espero gabe. Orain zure txanda da, ondoan izango nauzu nik zu izan zaitudan bezala. Hau hasi besterik ez da egin!

LABURDURAK	1
GLOSATEGIA	3
ARGITALPEN ORIGINALEN ZERRENDA	5
PROIEKTUAREN JUSTIFIKAZIOA ETA GARRANTZIA	7
SARRERA	9
1. Eritasun zeliakoa	11
1.1. Ezaugarri klinikoak eta diagnosia	11
1.2. Epidemiologia	13
1.3. Tratamendua	14
2. Eritasun zeliakoaren patogenesisia	16
2.1. Glutena	17
2.2. Transglutaminasa	18
2.3. Immunitate-sistema egokitzaila	20
2.4. Berezko immunitate-sistema	20
2.5. Beste bidezidor biologikoak	21
3. Eritasun zeliakoaren Genetika	24
3.1. HLA eremuaren partaidetza	26
3.2. HLA eremutik kanpoko arrisku genetikoa	29
4. Gene erregulazioa	33
4.1. Transkripzio faktoreak	34
4.2. mikroRNak	35
4.3. Kromatinaren egitura	38
4.4. DNaren metilazioa	42
HELBURUAK	45
MATERIALAK ETA METODOAK	49
1. Materialak	51
1.1. Pazienteak	51
1.1.1. Onarpen etikoa	51
1.1.2. Pazienteak eta biopsia laginak	51
1.2. Zelula lerroak eta zelula kultura	52
1.3. Zelulen eta biopsien kitzikapena	53
1.4. DNA eta RNA erauzketa	54
1.5. Datu multzoak	55
2. Metodoak	56
2.1. Genoma osoko koadierazpena EZn	56
2.1.1. TF eta miRNAen identifikazioa koadierazpen aldaketetan	56
2.1.1.1. Koadierazpen analisia	56
2.1.1.2. Erregulatzaile hautagaien aukeraketa	58
2.1.2. Hautagaien egiaztatze esperimentalak	58
2.1.2.1. Geneen adierazpen analisia	58

	2.1.2.1.1. Gene hautagaiak eta entseguak	58
	2.1.2.1.2. Alderantzizko transkripzioa edo cDNAREN	
sintesia		61
	2.1.2.1.3. Adierazpen analisia: qPCR	62
	2.1.2.1.3.1. <i>Fluidigm BioMark dynamic array</i>	
sistema		62
	2.1.2.1.3.2. <i>Eco Real-Time PCR</i> sistema	63
	2.1.2.2. TFen kokapen zelularra	63
	2.1.2.2.1. Immunofluoreszentzia azterketak	64
	2.1.2.2.2. Nukleo eta zitoplasmako proteinen erauzketa	64
	2.1.2.2.3. <i>Immunoblot</i> analisia	65
	2.1.2.3. Kromatinaren immunoprezipitazioa	66
	2.1.2.4. miRNA ituen adierazpena EZn	67
	2.1.3. Analisi estatistikoak	67
2.2. TADak EZn		67
	2.2.1. Kromatinako 3D egitura aldaketak EZn	68
	2.2.1.1. Koadierazpen analisia	68
	2.2.1.2. Ezaugarri genomikoan gainezarpena	69
	2.2.2. Hautagaien egiaztatze esperimntala	71
	2.2.2.1. Kromatinaren eskuragarritasuna	71
	2.2.2.2. DNasa I-ekiko hipersentikortasun eremuen ezabatzea	72
	2.2.2.2.1. sgRNA diseinua	73
	2.2.2.2.2. sgRNA klonazioa	76
	2.2.2.2.3. Transformazioa eta hautespena	77
	2.2.2.2.4. Zelula lerroen edizioa	77
	2.2.2.2.5. Klonen hedapena	78
	2.2.2.3. Geneen adierazpen analisia	78
	2.2.2.4. Zelula lerroen karakterizazioa	80
	2.2.3. Analisi estatistikoak	80
2.3. Metilazio patroien aldaketa akutuak EZn		81
	2.3.1. Bisulfito bidezko konbertsioa	81
	2.3.2. Hautatutako eskualdeen anplifikazioa, kuantifikazioa, purifikazioa eta normalizazioa	82
	2.3.3. Metilazio azterketa belaunaldi berriko sekuentziazioa erabiliz	85
	2.3.4. Analisi estatistikoak	86
EMAITZAK		87
1. Genoma osoko koadierazpena EZn		89
1.1. Gliadinaren kitzikapenak eragindako koadierazpen aldaketak EZn		89
1.2. Koadierazpen aldaketetan eragina duten elementu erregulatzailen identifikazioa		92

1.3. Erregulatzailer hautagaien eta TFen itur geneean aukeraketa ondorengo analisietarako	95
1.4. Transkripzio faktoreak	95
1.4.1. Moduluetako funtzio biologikoak	97
1.4.2. TF hautagaien eta haien itur geneen adierazpena EZn	99
1.4.3. TF hautagaien kokapena lerro zelularretan	101
1.4.4. TF hautagaien itur geneekiko lotura	104
1.5. mikroRNak	105
1.5.1. miRNA hautagaien adierazpena EZn	105
1.5.2. miRNAen itur geneen adierazpena EZn	106
2. TADak EZn	111
2.1. Eskualde eta gene hautagaien identifikazioa eta karakterizazioa	111
2.2. Identifikatutako eskualdeen kromatinaren eskuragarritasuna	111
2.3. Eskualde hautagaien edizioa	118
2.3.1. Gene edizioen konfirmazioa	118
2.3.2. Adierazpen eta koadierazpen analisiak	120
2.4. Zelula lerroen genotipazioa	121
3. Metilazio patroien aldaketa akutuak EZn	122
3.1. Bisulfito konbertsioa eta NGS metilazio-espezifikoa	122
3.2. Metilazio aldaketak EZn	124
EZTABAIDA	131
ONDORIOAK	153
BIBLIOGRAFIA	159

LABURDURAK

3D	Hiru dimentsioko
AGA	(Anti-Gliadin Antibodies) anti-gliadina antigorputzak
APC	(Antigen Presenting Cell) zelula antigeno aurkezlea
EZ	Eritasun Zeliakoa
ChIP	Kromatinaren immunoprezipitazioa
ChIP-seq	Kromatinaren immunoprezipitazio-sekuentziazioa
CTCF	<i>CCCTC-binding factor</i>
DC	(Dendritic Cell) zelula dendritikoa
DCGL	(Differential Co-expression Analysis and Differential Regulation Analysis of Gene Expression Microarray Data) Gene adierazpenaren microarray datuen koadierazpen diferentzialaren analisia eta erregulazio diferentzialaren analisia
DKG	Diferentzialki koadierazitako genea
DHS	(DNase I hypersensitive site) DNasa I hipersentikortasun eremua
DMEM	(Dulbecco's Modified Eagle's Medium) Dulbeccok modifikatutako Eagle medioa
DMMT	(DNA Methyltransferase) DNA metiltransferasa
DMP	(Differentially Methylated Position) diferentzialki metilatutako posizioa
DMR	(Differentially Methylated Regions) diferentzialki metilatutako eskualdea
EAE	(Experimental Autoimmune Encephalomyelitis) entzefalomiелitis autoimmune esperimentalak
EGF	(Epidermal Growth Factor) hazkunde faktore epidermikoa
EMA	(anti-endomysium autoantibodies) anti-endomysio autoantigorputzak
eQTL	<i>Expression Quantitative Trait Locus</i>
ESPGHAN	(European Society for Pediatric Gastroenterology, Hepatology and Nutrition) Gastroenterologia, Hepatologia eta Elikadura Pediatrikoaren Europako Elkarteak
FBS	(Fetal Bovine Serum) Behi-sero fetala
FDR	(False Discovery Rate) aurkikuntza faltsuen proportzioa
FPKM	<i>Fragments Per Kilobase Million</i>
GEO	<i>Gene Expression Omnibus</i>
GGD	Glutenik Gabeko Dieta
GO	<i>Gene Ontology</i>
GWA	(Genome Wide Association) Genoma osoko asoziazioa
HLA	(Human Leucocyte Antigen) Giza Antigeno Leukozitarioa
IBD	(Inflammatory Bowel Disease) Hesteetako hanturazko gaixotasuna
IELS	(Intraepithelial Lymphocytes) Epitelio barneko linfozitak
IFN	Interferoia
LD	(Linkage Disequilibrium) lotura desoreka

lncRNA	(Long non-coding RNA) RNA ez-kodetzaile luzea
MHC	(Mayor Histocompatibility Complex) Histokonpatibilitate Konplexu Nagusia
miRNA	mikroRNA
MMP	(Matrix Metalloproteinase) matrizeko metaloproteinasa
ncRNA	(Non-coding RNA) RNA ez-kodetzailea
NEAA	(Non-Essential Amino Acids) Aminoazido ez-esentzialak
NGS	(Next-Generation Sequencing) Belaunaldi berriko sekuentziazioa
NK	<i>Natural Killer</i>
NKR	(NK Receptor) NK hartzailea
PAM	<i>Protospacer adjacent motif</i>
PCR	(Polymerase Chain-Reaction) Polimerasaren kate-erreakzioa
PMA	(Phorbol 12-Myristate 13-Acetate) forbol 12-miristato 13-azetatoa
pre-miRNA	(Precursor miRNA) miRNA prekurtsorea
pri-miRNA	(Primary miRNA) miRNA primarioa
PT-BSA	(Pepsin-Trypsin-digested Bovine Serum Albumin) Pepsinaz eta tripsinaz liseritutako betabereen sero albumina
PT-G	(Pepsin-Trypsin-digested Gliadin) Pepsinaz eta tripsinaz liseritutako gliadina
qRT-PCR	(Quantitative Real-Time PCR) denbora errealeko PCR kuantitatiboa
RCD	(Refractory CD) EZ errefraktarioa
RISC	<i>RNA-Induced Silencing Complex</i>
RNA-seq	RNAren sekuentziazioa
ROS	oxigeno espezie errektiboak
RPMI	(Roswell Park Memorial Institute) Roswell Park Memorial Institutua
SNP	(Single Nucleotide Polymorphism) Nukleotido Bakarreko Polimosfismoa
SRA	<i>Sequence Read Archive</i>
T1D	(Type 1 Diabetes) 1 motako diabetesa
TAD	<i>Topologically Associating Domain</i>
TBST	(Tris-Buffered Saline with 0.05% Tween) %0,05 Tween 20-dun Tris-salino tanpoia
TF	Transkripzio faktorea
TFBS	(TF-Binding Sites) TF lotura gunea
TG2	(Tissue Transglutaminase type 2) 2 motako transglutaminasa
TGA	(Anti-Tissue Transglutaminase Autoantibodies) anti-transglutaminasa autoantigorputzak
TLR	(Toll-Like Receptor) Toll-moduko hartzailea
tTG	(Transglutaminase) transglutaminasa
UCSC	(University of California Santa Cruz) California Santa Cruz-eko unibertsitatea
WGCNA	<i>Weighted Correlation Network Analysis</i>
WT	(Wild Type) Jatorrizko

Genoma osoko ko-aderazpena eritasun zeliakoan atalerako GLOSATEGIA

Esperimentuak

- **Esperimentu akutua:** *in vitro* gliadina kitzikapenaren aurreko erantzunaren genomako adierazpen azterketa. Esperimentua bi zatitan banatu ziren eta 10 µg/ml gliadinarekin edo gliadina gabe 4 orduz inkubatu ziren glutenik gabeko dietan (GGD) zeuden 10 pazienteren duodenoko biopsien datuak dituen Human U133 Plus 2.0 Array-ean (Affymetrix, Santa Clara, CA, USA) datza. Datuak EBI Array Express datu-basetik (<http://www.ebi.ac.uk/microarray-as/ae/>) lortu ziren, E-MEXP-1823 esperimentu kodetik hain zuzen (Castellanos-Rubio et al., 2008).
- **Epe-luzeko esperimentua:** dietako gliadinaren kitzikapen kronikoaren aurreko erantzunaren genomako adierazpen azterketa. Esperimentua eritasun zeliako (EZ) aktiboa duten 9 pazienteren eta bi urtez baino gehiagoz GGDn egon diren 9 pazienteren duodenoko biopsien datuak dituen Human U133 Plus 2.0 Array-ean (Affymetrix, Santa Clara, CA, USA) datza. Datuak EBI Array Express datu-basetik (<http://www.ebi.ac.uk/microarray-as/ae/>) lortu ziren, E-MEXP-1828 esperimentu kodetik hain zuzen (Castellanos-Rubio et al., 2008).

Laginak

- **Gliadinarekin kitzikatutako laginak.** Esperimentu akaturako erabiliak; gliadinarekin inkubatutako GGDn dauden pazienteen biopsiak.
- **Kitzikatu gabeko lagina.** Esperimentu akaturako erabiliak; gliadina gabe inkubatutako GGDn dauden pazienteen biopsiak.

- **Lagin aktiboak.** Epe-luzeko esperimenturako erabiliak; endoskopia egin zen momentuan glutendun dietan zeuden EZdun paziente aktiboak.
- **GGD laginak.** Epe-luzeko esperimenturako erabiliak; bi urtez baino gehiagoz GGDn egon diren EZdun pazienteen biopsiak.

Konparaketak

- **Gliadinarekin kitzikatutako laginak vs. kitzikatu gabeko laginak eta alderantziz.**
- **Lagin aktiboak vs. GGD laginak eta alderantziz.**

Behaketak

- **Koadierazpen moduluak:** ko-adierazitako gene taldea. Multzokatzea, lagin anitzetan antzeko adierazpen patroiak dituzten geneak taldekatzeko egiten da, horrela gene pareak egin ordez gene taldeak egiteko. WGCNA paketeak koadierazpen moduluak sortzen ditu adierazpen datuekin sortutako korrelazio sare batean multzokatze hierarkikoa erabiliz. Multzokatze hierarkikoa multzo bakoitza azpi-multzoetan banatzen du zuhaitz bat sortuz, zuhaitzeko adar bakoitzak koadierazpen modulu bat irudikatzen duelarik. Moduluak, orduan, adarrak altuera jakin batean moztuz definitzen dira. Modulu hauek erregulatzailerak eta aberasgarritasun funtzionala identifikatzeko erabili daitezke.
- **Diferentzialki ko-adierazitako geneak (DKGak):** bi egoeren artean koadierazpen moduluko kideekin koadierazpen erlazioak aldatzen dituzten geneak. DKGak oso korrelazionatuta daude egoera batean, baina ez daude korrelazionatuta beste egoeran.
- **DKGen moduluak:** DKGek sortutako koadierazpen moduluak. Bi egoera desberdinen arteko konparazioz identifikatutako DKGak beraien artean sortzen zituzten koadierazpen moduluen arabera taldekatu ziren.

ARGITALPEN ORIGINALEN ZERRENDA

Tesia hurrengo argitalpenetan oinarritu da:

Romero-Garmendia I, Garcia-Etxebarria K, Hernandez-Vargas H, Santin I, Jauregi-Miguel A, Plaza-Izurieta L, Cros MP, Legarda M, Irastorza I, Herceg Z, Fernandez-Jimenez N, Bilbao JR. Transcription factor binding site enrichment analysis in co-expression modules in celiac disease. *Genes*. 2018 May; 10,9(5).

Romero-Garmendia I*, Jauregi-Miguel A*, Santin I., Bilbao J. R., Castellanos-Rubio A. Subcellular Fractionation from Fresh and Frozen Gastrointestinal Specimens. *J. Vis. Exp.* 2018 Jul 15;(137).

Fernandez-Jimenez N*, Garcia-Etxebarria K*, Plaza-Izurieta L*, **Romero-Garmendia I**, Jauregi-Miguel A, Legarda M, Ecsedi S, Castellanos-Rubio A, Cahais V, Cuenin C, Degli Esposti D, Irastorza I, Hernandez-Vargas H, Herceg Z, Bilbao JR. The methylome of the celiac intestinal epithelium harbours genotype-independent alterations in the HLA region. *Sci Rep*. 2018 Under review.

Tesi honetan zehar, hurrengo argitalpenetan ere parte hartu dut:

Santin I, Jauregi-Miguel A, Velayos T, Castellanos-Rubio A, Garcia-Etxebarria K, **Romero-Garmendia I**, Fernandez-Jimenez N, Irastorza I, Castaño L, Bilbao JR. A Celiac Disease Associated lncRNA Named HCG14 Regulates NOD1 Expression in Intestinal Cells. *J Pediatr Gastroenterol Nutr*. 2018 Aug;67(2):225-231.

Tentelier C, Barroso-Gomila O, Lepais O, Manicki A, **Romero-Garmendia I**, Jugo BM. Testing mate choice and overdominance at MH in natural families of Atlantic salmon *Salmo salar*. *J Fish Biol*. 2017 Apr;90(4):1644-1659.

Garcia-Etxebarria K, Jauregi-Miguel A, **Romero-Garmendia I**, Plaza-Izurieta L, Legarda M, Irastorza I, Bilbao JR. Ancestry-based stratified analysis of ImmunoChip data identifies novel associations with celiac disease. *Eur J Hum Genet.* 2016 Dec; 24(12):1831-1834.

Plaza-Izurieta L, Fernandez-Jimenez N, Irastorza I, Jauregi-Miguel A, **Romero-Garmendia I**, Vitoria JC, Bilbao JR. Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in celiac disease. *Eur J Hum Genet.* 2015 Aug;23(8):1100-5.

PROIEKTUAREN JUSTIFIKAZIOA ETA GARRANTZIA

Eritasun zeliakoa (EZ), prebalentzia altuko gaixotasun immunitario kronikoa da. Prebentzioa gaixotasuna desagerrarazteko ezinbestekoa izango dela uste da, eta beraz, aurreikuspenarako zein diagnosi goiztiarrerako tresna eraginkorrak garatu beharko dira. Denboraren eskalan, sintoma klinikoaren presentziak gaixotasunaren garapen prozesua oso aurreratuta dagoela esan nahi du. Gaixotasunaren estadio aktibo honen aurrerik, ordea, predisposizio genetikoak duten indibiduoetan erantzun immunitarioak eragindako heste-ehunaren suntsipenaren adierazle diren ehuneko transglutaminasaren aurkako autoantigorputzak (TGA) moduko markatzaileak antzeman daitezke.

Hori dela eta, eritasunarekiko suszeptibilitatearen atzean dauden geneen, bidezidorren eta mekanismo erregulatzaileen identifikazioa ezinbestekoa da. Modu horretan, EZren garapenean eragiten duten mekanismo patogenikoak hobeto ulertuko dira, eta gaixotasuna pairatzeko arriskua duten indibiduoak identifikatzeko markatzaile genetiko eta epigenetikoak lortuko dira. Gainera, bidezidor konplexuen eta gene taldeen erregulatzaile nagusien aurkikuntzak interbentziorako iturri berrien identifikazioan lagunduko du. Ziurrenik, glutenik gabeko dieta ordezkatzeko duen tratamendurik sortuko ez den arren, tesi honetan egindako moduko lanek dietaren transgresio puntualek eragindako sintomak arintzeko alternatiba terapeutiko berrietarako ateak zabalik uzten dituzte. Horrez gain, aurkikuntza berriak beste gaixotasun konplexu eta autoimmune batzuetara estrapola daitezke, eta gaixotasun genetikoak ulertzeko beste ikuspuntu batzuk aldarrikatu.

EZren genetika eta epigenetika aztertzeke helburuz, lan hau erabakigarri izan daitezkeen faktore funtzionalen bilaketan fokatu da, hurbilketa bioinformatiko berriak eta hautagaien balioztatze funtzionalak erabiliz. Zehazki, doktorego tesi honetan gene erregulazioaren lau mekanismotan sakondu dugu: transkripzio

faktoreak (TF), mikroRNAk (miRNA), kromatinaren antolaketa eta DNAREN metilazioa.

Sarrera

1. Eritasun zeliakoa

Eritasun zeliakoa (EZ; MIM 212750) edo glutenarekiko sentikorra den enteropatia, genetikoki suszeptibleak diren indibiduoetan, dietan hartutako gariaren glutenarekiko zein zekale eta garagarrean agertzen diren antzeko proteinekiko intolerantzia dela medio garatzen den gaixotasun kroniko eta inflamatorioa da. Ezaugarri nabarmenena heste meharreko biloxken zapaltzea da.

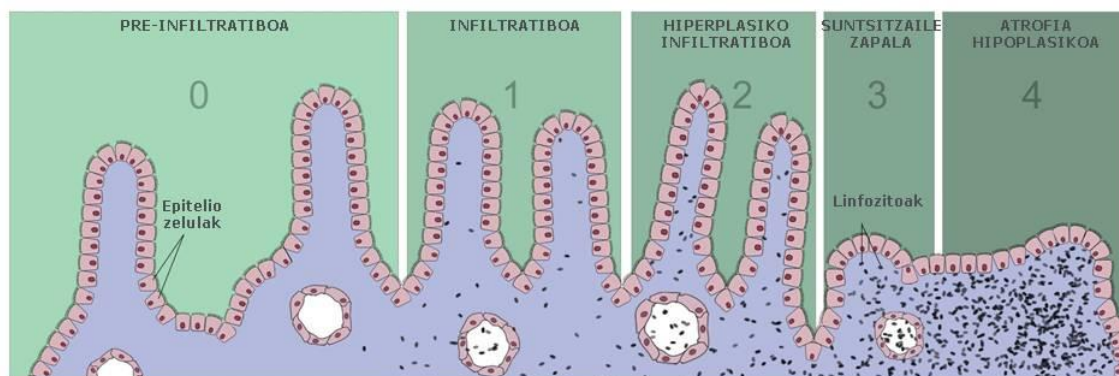
1.1. Ezaugarri klinikoak eta diagnosis

Dietan hartutako glutenaren eragin kaltegarriak ez ziren 1950era arte antzeman (Dicke, 1950), EZren agerpen kliniko lehengo irudi orokorrak 60 urte lehenago argia ikusi zuen arren (Gee, 1888). Tratatu ezean, honako sintomak aurkeztu ohi ditu EZk: batetik, heste meharrari dagozkion ezaugarriak, hala nola, beherakoa, distentsio abdominal edota goragalea, eta bestetik, malabsortzioak eragindakoak, garapenari dagozkionak (pisu baxua, gantzik eza, ilearen fintasuna) edo urritasun psikomotorrari dagozkionak (giharren ahultzea), besteak beste (Feighery, 1999). EZri bestelako zenbait sintoma atipiko ere egotzi izan zaizkio. Horien artean, asaldura neurologikoak, hortzetako esmaltearen akatsak, antzutasuna, osteoporosia, giltzadura mailako sintomak eta gibelesko entzimen kontzentrazio altuak dira nabarmenenak (Mäki eta Collin, 1997).

Ikuspuntu histologikotik, suszeptibilitatea duen pertsonak glutendun dieta hartzen duenean, heste meharraren mailaz mailako endekatzea gertatzen da. Prozesu horren emaitza biloxken atrofia eta kripten hiperplasia ezaugarri dituen hestearen lesioa izaten da. Glutenak eragindako hesteko mukosaren kaltearen maila eta larritasuna Marsh-Oberhuber sailkapenean dago definituta (Marsh, 1992); horren arabera, heste osasuntsu eta normala Marsh 0 mailan dago eta mukosaren atrofia hipoplasikoa Marsh 4ean (1. irudia). Sailkapen honek aldaketa batzuk jasan dituen arren, EZren hasierako estadioetan pazienteak identifikatzeko

eta tratatzeko erabiltzen da oraindik ere. Gaixotasunaren hasieran, epitelio barneko linfuzitoen infiltrazioa ikus daiteke (Marsh 1). Ondoren, kripten hiperplasia gertatzen da (Marsh 2), biloxken tamaina mantendu arren. Hurrengo estadioan (Marsh 3) kripten hipertrofiaz gainera, lamina propioa hanpatuta dago eta biloxken atrofia partziala, ia osoa edo osoa gertatzen da. Estadiotik aurreratuenean (Marsh 4), heste meharraren arkitekturan hipoplasia ikusten da. Hesteko gertaerez gain, EZn glutenaren menpeko diren hainbat autoantigorputz ere agertzen da, anti-endomisio (ingelesez *anti-endomysium autoantibodies* edo EMA) edo anti-transglutaminasa (ingelesez *anti-transglutaminase autoantibodies* edo TGA) kasu (Stenman et al., 2008).

GOI JEJUNOAREN MUKOSAREN IMMUNOPATOLOGIA



1. irudia. Glutenak induzitutako aldaketak estadio ezberdinetan, Marsh sailkapenaren arabera. www.thegluten.com orrialdetik moldatua.

Ezaugarri patologiko horiek kontuan hartuta, Gastroenterologia, Hepatologia eta Elikadura Pediatrikoaren Europako Elkarteak (ingelesez *European Society for Pediatric Gastroenterology, Hepatology and Nutrition* edo ESPGHAN) irizpide diagnostikoak finkatu ditu. Aurretik, irizpide horiek heste meharreko biopsietako kalte histologiko eta emaitza serologikoetan oinarritu dira, azken horiek ezinbestekoak izan ez badira ere. Gainera, jarraipen biopsiak ere egin izan dira diagnosiaren osteko jarraipenean zehar kontrol modura (Mäki, 1995). 2012an onartutako irizpideen arabera, antigorputzek eta ezaugarri klinikoek lehentasuna dute eta duodenoaren biopsia egiteari utz dakioke baldin eta haurra sintomatikoa bada, bere TGA mailak normaltasun atalasea baino 10 aldiz altuagoak badira eta

EMA positiboa bada lagin independente batean. Giza Antigeno Leukozitarioaren (ingelesez *Human Leucocyte Antigen* edo HLA) genotipoa aztertzea ere oso baliagarria izan daiteke, EZ oso ezohikoa delako arrisku alelo genetikorik gabe (Husby et al., 2012).

1.2. Epidemiologia

Aurreko mende bukaerara arte, EZ ezohiko gaixotasuna zela uste zuen komunitate zientifikoak, ia modu eksklusiboan europar jatorriko pertsonen 1/1000 maiztasunarekin pairatzen zutena (Feighery, 1999). Hala ere, osteko populazio ikerketek Europa Mendebaldeko maiztasuna %1 ingurukoa dela erakutsi dute, eskualdeen artean maiztasunean ezberdintasunak dauden arren (Dubé et al., 2005).

Ikuspuntu klasikoaren arabera, EZ Europar jatorriko pertsonetan garatzen den gaixotasuntzat jo izan den arren, zoriz aukeratutako Ipar Amerika eta Australiako populazioetan 1:100 eta 1:200 bitarteko maiztasunak ikusi dira (Cataldo eta Montalto, 2007; de Kauwe et al., 2009). Bestetik, EZ Latinoamerikan (Galvao et al., 1992; Rabassa et al., 1981; Sagaro eta Jimenez, 1981), Afrikako iparraldean (Al-Tawaty eta Elbargathy, 1998; Suliman, 1978), eta Ekialde Hurbilean (Al-hassany, 1975; Khuffash et al., 1987) ere arraroa eta ezohikoa dela pentsatu izan da, leku hauetan kasu mugatu eta puntualak besterik ez baitira behatu. Gainera, Ekialde Urrunean (Txina, Japonia, Korea...) existitzen ez zela pentsatu izan da (Fasano eta Catassi, 2001). Dena den, bahetze edo *screening* berriek munduko leku hauetan guztietan EZren maiztasuna gutxietsita zegoela frogatu dute, maiztasuna mendebaldeko herrialdeetan erregistratutakoaren antzekoa dela erakutsita. Gluten asko duen mendebaldeko dieta modernoa munduko leku guztietara hedatu da, eta ondorioz, EZ Osasun Publikoko arazo bihurtu da, baita garapen bidean dauden herrialdeetan ere (de Kauwe et al., 2009).

Zenbait ikerlanek iradoki duenez, EZren haurtzaroko agerpenaren maiztasuna haziz joan da 1980 eta 1990eko hamarkadetan zehar, eta hazkunde hori umeen elikadura irizpideekin erabat lotuta dagoela ikusi da (Ivarsson et al., 2000). Eskuragarri dauden datu berriek EZren intzidentzia handitzen ari dela berretsi dute, eta EZ aurretik uste zena baino ohikoagoa dela ikusi da (Catassi et al., 2014). Bestetik, helduen diagnosi kopuruak ere gora egin du informazioa eskuragarri dagoen lurraldeetan (Bodé eta Gudmand-Høyer, 1996; Collin et al., 1997; Murray et al., 2003). Genetikoak ez diren arrisku faktoreen artean, urtaroen araberrako patroiak eta zenbait infekzio biriko aipatu dira (Plot eta Amital, 2009), baita glutenaren kontsumoan eman diren aldaketak eta haurren elikadura ere (Catassi et al., 2014).

EZren prebentzio primarioaren aukera ikertzeko Prevent CD (www.preventcd.com) Europako ikerketa proiektu multizentrikoa abiatu zen (Hogen Esch et al., 2010). Lan horren funtsa honakoa zen: genetikoki suszeptibleak ziren haurren dietan (ahal izanez gero oraindik bularra ematen zieten bitartean) gluten kantitate txikiak sartzea glutenarekiko tolerantzia indultzeko. Emaitzen arabera, ez gluten kantitate txikien bidezko indukzioak, ez eta edoskitzeak ere, ez dute EZren arriskua murrizten (Vriezinga et al., 2014).

Genero ikuspuntutik, emakumeek gizonek baino maizago pairatzen dute EZ. Garai emankorrean dauden gaixoen artean proportzioa 3:1 arrazoiria iristen da (Feighery et al., 1998).

1.3. Tratamendua

Egun, tratamendu eraginkor bakarra biziarteko glutenik gabeko dieta (GGD) hertsia da, gari, zekale eta garagarraren produktuak dietatik ezabatuz (Feighery, 1999; Di Sabatino eta Corazza, 2009). Sintomen hobekuntza edo desagertzea egun edo aste batzuen buruan gertatzen da, askotan markatzaile serologikoen eta duodenoaren atrofiaren normalizazioarekin batera (Murray et al., 2004).

Tratamendua modu zorrotzean jarraitzen ez bada konplikazioak ager daitezke, hala nola, garaiera baxua, desnutrizioa, osteoporosia, gaixotasun autoimmunitario sekundarioak, minbiziak eta arazoak haurdunaldian (Lerner, 2010).

Tratamendua eraginkorra bada ere, GGD zorrotza betetzea ez da erraza, eta EZ pairatzen duten pazienteen bizi kalitatea hobetzeko konponbideak beharrezkoak dira (Samasca et al., 2014). Gainera, gaixoen proportzio txiki batek ez dio tratamenduari ondo erantzuten eta EZ errefraktarioa (ingelesez *Refractory CD* edo RCD) garatzen du; hau da, sintoma iraunkor edo errepikakorrek agertzen dituzte, GGD hertsian egon arren, heste hantura eta biloxken atrofia kronikoa kasu (Rubio-Tapia eta Murray, 2010). Horregatik guztiagatik, terapia alternatibo seguru, eraginkor eta eskuragarrien garapena ezinbestekoa da.

EZren patologiar buruzko ezagutza handitu ahala, GGDren ordezkotako edo osagarri izan daitezkeen terapia berriak garatzen joan dira (Freeman, 2013). Hainbat droga garapen eta azterketa fase ezberdinetan dago, larazotido azetatoa esaterako, zelulen geruza bakarreko iragazkortasun parazelularra blokeatu eta epitelioan zeharreko gliadina peptidoen fluxua ekiditen duen peptido sintetikoa (Paterson et al., 2007). Plazeboaren kontrako ikerlan batean, 0,5 mg larazotido azetato egunean hiru aldiz hartzeak GGDn zeuden gaixo sintomatikoaren egoera arintzen zuela ikusi zen. Lotura hertsien erregulazioa modulatzeko agente hori aukera terapeutikoa izan liteke GGDn egon arren, EZren sintomak mantentzen dituzten pazienteentzat; hala ere, azterketa gehiago egitea beharrezkoa da (Leffler et al., 2015).

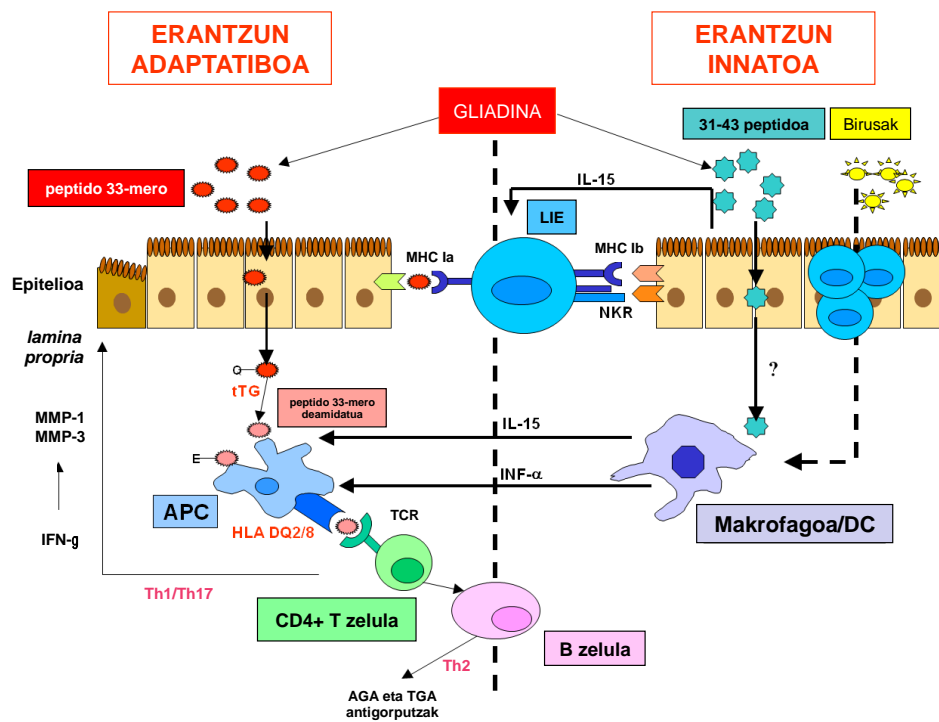
Berriki, Nexvax2 txerto terapeutikoa frogatu da. Txerto honek gluten-espezifikoko diren CD4+ T zelulenzako epitopoak ditu, modu horretan glutena dieta bidez sartzen denean prozesu immunea inaktibatzeke helburuarekin (Goel et al., 2017). Momentuz emaitza onak jaso dira lehenengo faseko frogetan, baina oraindik ikerketa gehiago behar dira EZ sendatu dezakeen demostratzeko.

2. Eritasun zeliakoaren patogenesisia

EZren garapenean parte hartzen duten mekanismoen ezagutzan egin diren aurrerapenen ondorioz, egun EZ HLArekin erlazionaturiko gaixotasunen artean hobekien ezagutzen dugunetako bat da. Hala ere, zenbait prozesu patogeniko azaltzeke dago oraindik.

T zelulek EZn parte hartzen dutela aspalditik da ezaguna. Laburbilduz, gliadinaren peptidoek epitelioa zeharkatzen dute, eta lamina propiora iristean, ehunetako transglutaminasak (TG2) deamidatu egiten ditu. HLA-DQ2 edota HLA-DQ8 hartzailekun zelula antígeno aurkezleek (ingelesez *Antigen Presenting Cell* edo APC), deamidatutako peptidoak CD4+ T zelula patogenikoei aurkezten dizkiete eta horren ondorioz, lamina propioaren infiltrazioa, kripton hiperplasia eta biloxken atrofia eragingo dituzten Th1 eta Th17 erantzunak abiaraziko dira (Castellanos-Rubio et al., 2009; Schuppan et al., 2009).

Hainbat azterlanek erakutsi duenez, berezko immunitate-sistemak ere berebiziko garrantzia du EZn, eta badirudi gliadina T zelulen menpekoa ez den erantzuna aktibatzeke gai dela (Hüe et al., 2004; Maiuri et al., 2003) (2. irudia). Gliadina peptidoen aurkako berezko erantzun immunearen bitartekari nagusia interleukina 15 (IL-15) zitokina da, eta horrek epitelio barneko linfzitoen (ingelesez *intraepithelial lymphocytes* edo IEL) funtzioa molda dezake (Escudero-Hernández et al., 2017).



2. irudia. Eritasun zeliakoaren mekanismo patogenikoak: adar immunitario egokitzaila eta berezkoa. APC, zelula antigeno aurkezlea; MMP, matritzeko metaloproteinasak; IFN, interferoia; Th, T zelula laguntzailea; HLA-DQ2/8, DQ2/8 giza antigeno leukozitarioa; AGA, anti-gliadina antigorputzak; TGA, anti-transglutaminasa antigorputzak; IEL, epitelio barneko linfotzitoa; IL-15, interleukina 15; NKR, Natural killer (NK) hartzailea; DC, zelula dendritikoa.

Dena den, patogenesiaren alderdi asko argitzeke dago oraindik, bereziki glutenaren eta zelula epitelialen arteko interakzioak, lamina propiorako gluten peptidoen fluxua, TG2aren aktibazioa, IL-15ren adierazpena kontrolatzen duten mekanismoak eta autoantigorputzen ekoizpena.

2.1. Glutena

Suszeptibilitate genetikoak duten indibiduoek pairatzen dute EZ, dietako glutenaren proteinekiko kontaktuaren ondorioz. Glutena alkoholetan disolbagarriak diren glutenina monomeroen eta prolaminatan aberatsak diren gliadina polimerikoen nahastura da, eta gari, zekale eta garagarrean dago.

Hesteko proteasek ezin dituzte proteina hauek guztiz endekatu eta zati luzeak (10-50 aminoazido) geratu ohi dira hesteko lumenean (Shan et al., 2002). Proteina zati horiek transglutaminasaren substratu egokiak dira, eta deamidatu egiten ditu HLA-DQ2 eta HLA-DQ8 molekulekiko afinitatea handitu eta ondorioz, CD4+ T zelulek gidatzen duten Th1 erantzuna sustatzen da. Erantzun honek hesteko mukosaren hantura, malabsortzioa eta EZren beste sintoma batzuk dakartza. Gainera, glutenak CD8+ T zelulen erantzuna ekar dezake lamina propioan eta Histokonpatibilitate Konplexu Nagusiarekiko (ingelesez *Major Histocompatibility Complex* edo MHC) independentea den mekanismo baten bidez epitelio barneko linfuzitoen populazioa emendatu dezake (Jabri eta Sollid, 2009; Schuppan et al., 2009).

In vitro eginiko zenbait ikerketan glutenak berezko immunitate-sistema kitzika dezakeela ikusi da. Gliadinatik eratorritako peptidoen artean, α -gliadina molekularen 31-43 arteko aminoazidoak hartzen dituen peptidoa izan da ikertuena. Peptidoak ez du T zelulen erantzunik eragiten, bai ordea berezkoa eta ondorioz, epitelioaren apoptosia sustatzen du EZ duten pazienteetan (Maiuri et al., 2003). 2013an burututako ikerketa lan batean 8-mer gliadina peptidoa identifikatu zen. Peptidoaren forma deamidatuak erantzun immune humoralak sortarazten du, IgA antigorputz espezifikoak, hain zuzen ere (Vallejo-Diez et al., 2013).

2.2. Transglutaminasa

TG2 nonahi adierazten den proteina multifuntzionala da, eta orokorrean espazio estrazelularrean da aktibo. Proteina baten glutaminaren eta bigarren proteina baten lisinaren arteko lotura itzulezin eta kobalentea katalizatzen du (Folk eta Cole, 1966; Folk eta Chung, 1985). Immunitate-sistema egokitzailearen aktibazioan berebiziko garrantzia du; izan ere, gliadina peptidoen eraldaketa entzimatikoa eragiten du, eta HLA-DQ2 eta HLA-DQ8 molekulekiko afinitatea emendatzen du (Arentz-Hansen et al., 2000). Glutena prolina eta glutaminetan

aberatsa da eta gainera, HLA-DQ2 eta HLA-DQ8 molekulekin lotzeko beharrezkoak diren aminoazido negatibo gutxi ditu. Gliadinaren peptido espezifikoaren TG2ren bidezko deamidazioak karga negatiboak emendatzen ditu, eta horrela peptidoen HLA-DQ2 eta HLA-DQ8 molekulekiko lotura eta CD4+ T zelulei aurkezpena faboratzen dira (Molberg et al., 1998; Ráki et al., 2007; Wal et al., 1998). Beraz, TG2k gliadina eraldatzen du, molekula ez-kitzikatzaile izatetik, zitokinen tokiko ekoizpen masiboa eragiteko eta enterozitoen desberdintzapena eta proliferazioa abiarazteko gai den T zelulen antigeno izatera.

Bestetik, TG2k katalizatutako erreakzioan parte hartzen duten peptidoek eta entzimak berak, TG2-gliadina konplexuak eratzen dituzte, TGAen ekoizpena bultzatuz (Caputo et al., 2009). Zenbait ikerketek TGAek patogenesisian eta hesteko mukosaren ohiko lesioan izan dezaketean papera azpimarratu dute; izan ere, ziklo zelularrean, apoptosian, angiogenesisian eta heste iragazkortasunean eragiten dutela ikusi da, eta eritasunaren patologian paper garrantzitsua izan dezaketela iradoki da (Caputo et al., 2009; Lindfors et al., 2009). Zenbait ikerketek TGAek EZn rol patogenikoa izan dezaketela proposatu dute; izan ere, TG2ren aktibitatea modulatu dezakete. Autoantigorputzek α -gliadinaren 31-43 peptidoaren (berezko immunitatearen eragilea) hartzea ere aldatzen dutela proposatu da, horrela heste zelulak p31-43k sortutako kaltetatik babestuz. Duela gutxi gauzatutako ikerlan batean, TGAek p31-43k sortzen dituen kalteen aurrean zelulak babesten dituen ikertu da, horretarako fibroblastoak erabiliz eta EZ eredu bat sortuz. Antigorputzek, indibiduo osasuntsuetatik eratorritako fibroblastoetan, p31-43ren hartzea txikitzen zutela ikusi zuten, ez ordea EZ zuten pazienteetatik eratorritakoetan, eta gaixotasunean antigorputzen rol babeslearen galera ematen dela iradoki zuten (Paolella et al., 2017).

2.3. Immunitate-sistema egokitzailea

Immunitate-sistema egokitzaileak T zelulen bidezko immunitatea eta immunitate humorala induitzen ditu, eta biak EZdun pazienteen heste meharreko mukosan aktibatzen dira gliadinaren estimuluz. Hesteko CD4⁺ T linfozitoek APCetako HLA-DQ2 eta HLA-DQ8 heterodimeroei lotuta aurkezten diren deamidatutako gliadina peptidoak ezagutzen dituzte (Lundin et al., 1990; Mazzarella et al., 2003). Gliadinarekiko espezifiko diren mukosa zeliakoko T zelulak Th1 fenotipokoak dira batez ere eta hantura sustatzen duten zitokinak ekoizten dituzte, IFN- γ kasu (Nilsen et al., 1995; Troncone et al., 1998). Zitokina horretaz gain, Th1 sustatzen duten beste zitokina batzuk ere gainadierazita daude, IL-18 eta IFN- α esaterako (León et al., 2006; Monteleone et al., 2001; Steinman, 2007).

Bestetik, IL-6 eta TGF β ren presentzian desberdintzatu eta IL-17 zitokina ekoiz dezakeen bestelako CD4⁺ T zelula laguntzaileen lerroa identifikatu da: Th17 linfozitoak. Ikusi denez, lerro hau lehen IL-12/INF- γ sareari egozten zitzaizkion zenbait funtzioen erantzulea da (Castellanos-Rubio et al., 2009). Th1 eta Th17 erantzunek gaixotasun aktiboan parte hartzen dute, beste zenbait gaixotasun autoimmunitariotan ikusi den bezala, hain zuzen (Harris et al., 2010; Monteleone et al., 2010; Sjöström et al., 1998).

2.4. Berezko immunitate-sistema

Berezko erantzun immunitarioa patogenoen aurkako lehen harresia da eta agente infekziosoekin topo egin osteko lehenengo estadioetan aktibatzen da. EZn berezko immunitate-sistemak T $\alpha\beta$ linfozitoekiko modu independentean aurre egiten dio gliadinaren estimulari, eta gerora sortuko den T zelulen aktibaziorako inguru proinflamatorio egokia sortzen laguntzen du HLA-DQ2 edo HLA-DQ8 daramaten pazienteetan.

Zenbait *in vivo* azterketak erakutsi duenez, gliadinaren 31-43 peptidoak gaixotasunaren sintomak agerraraz ditzake, eta biopsia kulturetan hainbat aldaketa behatu da (Maiuri et al., 1996; Picarelli et al., 1999; Sturgess et al., 1994). Peptido honek ez du T zelulen bidezko erantzunaren estimulazioa eragiten (Anderson et al., 2000; Arentz-Hansen et al., 2000), eta beraz, pentsatzekoa da 31-43 peptidoaren toxizitatea berezko immunitatea aktibatzeke gaitasunari dagokiola. Berriki, peptido honek EZdun pazienteen hestean eta Caco-2 lerro zelularrean IFN- α aktibatzeke gaitasuna duela ikusi da, Toll-like hartzaile 7 (ingelesez *Toll-like receptor 7*) edo TLR7ren ligando birala den loxoribinaren (LOX) kooperazioz, TLR7 bidezidorra aktibatu eta garraio endozitikoan eragiten duelako. Honek, infekzio biralek eta elikadurako proteinek berezko immunitate sistema potentziatu dezaketela iradokitzen du (Nanayakkara et al., 2018).

Monozitoetan, makrofagoetan eta zelula dendritikoetan egindako azterketek TLR4 eta bere seinale transduktore nagusia den MyD88 inplikatu dituzte zerealen proteinekiko garatzen den erantzunaren hartzaile eta eragile primario modura (Schuppan et al., 2009). Glutenak eragindako IELen berezko aktibazioak *MICA* genearen adierazpena induzitzen du hesteko epitelioan, hain zuzen ere, NK (ingelesez *Natural Killer*) eta T $\gamma\delta$ zeluletan eta CD4⁺ nahiz CD8⁺ diren zenbait T zelulen azpitaldetan adierazten den NKG2D hartzailearen ligandoarena. MICA epitelialak, gainadierazitako IL-15ekin batera, NKG2Dren aktibazioa sustatzen du IELetan, linfozitoak bitartekari modura dituen zitotoxizitate antigeno-espezifikoa bultzatuz. Azkenik, IL-15ekin batera diharduen IL-21 ere berezko immunitatearen gidaritzan aipagarria dela ikusi da (Fina et al., 2008; He et al., 2004; Sarra et al., 2013).

2.5. Beste bidezidor biologikoak

EZn identifikatutako suszeptibilitate gene gehienak erantzun immunitarioan parte hartzen dutela ikusi da, baina heste ehunaren suntsipenean eta elikagaien malabsortzioan parte hartzen duten mekanismo sekundarioen funtzionamendua

ez da guztiz ezaguna. Nahiz eta gaixotasunaren eragile nagusia transglutaminasak deamidatutako gliadina peptidoen aurkako erantzun autoimmunea izan, badira EZren ezaugarri anatomiko eta histologikoetan eragiten duten beste bidezidor biologiko eraldatu batzuk ere. Mekanismo patogeniko hauek oraindik ere ez dira guztiz argitu.

Azken urteetan, genoma osoaren hainbat adierazpen analisi egin dira ehun ezberdinetan, *microarray*-etatik hasi eta RNAREN sekuntziazioraino (RNA-seq). Horietariko batzuk, EZrekin erlazionatutako aldaera genetikoek gene isolatuen, gene taldeen, erregulazio koordinatua duten geneen eta bidezidorren adierazpenean duren eragina argitzeko helburuarekin egin dira (Castellanos-Rubio eta Bilbao, 2017).

EZn eginiko lehenengo *microarray* esperimentuan EZdun paziente aktiboen, GGDn zeuden EZdun pazienteen eta EZ pairatzen ez zuten kontrolen biopsiak erabili ziren. Lan horretan T zelulen aktibazioarekin, B zelulen heltze prozesuarekin eta epitelioko zelulen diferentziazioarekin erlazionatutako geneak identifikatu ziren (Juuti-Uusitalo et al., 2004). Gaixotasun aktiboan eta GGDn zeuden pazienteetan, adierazpena aldatuta zuten geneak (konstitutiboki aldatutako geneak) EZren etiologian potentzialki garrantzitsuak izan daitezkeela iradoki zuten. Berrikiago ikusi da konstitutiboki aldatutako geneak normalean bidezidor biologikoen zentroan kokatzen direla, gaixotasun aktiboan soilik aldatuta agertzen direnak periferikoagoak diren bitartean (Fernandez-Jimenez et al., 2014).

Bigarren *microarray* batean (Diosdado et al., 2004), adierazpen genikoa EZ pairatzen zuten 15 pazienteen (aktiboak eta GGDn zeudenak), eta EZ ez zuten 7 kontrolen duodenoko biopsietan aztertu zen. Alde batetik, hesteko kalteak aztertu zituzten gluten-ingestioa kontuan hartu gabe, eta T zelula eta makrofagoekin erlazionatutako geneen emendioa gertatzen zela ikusi zuten, baita Th1 erantzunarekin eta zelulen proliferazioarekin erlazionatutakoena ere.

Bestetik, glutenaren efektuak aztertu zituzten, eta ziklo zelularrekin eta zatiketa zelularrekin erlazionatutako geneak identifikatu zituzten.

Hirugarren *microarray* batean EZ pairatzen zuten pazienteen heste biopsiak erabili ziren (Castellanos-Rubio et al., 2010). Esperimentu honetan, duodenoko mukosan gliadinaren kitzikapenaren ondorioz gertatzen diren erantzun akutua eta epe-luzeko erantzuna aztertu ziren. Ziklo zelularrean, apoptosian, matrize estrazelularreko eta komunikazio zelularreko bidezidorretan aldaketak aurkitu ziren glutenaren aurreko bi erantzun motetan. TGF- β , Jak-Stat eta NF κ B-rekin erlazionatutako seinalizazio bidezidor konplexuen kasuan, batez ere epe-luzeko erantzunean aurkitu ziren aldaketak. Beraz, haien partaidetza EZren jatorriari baino, gaixotasunaren iraunarazteari dagokiola dirudi.

EZdun pazienteetan proliferazio eta diferentziazioan parte hartzen duten gene berriak aurkitzeko, diferentziazio eredu diren TGF- β 1ekin induzitutako T84 zelula epitelialetan *microarray* analisiak ere egin ziren (Juuti-Uusitalo et al., 2007). Lan horretan hazkunde epidermikoko hazkuntza faktorearen (ingelesez *epidermal growth factor* edo EGF) bitarterko seinalizazio bidezidorreko partaide diren zenbait gene identifikatu zituzten. Gliadinarekin kitzikatutako Caco-2 zeluletan beste *microarray* esperimentu bat egin zen itu geneak identifikatzeko (Parmar et al., 2013). Zenbait gene identifikatu bazen ere, ezin izan zuten gliadinaren efektuak zirenik ziurtatu. Hala ere, *in vitro* ereduakin lortutako emaitzak kontu handiz hartu behar dira. EZdun gaixoen eta kontrolen heste biopsietatik isolatutako epitelio zeluletan egin zen beste ikerlan batean, proliferazioan, heriotza zelularrean, eta diferentziazio bidezidorretan adierazpen aldaketak ikusi ziren, besteak beste (Bracken et al., 2008).

RNAREN sekuentziazioa *microarray*-ekiko alternatiba indartsu bat da, hondoko seinale baxuagoak eta sentikortasun altuagoa eskaintzen dituena. Gainera, RNAREN sekuentziazioak gaixotasunaren patogenesisian garrantzitsuak izan daitezkeen transkribatutako eskualde berriak identifikatzeko aukera ematen du

(Wang et al., 2009). Egun, EZn RNA sekuentziazio lan bakarra publikatu da (Quinn et al., 2015). Lan honetan EZdun paziente eta kontrolen odol periferikoko zelula mononuklearretatik lortutako CD4+ T linfozitoen transkriptoma aztertu zuten oinarrizko baldintzetan eta anti-CD3/CD28 antigorputzez eta forbol 12-miristato 13-azetatoz (ingelesez *Phorbol 12-Myristate 13-Acetate* edo PMA) kitzikatu ondoren. Autoreek ez zuten kitzikatu gabeko zeluletan aldatutako bidezidorrik aurkitu, antigorputzekin kitzikatutakoan zitokina-zitokina hartzaile interakzio bidezidorra identifikatu zuten bitartean. PMArekin estimulatutako laginetan immunitatearekin erlazionatutako geneak aurkitu ziren. Antigorputz-estimulazioan aldatuta zeuden geneak koadierazpen moduluetan taldekatu ziren. Talde hauetan immunitatearekin erlazionatutako bidezidorrak eta aurretik EZren arriskuarekin asoziatutako geneak identifikatu ziren. Zehazki, *BACH2* gene hautagaia (Dubois et al., 2010) azpiadierazita zegoen frogatutako kondizio guztietan, honek T zelulen diferentziazioan eta gaixotasun autoimmuneen prebentzioan duen garrantzia azpimarratuz.

Laburbilduz, esperimentu ezberdinetan jasotako emaitzak elkartzea zaila bada ere, erantzun immunearen areagotzea eta seinalizazio eta ziklo zelularreko aldaketak dira nabarmen. Ikerketa hauek aldatutako gene eta bidezidor ugari identifikatu dute eta eragile ez genetikoaren, polimorfismo genetikoaren, eta adierazpen genetikoaren arteko elkarrekintza konplexuak ulertzen hasiak gara. Gertaera horietako partaide genomiko nagusien identifikazioak eritasun autoimmune honen eta beste batzuen itxaropen terapeutiko potentzialen identifikazioan lagun lezake.

3. Eritasun zeliakoaren Genetika

EZren herentzia eredu oraindik ezezaguna bada ere, Genetikak gaixotasun honekiko suszeptibilitatean parte hartzen duela aspalditik da jakina. Familia afektatuetan gauzaturiko prebalentzia ikerketak eta bereziki, bikiak konparatu

dituztenak, berebizikoak izan dira ingurumen faktoreen eta faktore genetikoen ekarpenak zenbatekoak diren estimatzeko. Azterlan horien arabera, Genetika funtsezkoa da EZren hasieran zein garapenaren fase guztietan.

Orokorrean, komunitate zientifikoak EZren agerpenean bat datozen biki monozigotikoen proportzioa %75-86 dela onartzen du. Biki dizigotikoen kasuan, aldiz, kopuru hau nabarmen jaisten da, %16-20ra arte, hain zuzen. Biki mota bien konkordantzia desberdintasunek EZren osagarri genetikoa estimatzen lagundu dute eta orain arte bestelako gaixotasun autoimmuneetan behatutakoa baino pisuzkoagoa dela argi utzi dute. Adibidez, 1 motako diabetea (T1D) osagarri genetikoa baxuagoa da (konkordantzia biki monozigotiko eta dizigotikoetan %30 eta %6, hurrenez hurren) (Sollid eta Thorsby, 1993). Gainera, anai-arreben eta biki dizigotikoen arteko konkordantzia ratioak oso antzekoak dira EZn eta beraz, ingurumen faktoreek gaixotasuna garatzeko arriskuan duten ekarpena minimoa dela ondoriozta daiteke. Laburbilduz, azken urteetan jasotako ebidentziek EZk osagai genetikoa oso garrantzitsua duela eta gaixotasunaren heredagarritasuna (faktore genetikoei egotzi ahal zaien EZ pairatzeko arriskuaren proportzioa, ingurumen faktoreei egotzi ahal zaienarekin alderatuta) %87 ingurukoa dela eraman dute aditzera (Dieli-Crimi et al., 2015; Greco et al., 2002; Nistico et al., 2006).

Gaixotasuna garatzeko arrisku genetikoen parte handiena HLA-DQ2 edota HLA-DQ8 aleloen presentziari dagokio. Dena den, eta gaixotasunaren patogenesisian HLA molekula horien papera oinarrizkoa bada ere, herentziari egiten dioten ekarpena ez da oso handia. HLA lokus klasikoek EZren arrisku genetikoen %23 azaltzen dutela kalkulatu da. Berriki %18ko ekarpena egiten duten 5 HLA aldaera aurkitu dira (Gutierrez-Achury et al., 2015) eta beraz, guztira EZrekiko arriskuaren %40a azaldu daiteke. Ondorioz, HLA lokusetik kanpo eragin txikiko lokus ezezagun ugari egon behar dutela ondorioztatu da.

Azken urteetan EZn identifikatzeke dauden osagai genetikoak aurkitzeko lan handia egin da. Saiakera hauen artean lotura genetikoko ikerketak, gene kandidatuen asoziazio ikerketak, eta azken hamarkadan genoma osoko asoziazio azterketak (ingelesez *Genome Wide Association Studies* edo GWAS) egin dira. Estrategia hauek erabilia genomaz zeharreko zenbait *loci* asoziatu dira EZrekin; batzuk ziurtasunez eragiten dute gaixotasunean, beste batzuk aldiz gehiago ikertu behar dira.

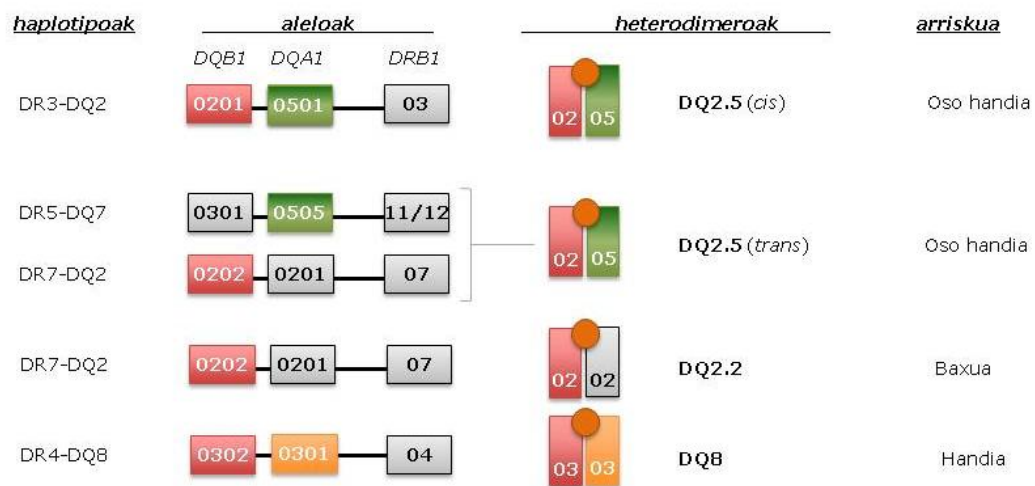
3.1. HLA eremuaren partaidetza

HLA da gizakian MHCk hartzen duen izena; 6p21 eremu kromosomikoan kokatzen den superlokusa da, eta immunitate-sistemarekin zerikusia duten gene ugari dauka. Eskualde honetan, lotura desoreka (ingelesez *Linkage Disequilibrium* edo LD) handia da eta kontserbatutako haplotipo blokeak daude. HLA geneek giza zelula gehienetan adierazten diren proteina antigeno-aurkezleak kodetzen dituzte. Hauek ezinbestekoak dira molekula exogeno eta endogenoen artean bereizteko.

HLA geneek gaixotasun autoimmunitario eta inflamatorio askotan hartzen dute parte, eta malaria edo hiesa bezalako eritasun infekziosoak pairatzeko arriskua ere emanda dezakete. Hala ere, eremuko konplexutasun genetiko handia dela eta, gaixotasun hauen suszeptibilitatea azaltzen dituzten faktore genetiko zehatzak eta mekanismo patogeniko gehienak ezezagunak izaten jarraitzen dute gaur egun. Izan ere, HLAk genoma osoko gene dentsitatearik altuena eta adierazpen maila goratua ditu (Horton et al., 2004).

Lehen esan bezala, HLA suszeptibilitate eremurik garrantzitsuena da EZn, eta gaixotasunaren osagarri genetikoaren %40 inguru azaltzen du (%23 aldaera klasikoek eta %18 berriek). HLA eta EZren arteko lehenengo asoziazio ebidentziek 1973an ikusi zuten argia (Ludwig et al., 1973). Beranduago

publikatutako ikerlanek HLA-DQ2 eta HLA-DQ8 kodetzen dituzten II klaseko HLA geneak zuzenean inplikaturik dauden faktoreak direla erakutsi dute (Sollid, 1989; Spurkland et al., 1992) (3. irudia). HLA-DQ2 eta HLA-DQ8 aldaerak LDn daude HLA-DR3 eta HLA-DR4ekin, hurrenez hurrez, eta hori dela eta, askotan arrisku aleloek DR3-DQ2 eta DR4-DQ8 haplotipo izenak hartzen dituzte (Sollid, 1989).



3. irudia. HLA eta EZren arteko asoziazioa. HLA-DQ2 molekula EZ garatzeko arrisku faktorerik garrantzitsua 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 *trans* eran, DR5-DQ7 eta DR7-DQ2.2 heterozigotoetan. HLA-DQ2.2 dimeroak, HLA-DQA1*02:01 eta HLA-DQB1*02:02 aleloek kodetutako 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., 2011 publikaziotik moldatua.

HLA-DQ heterodimeroen funtzio nagusia T zelulei antigeno exogenoak aurkeztea da. Polimorfismo genetikoek heterodimeroen lotura gunea aldarazten dute, bertara modu eraginkorrean lotu daitezkeen eta ondoren aurkezten diren peptidoen multzoan eraginez. Aurretik esan bezela, EZn HLA-DQ2 eta HLA-DQ8 molekulak dira arrisku genetikoarekin asoziatutako faktore nagusiak, hala ere, HLA-DQ2 da eremuko asoziaziorik indartsuena (Louka eta Sollid, 2003).

HLA-DQ2.5 heterodimeroa, DQA1*05 eta DQB1*02 aleoen konbinazioz osatua, gaixoen %90ean agertzen da. Heterodimeroaren α eta β kateak kromosoma berean egon daitezke kodetuak, eta beraz *cis* eran, baina, heterodimeroko bi kateak *trans* moduan ere kodetuta egon daitezke. HLA-DQ2.5 dimero bi mota hauen arteko desberdintasunek DQ α kateko aminoazido batean (DQA1*05:01 *vs.* DQA1*05:05) eta DQ β katearen mintzaren eremuko beste batean (DQB1*02:01 *vs.* DQB1*02:02) eragiten dute, baina ez du ematen ondorio funtzionalik dutenik, eta arrisku efektu berdinarekin asoziatuta daude. Dena den, HLA-DQ2ren beste aldaera batek, HLA-DQ2.2 dimeroak, ekartzen duen arriskua oso txikia da (Sollid, 2002; Sollid eta Thorsby, 1993).

HLA-DQ2 eramaileak ez diren gaixo gehienak HLA-DQ8 eramaileak dira, eta DQA1*03:01 eta DQB1*03:02 aleloak kodetzen dituen haplotipoaren kopia bat dute gutxienez (Mäki eta Collin, 1997). Pazienteen proportzio txiki bat DQ2 eta DQ8rako negatiboa da, baina kasu gutxi hauetan ere, gaixoak DQ2 molekula kodetzen duten aleloetako baten eramaileak direla ikusi da (DQA1*05 edo DQB1*02) (Karell et al., 2003; Spurkland et al., 1992).

EZrekiko suszeptibilitate mailaren eta HLA-DQ2.5 heterodimero kopuruaren arteko harremana ere ikusi da. Bi DR3-DQ2 haplotipo dituzten indibiduo homozigotoek eta DR3-DQ2/DR7-DQ2 heterozigotoek adierazten dituzte DQ2.5 heterodimero mailarik altuenak, eta beraz, EZ pairatzeko arriskurik handiena dute (van Belzen et al., 2004; Lundin et al., 1993; Ploski et al., 1993). Zentzu horretan, aipagarria da EZ errefraktarioa duten pazienteek (GGDren aurrean erantzuten ez dutenak) maizago agertzen dutela DR3-DQ2 haplotipoa homozigosian (%44-62) beste gaixo zeliakoekin alderatuta (%20-24). DQ8 molekulekin ere dosiaren menpeko antzeko efektua proposatu da.

DQ molekulak kodetzen dituzten geneez aparte, HLA eremuak baditu immunitate-sisteman parte hartu eta EZ pairatzeko arriskua emenda dezaketen

beste gene asko ere. MHC eskualde osoan egindako mapaketa zehatz batean HLA-DQ geneen bi alboetara arrisku aldaera berriak identifikatu ziren, telomero aldeko I klaseko HLA eskualdean eta zentromero aldeko HLA-DPB1 genean. Lehen esan bezala, aldaera berri hauek EZn suszeptibilitate genetikoaren %18a azaltzen dute (Gutierrez-Achury et al., 2015).

HLA geneek suszeptibilitate genetikoari egiten dioten ekarpena berebizikoa bada ere, zeliakoa ez den populazioaren %20-30ak HLA-DQ2 aldaera aurkezten du eta beraz, HLA geneak garrantzitsuak izan arren, gaixotasuna garatzeko nahikoak ez direla esan daiteke (Sollid, 2002).

3.2. HLA eremutik kanpoko arrisku genetikoa

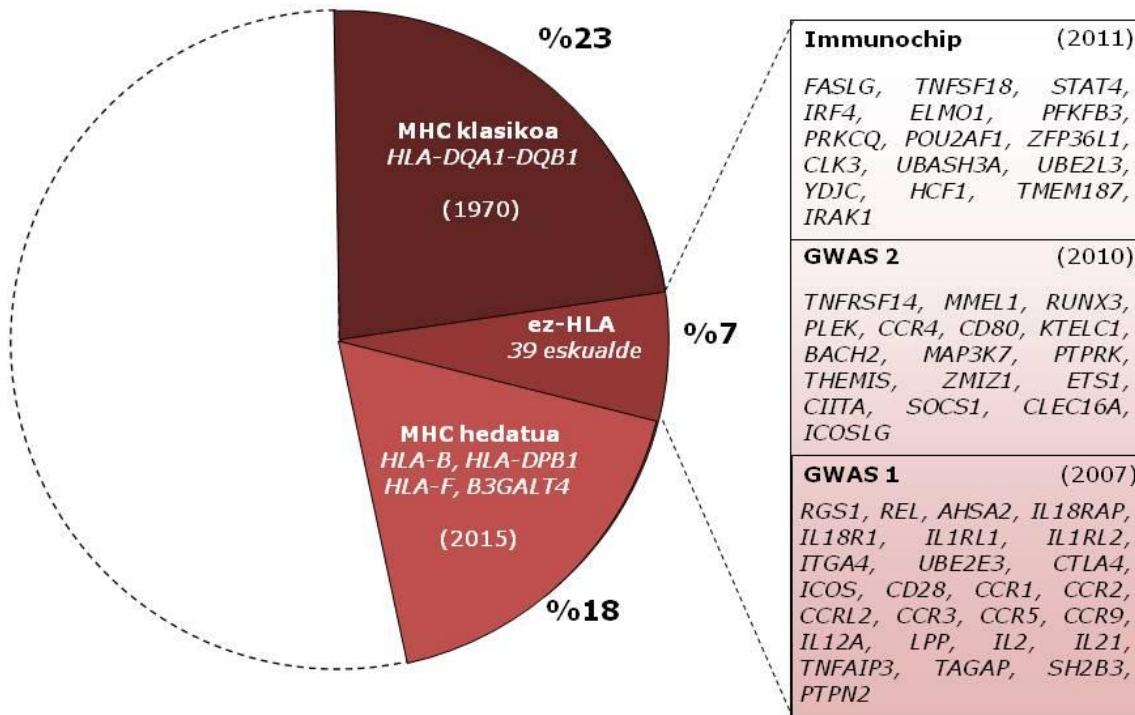
HLAk EZren osagai genetikoaren %40a baino azaldu ezin duenez, azken hamarkadan saiakera ugari egin da HLA eremutik kanpo egon daitezkeen eta EZren Genetika konplexua argitzen lagundu dezaketen suszeptibilitate geneak aurkitzeko. Hasiera batean bi estrategia erabili ziren helburu honetarako: 1) afektatutako familietan lotura ikerketak eta 2) populazioaren baheketan oinarritutako asoziazio ikerketak.

Familietako lotura azterketek gaixoen belaunaldi ezberdinetan zehar modu errepikakor eta kontsistentean heredatzen diren eremu kromosomikoak identifikatzea baimentzen dute. Horrela, gaixotasunaren hasieran zein garapenean garrantzitsuak diren eremuak aukera eta muga daitezke. Eremu genomiko horietan kokaturiko geneei hautagai posizional deritze; izan ere, beren kokapena da, hain zuzen, hautagai izaera ematen diena. MHC lokusean izan ezik, lotura ikerketen emaitzak kontraesankorrak izan dira. Hala ere, HLAz gain (CELIAC1), hiru eskualde kromosomiko agertu dira EZri lotuta ikerketa batean baino gehiagotan: 5q31-33 (CELIAC2) (Greco et al., 1998); 2q32 (CELIAC3; T linfzito erregulatzaileak diren *CD28*, *CTLA4* eta *ICOS* geneak dituen)

(Holopainen et al., 2004) eta 19p13 (CELIAC4; *MYO9B* genea duena) (Van Belzen et al., 2003).

Berrikiago, errendimendu handiko teknologien garapenak GWA azterketak ahalbidetu ditu, eta gaixotasun konplexuen azterketa genetikoak bideratzeko modua erabat aldatu da. GWA azterketak markatzaile genetiko asko (polimorfismoak) DNA lagin multzo handietan (kasu-kontrol) aztertzeko tresna oso baliagarriak dira, horrela, gaixotasun zehatzekin asoziatutako aldaera genetikoak aurkitzeko.

EZn burututako bi GWA azterketetan 5.000 kasu eta 10.000 kontrol inguru aztertu ziren, eta HLA eremutik kanpo dauden 26 eskualde asoziatu ziren gaixotasunarekin (Dubois et al., 2010; van Heel et al., 2007). Urtebete beranduago, Immunochip izeneko plataformaren bidez 12.000 kasu eta 12.000 kontrol inguru genotipatu ziren. Kasu honetan, immunitatearekin erlazionatutako 10 gaixotasun ezberdinetan asoziatuta dauden 186 lokusen mapaketa fina burutu zen. Azken hurbilketa honetan EZri asoziatutako 13 eskualde berri identifikatu ziren (Trynka et al., 2011). Berriki, Immunochiparen berrazterketa batek EZrekin erlazionatutako bost eskualde gehiago aurkitu zituen (Garcia-Etxebarria et al., 2016). Guztira, HLAtik kanpoko 44 eskualde genomiko asoziatu dira EZrekin. Hauek 62 asoziazio seinale independente dituzte, eta guztiak batuz %5-7ko ekarpena egiten diote arrisku genetikoari (4. irudia). Hemeretzi eremuetan gene hautagaiak nabarmendu diren arren, asoziatutako 3 SNP besterik ez daude proteinak trunkatzeko ahalmena duen exoi baten barruan. Gene hautagai batzuk asoziatutako seinaleak 5' edo 3' inguruko eremu erregulatzailan daudelako aukeratu dira.



4. irudia. Aurrerapenak EZren Genetikan. ImmunoChip ikerlanaren ostean HLAtik kanpoko 39 eskualde asoziatu dira gaixotasunarekin. Beranduago, 5 *loci* gehiago aurkitu dira (Garcia-Etxebarria et al., 2016).

Bestetik, *NCF2* geneko rs17849502 aldaera ez-sinonimoaz gain, ez da arriskuan eragina duen beste aldaera arrarorik aurkitu. Azken hori arriskuaren handipen txiki batekin asoziatu da (Hunt et al., 2013).

GWA eta jarraipen azterketak suszeptibilitate lokusak identifikatzeko estrategia arrakastatsuak izan diren arren, aurkitutako aldaerek gaixotasunaren ekarpen genetikoaren proportzio txikia azaltzen dute. Beraz, arriskuan parte hartzen duten lokus gehienak identifikatzeke daude oraindik. Gainera, LD dela eta, identifikatutako SNPak ez dute zertan gaixotasunaren benetako eragile izan behar, zehaztasun handiagoz behatu behar diren eremu asoziatuen markatzaileak baizik. Bestalde, asoziatutako SNP gehienak (%80) sekuentzia kodetzaileetatik kanpo daude (gune erregulatzaile edo gene arteko eremuetan) eta horregatik, euren funtzioa geneen adierazpenaren erregulazioarekin loturik dagoela ondorioztatu da. Honen adibide da NFκB bidezidorraren erregulatzaile den

IL18RAP genearen ondoko RNA ez-kodetzaile luzea (ingelesez *long non-coding RNA* edo lncRNA), EZrekin asoziatuta dagoen SNP bat daramana. lncRNA honek transkripzioko hasiera gunera lotzen den proteina-konplexuari eragiten dio, eta gene inflamatorio batzuen adierazpena bultzatzen du. EZren arrisku aleloak efizientzia txikiagoz lotzen du proteina-konplexua, eta gene inflamatorioen adierazpena emendatzen du eta gaixotasunaren garapena bultzatu (Castellanos-Rubio et al., 2016).

Orokorrean, asoziazio seinale hauen efektuak ulertzeko ikerketa gehienak EZrekin asoziatutako aldaerek inguruko geneen mRNA adierazpen mailak erregulatzen ote dituzten aztertzeraz mugatu dira (ingelesez *expression Quantitative Trait Loci* edo eQTL). Emaitza mugatuak lortu dira eta EZren patogenesisia uste baino konplexuagoa dela dirudi, non gene taldeen edo bidezidorren erantzun koordinatuaren aldaketak partaide funtzional izan daitezkeen (Plaza-Izurieta et al., 2015).

Testuinguru honetan, gure taldeak GWA azterketa batean identifikaturiko hg19 chr6:128.31-128.70 eremuan kokatutako immune sistemako *PTPRK* eta *THEMIS* geneak aztertu zituen. Gene horien adierazpena EZdun pazienteetan positiboki korrelazionatuta dagoela ikusi zuen, kontroletan hau ematen ez zen bitartean (Bondar et al., 2014). Beste esperimentu batean, NFκB bidezidorrarekin erlazionatutako 93 genek mRNA mailan korrelazio sendoa erakutsi zuten EZ pairatzen ez zuten indibiduen heste mukosan, baina gaixotasun aktiboan erregulazio hau deuseztatuta zegoen (Fernandez-Jimenez et al., 2014). Aurkikuntza hauek gaixotasunean koadierazpena, hau da, bi generen edo gehiagoren adierazpenaren korrelazioa, aldatuta dagoela adierazten dute, eta transkripzio faktoreek (TF), RNA ez kodetzaileek (ingelesez *non-coding RNA* edo ncRNA), kromatinaren egiturak, DNAREN metilazioak, etab. gene taldeen adierazpena kontrola lezakete.

Laburbilduz, EZren Genomika guztiz ulertu gabe segitzen dugu, eta gutxi gorabehera, arrisku genetikoaren erdia da ezaguna. Beraz, genomaren erregulazio sistema konplexuan partaide diren elementuak identifikatzeko daude, eta ikertzaileek azterketa sakonak burutu beharko dituzte aldaera estruktural eta epigenetikoak zein genoma ez kodetzaile zabala aztertzeko.

4. Gene erregulazioa

Gene erregulazioa prozesu konplexua da; izan ere, gene bakoitzaren produktu diren mRNA edota proteinen ekoizpen mailan eragiten duten elementu guztiak biltzen ditu. Geneen adierazpena modulatzeko duen erregulazio honek areagotzaileak, promotoreak, TFak eta haien lotura guneak, lncRNAk edo mikroRNA (miRNA) bezalako RNA erregulatzaileak, DNA genomikoaren metilazioa, histonen aldaketak, etab. barnebiltzen ditu. Erregulazioan eman daitezkeen aldaketek gaixotasun ugari eragin dezakete; adibidez, minbizia, autoimmunitatea, gaixotasun neurologikoak, garapen sindromeak, diabetesa, gaixotasun kardiobaskularrak, edota obesitatea (Lee eta Young, 2013).

Gainera, ikusi da gene indibidualek ez dutela bakarrik funtzionatzen, elkarrekin interakzionatzen dute ordea, eta taldeak, gene-sareak edo bidezidorrak sortzen dituzte. Interakzio hauek guztiek giza osasunarengan eragin lezakete. Bataz bestean, gene bakoitzak beste 4-8 generekin interakzionatzen duela uste da (Arnone eta Davidson, 1997), eta 10 bat funtzio biologikotan parte hartzen duela estimatzen da (Miklos eta Rubin, 1996). Horregatik, gene talde edo bidezidorren erantzun koordinatua gaixotasun konplexuetan aztertzeke dagoen parte-hartzaileak izan daiteke, eta gene taldeen erregulazioa aztertu beharko litzateke (Li et al., 2018).

4.1. Transkripzio faktoreak

TFak II RNA polimerasaren bidez katalizatutako gene adierazpena kontrolatzen duten proteina erregulatzailerak dira (Fulton et al., 2009; Vaquerizas et al., 2009). DNA sekuentzia espezifikoetara lotzen dira, eta beraien aktibitateak zelulen funtzioa eta ingurumenaren aurreko erantzuna baldintzatzen ditu. TF batek gene desberdinak erregula ditzake zelula ezberdinetan (Geertz et al., 2012), eta gene baten adierazpen maila ezberdina mantendu zelula mota bakoitzean.

TFek DNAREN sekuentzia espezifikoetara lotzeko domeinuak dituzte eta transkripzioa aktibatu edo inhibit dezakete. DNARI lotzeko domeinuaz gain, aktibazio/errepresio domeinuak ere badituzte. Azken horiek, kofaktore deritzen eta aktibazioan eta inhibizioan parte hartzen duten proteina-konplexuekin interakzionatzen dute. Kofaktoreek transkripzio maila kontrola dezakete II RNA polimerasa aktibatuz edo inguruko kromatina egitura modifikatuz (Bhagwat eta Vakoc, 2015). Ikuspuntu klasikoaren arabera, TFak aktibatzaile edo inhibitzailetan sailkatu izan dira, baina hauetako asko kofaktore desberdinak erreklutatzen direla, eta sailkapena zalantzan dago (Frietze eta Farnham, 2011; Rosenfeld et al., 2006; Schmitges et al., 2016).

TF bakoitzak antzeko DNA sekuentziez osatutako bilduma ezagutzen du, motibo izenez ezagunak. Motibo hauek sekuentzia luzeagoak ikertzeko erabilgarriak dira, lotura-gune potentzialak identifikatzeko. Azken hamarkadan, bai motiboen zein genomako lotura-guneen ezagutza handitu da eta TF-DNA interakzio datu ugari bildu da. Azken urteetan, giza TF katalogoez gain (Fulton et al., 2009; Vaquerizas et al., 2009; Wingender et al., 2015), TRANSFAC (Matys et al., 2006), JASPAR (Mathelier et al., 2016), HT-SELEX (Jolma et al., 2013, 2015; Yin et al., 2017), UniPROBE (Hume et al., 2015), eta CisBP (Weirauch et al., 2014) moduko motiboen bildumak sortu dira. Sekuentziaren testuinguruak, alboetako sekuentziak, DNAREN konformazioak eta elkar eragiten duten kofaktore eta TFek, besteak beste, sekuentziaren ezagutza alda dezakete (Siggers

eta Gordân, 2014). Ezaugarri hauek, TFaren adierazpenarekin eta kromatinaren eskuragarritasunarekin batera, egoera-espezifikoak diren TFen lotura zehazten dute, eta ondorioz, gene erregulazioa (Wang et al., 2012).

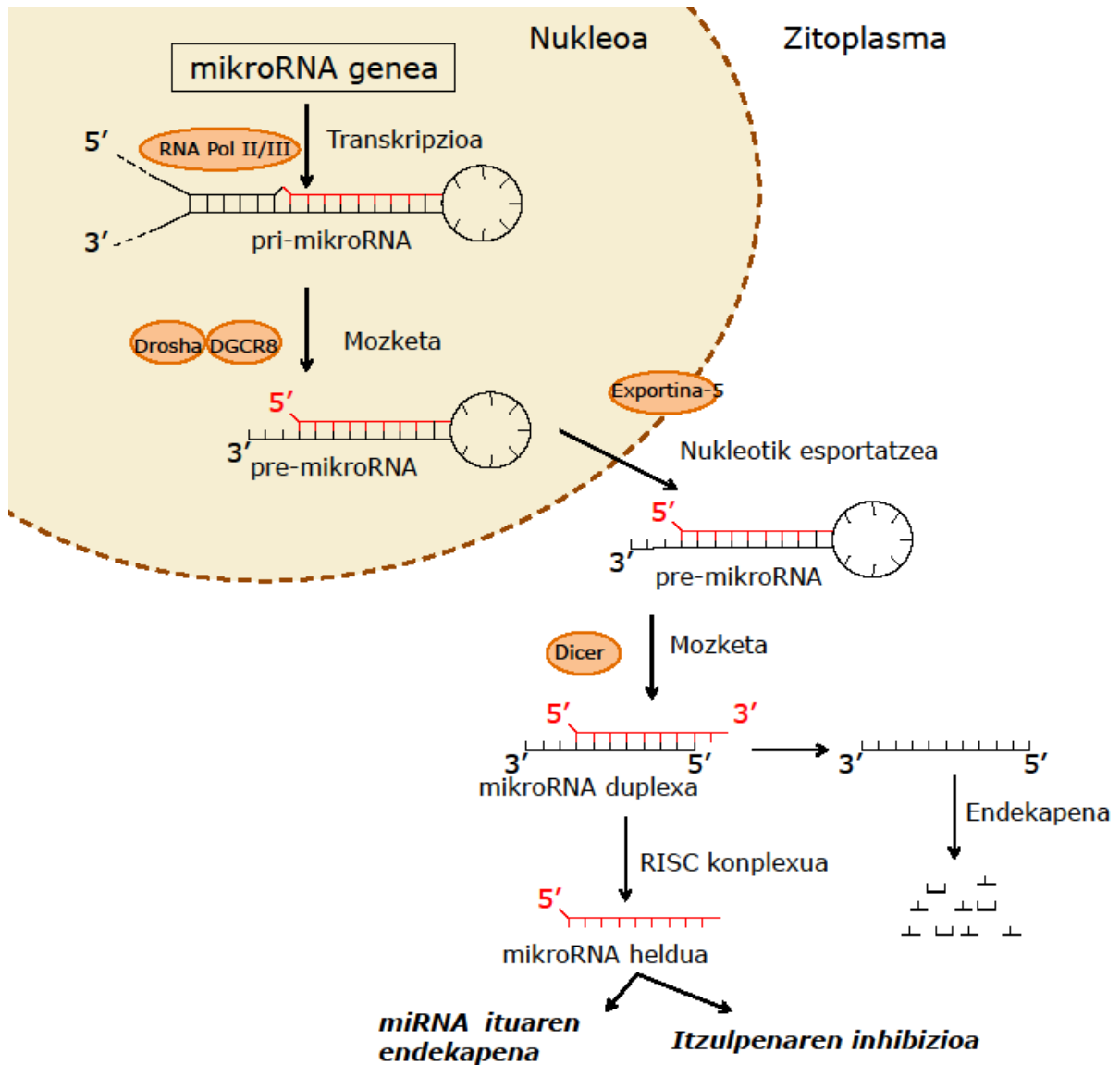
TF batzuk promotoreetara lotzen dira, transkripzioaren hasiera-komplexuaren eraketan laguntzeko; beste batzuk aldiz, erregulatzaileak diren sekuentzia areagotzaileetara lotzen dira transkripzioa estimulatu edo inhibitzeko. Sekuentzia erregulatzaileak transkribatzen den genetik milaka base paretara egon daitezke.

TFek giza gene guztien %8 osatzen dute, eta gaixotasun eta fenotipo ugarirekin erlazionatu dira. TFak kodetzen dituzten geneen mutazioak askotan hilgarriak dira. Honek TFak kodetzen dituzten lokusetan ikusten den kontserbazio maila altua azal lezake. (Bejerano et al., 2004). Ikerketa ezberdinek zenbait gaixotasun-prozesuren gidari modura identifikatu dituzte. Esaterako, bularreko minbizian *AGTR2*, *ZNF132*, *TFDP3* eta beste batzuk erregulatzaile modura identifikatu dira (Tovar et al., 2015); periodontitisean aldiz, 41 erregulatzaile aurkitu dira (Sawle et al., 2016). TF batzuk EZn inplikatu dira, T zelulen diferentziazioaren erregulatzailea den *BACH* barne. Azken hau, gaixotasunean azpiadierazita dago, eta TF horren iturri diren 98 genetan ere adierazpen aldaketak behatu dira (Quinn et al., 2015). Hala ere, EZren patogenesisian eragina izan dezaketen beste TF batzuen kontribuzioa ezezaguna da.

4.2. mikroRNAk

mikroRNAk gene adierazpena erregulatzen duten RNA ez kodetzaile laburrak dira (18-22 nukleotido). ncRNA labur hauek proteina kodetzaileen %60aren itzulpena erregulatzen dutela estimatzen da (Esteller, 2011), eta zenbait bidezidorretan aldaketak sor ditzakete TFen aktibitatea erregulatuz.

miRNAk kodifikatzen dituzten geneak transkribatzean, transkripto primario luzea (ingelesez *long primary transcript* edo pri-miRNA) sortzen da. pri-miRNAk gene proteina-kodetzaile zein ez kodetzaileen exoi (%13-20) eta introietan (%80-87) daude kodetuak (Rodriguez et al., 2004). pri-miRNAk II eta III polimerasen bidez transkribatzen dira (Borchert et al., 2006) eta 70 eta 100 nukleotido arteko luzera dute. Nukleoan pri-miRNA sortu eta Drosha eta DGCR8 entzimek prozesatu egiten dute, 70 nukleotido inguruko RNA aitzindari bat edo gehiago sortzeko (ingelesez *precursor-miRNA* edo pre-miRNA). Exportina-5ak pre-miRNAk lotu eta nukleotik zitoplasmara translokatzeko ditu (Sun et al., 2010). Behin pre-miRNA zitoplasman dagoela, Dicer entzimak moztu egiten du 20 bp inguruko miRNA/miRNA* duplexa sortzeko. miRNA duplex honek bi miRNA heldu desberdin eman litzake, baina normalean kate bakarra lotzen da RISC konplexuan (ingelesez *RNA-induced silencing complex*) miRNA heldu eta funtzional bihurtzeko. miRNA helduen bidezko erregulazioa, osagarriak diren itu mRNA molekulen 3' UTR sekuentzietara lotuz gertatzen da (5. irudia). Parekatze perfektua ematen bada, itu mRNAren endekapena gertatuko da, baina osagarritasuna partziala denean, itzulpena inhibituko da mRNA mailan eraginik izan gabe (Dalmay, 2008). Ondorioz, miRNA mailan gerta daitezkeen aldaketek itu geneen gainadierazpen edo azpiadierazpena eragin dezakete.



5. irudia. miRNA molekulen prozesamenduaren bidezidor kanonikoa. Winter et al., 2009 publikaziotik moldatua.

miRNAk zenbait gaixotasunetan parte hartzen dutela ikusi da, hala nola, minbizian, gaixotasun neurologikoetan, gaixotasun kardiobaskularretan edota gaixotasun inflamatorioetan (Esteller, 2011). miRNAen erregulazioan eman daitezkeen aldaketak immunitateko akatsekin eta autoimmunitatearen garapenarekin ere erlazionatu dira. Hainbat ikerketa gaixotasun autoimmuneetan miRNAek izan dezaketen rolean fokatu dira, eta biomarkatzaile izan daitezkeen adierazpen profil ezberdinak identifikatu dira lupus eritematoso sistemiko, artritis erreumatoide eta Sjögren-en sindrome moduko gaitz autoimmuneetan (Chen et

al., 2016). EZn ere zenbait ikerketa egin da; horietako batzuek zenbait miRNA espezifikoren adierazpena aldatuta dagoela erakutsi dute (Buoli Comani et al., 2015; Capuano et al., 2011; Magni et al., 2014; Vaira et al., 2014), gaixotasunaren patogenesisian eragin dezaketela bai eta diagnosirako tresna izan daitezkeela ere (Felli et al., 2017).

4.3. Kromatinaren egitura

Giza genomaren funtzionamenduaz jabetzeko, zelularen nukleo barnean kromosomen hiru dimentsioko (3D) tolesdura eta antolamendu espaziala ulertzea ezinbestekoa da; izan ere, gene adierazpenaren erregulazioa kromosomen egiturarekin erlazionatua dago (Lupiáñez et al., 2016; Sexton et al., 2007).

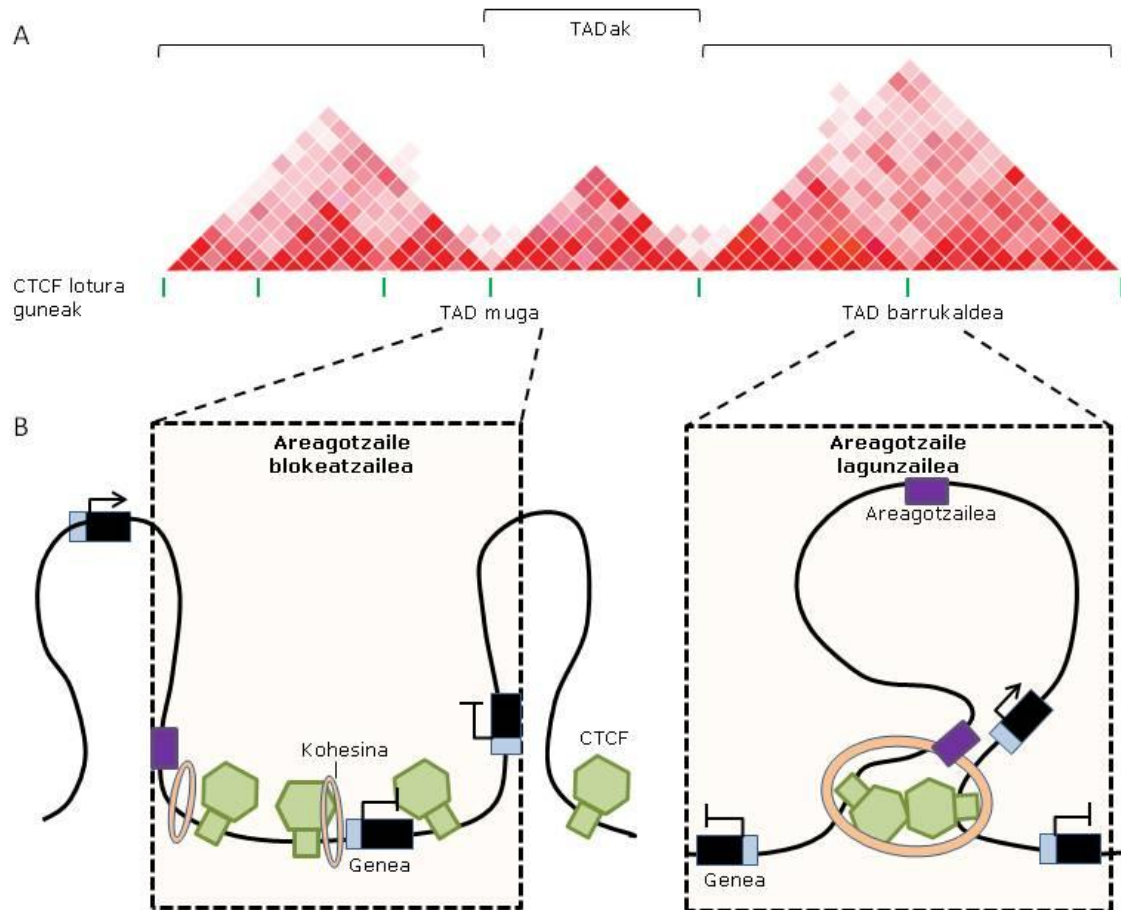
Nukleoan, heterokromatina (kromatina inaktiboa) periferiatik hurbil egon ohi da, eukromatina (kromatina aktiboa) barrurago kokatzen den bitartean (Bickmore, 2013). Kromosomei dagokienez, hauek kromatina aktibo eta irekia, edo kromatina itxia dituzten konpartimentu handiagoetan (A/B konpartimentuak) banatzen dira (Simonis et al., 2006). Konpartimentu aktibo eta irekiek beste konpartimentu aktibo eta ireki batzuekin elkar eragiten dute, konpartimentu itxiek beste itxi batzuekin egiten duten modura. Konpartimentu hauek zelula motaren arabera ezberdinak dira, zelula desberdinetan gene sorta ezberdinak adierazten baitira. Hala ere, kromosomaren megabasetik beheragoko antolamenduaren ulermena mugatua da oraindik.

Berriki, ingelesez *Topologically Associating Domain* edo TAD izena hartzen duten eskualde genomikoak deskribatu dira. Aurretik azaldutako A/B konpartimentuak TADez osatuta daude. Beraz, TADak kromatinaren oinarriko unitate estruktural zein funtzionalak dira (Dekker eta Heard, 2015; Dixon et al., 2016). Era berean, TADak kromatinaren frekuentzia handiko interakzioen domeinu espazialak dira (Dixon et al., 2012). Kromatinaren domeinu hauek, gutxi gorabehera, 1 Mb-ko luzerakoak dira, eta zelula mota eta espezie

desberdinen artean egonkorrak eta kontserbatuak dira. TADak kromatina-kromatina elkarrekintzetan aberatsak dira, domeinu ezberdinen arteko elkarrekintzak urriak diren bitartean.

Kromatinaren eskualdeen arteko elkarrekintzak gene erregulazioan garrantzitsuak direla uste da. Izan ere, TADek osatzen dituzte karakterizatutako areagotzaile-promotore bikote gehienak (Rao et al., 2014). Areagotzaile hauek domeinu erregulatzailerik handietan kokatzen dira, eta TADen barruan eragiten dute korregulazioa bideratzeko (Le Dily et al., 2014).

Kromatinaren 3D egitura hainbat faktoreren menpe dago (Phillips-Cremins et al., 2013), baina haien artean CTCF (ingelesez *CCCTC-binding factor protein*) eta kohesina dira garrantzitsuenak. CTCF da ornodunetan deskribatu den proteina isolatzaile nagusia, eta kromatinaren mugak definitu eta begizten eraketan eragiten du (Ong eta Corces, 2014). Hala ere, genomako CTCF lotura-guneen %15 besterik ez dago TAD mugetan, %85 TADen barruan dagoen bitartean (Dixon et al., 2012). CTCF hauen funtzioa TAD barruko areagotzaileak gene espezifikoen promotoreetara zuzentzea izan liteke (6. irudia).



6. irudia. CTCF-bidezko begizten eraketaren karakterizazioa. A) TADen eta haien arteko mugez irudi eskematikoa. B) TAD mugetan dauden CTCF lotura guneek mugak finkatzen dituzte, eta CTCF proteinek areagotzaileen blokeatzaile modura jarduten dute. Bestetik, TAD barruko CTCF proteinek areagotzaile-promotore begizten eraketan laguntzen dute. Laukizuzen urdinek promotorea irudikatzen dute. Ong eta Corces, 2014 publikaziotik eratorria.

Berriki, 3D genomaren hausturaren eta gaixotasunen arteko lotura ikusi da. Giza gaixotasunetan, alterazio hauek 3D genomaren antolatzaileetan egon daitezkeen mutazioen, SNPen, intserzio edo delezio txikien, eta kromosomopaten ondorioz sortzen dira. Kasu gehienetan zaila da gaixotasunari asoziatutako 3D genomaren egitura aldaketak gaixotasunaren kausa edo ondorio diren jakitea, baina genomaren egituraren aberrazioek funtzioan eragin lezakete, eta ondorioz, gaixotasunen garapenean (Krumm eta Duan, 2018).

Gutxienez bi mekanismo aurkitu dira genomaren 3D egituraren gertatzen diren alterazioen bidez gaixotasunari asoziatutako adierazpen aldaketak eragin eta

ondorioz, gaixotasuna garatzeko. Lehenengo mekanismoan, gaixotasunari asoziatutako SNPeK promotore-areagotzaile interakzioak aldatzen dituzte eta gene adierazpen okerra eta gaixotasuna eragiten dituzte. Ikerketa askok frogatu dute mekanismo hau giza gaixotasunetan, hala nola, arteria koronarioetako gaixotasunean (Mumbach et al., 2017), hesteetako hanturazko gaixotasunean (Meddens et al., 2016), eta artritis erreumatoide, I motako diabetes, artritis psoriatiko eta gaztaroko artritis idiopatiko moduko gaixotasun autoimmuneetan (Martin et al., 2015). Ikerketa horiek erakutsi dute gaixotasunekin asoziatutako SNP erregulatzailleek gene adierazpena aldatu dezaketela, gene eta elementu erregulatzailen arteko kromatina-interakzioen erregulazioaren bidez.

Bigarren mekanismoan, TADen mugetako alterazioek bata bestearen alboan dauden bi TADen arteko fusioa eragin dezakete, eta horrek normalean elkarrengandik isolatuta dauden areagotzaile eta geneen arteko komunikazioa ahalbidetuko du. Areagotzailearen adopzio edo bahiketa (ingelesez *hijacking*) honek gaixotasunean garrantzitsuak diren geneen aktibazio ektopikoa baimentzen du. Hau horrela, TADen haustura eta ondoriozko areagotzaileen bahiketak eragindako zenbait gertaera patogeniko aurkitu da (Lupiáñez et al., 2016; Valton eta Dekker, 2016). TADen mugetako CTCF moduko egitura-proteinen loturaguneen disrupzio genetikoek edo inaktibazio epigenetikoek, bata bestearen alboan dauden bi TADen fusioa eragin dezakete (Flavahan et al., 2016; Lupiáñez et al., 2015). TAD muga oso baten delezioak ere TAD egitura aldatzen du (Giorgio et al., 2014). Berrantolaketa genomikoak; inbertsioak, delezioak, duplikazioak eta translokazioak esaterako, TADak puskatzeko eta hauen arteko fusioak edo TAD berriak sortzeko gai dira (Lupiáñez et al., 2015; Weischenfeldt et al., 2017).

TADek gene erregulazioan duten paper garrantzitsua kontuan hartuta, TADen eraketa ezagutzea eta beren mugek begizten interakzioetan nola eragiten duten ikertzea erabilgarria izan liteke gene erregulazioa ulertzeko, baldintza fisiologiko zein patologikoetan.

4.4. DNAREN metilazioa

DNAREN metilazioa marka epigenetikoa da; DNA metiltransferasek (DNMTs) DNAREN zitosina basearen C-5 posizioan metilo talde bat ezartzearen ondorioa (Robertson, 2005). DNAREN metilazioaren %98 zelula somatikoan CpG dinukleotidoetan ikusten da (Lister et al., 2009).

Orokorrean, DNAREN metilazioa transkripzioaren errepresioarekin erlazionatu da TF eta DNARI lotzen zaizkion beste proteinak oztopatzen dituelako (Bird, 2002). Hala ere, genoma osoko metilazio azterketek erakutsi dute DNAREN metilazioa ez dagoela beti gene adierazpenarekin negatiboki korrelazionatuta. DNAREN metilazioak gene adierazpena oztopatzen duen ikuspegi tradizionala zalantzan jarri da, DNAREN metilazioa gene adierazpenarekin positiboki ere korrelaziona daitekeela ikusi baita (Irizarry et al., 2009).

DNAREN metilazioa garapen normalerako beharrezkoa da, eta inpronta genomiko, X kromosomaren inaktibazio eta elementu errepikakorren transkripzio eta transposizioaren murrizketa moduko oinarrizko prozesuetan oso garrantzitsua da. Informazio epigenetiko hau behar bezala finkatzen ez bada, minbizi moduko gaixotasunak sor daitezke (Delpu et al., 2013; Robertson, 2005). Hauetan, metilazioaren galera ohiko gertaera da, eta gaixotasunaren larritasunarekin eta mota ezberdinetako tumore metastasikoekin korrelazionatu da (Widschwendter et al., 2004; Zhang et al., 2015).

Metilazioa zelula immuneen funtzionamendu egokian kritikoa da. Ikerketa berriek, DNAREN metilazio mailan ematen diren alterazioak autoimmunitatearekin erlazionatu dituzte. Ingurumen faktoreek eta polimorfismo genetikoek metilazio patroia okerrak eragin ditzakete, eta hauek immunitatearekin erlazionatutako zenbait generen transkripzio mailetan eragin dezakete (Sun et al., 2016).

EZri dagokionez, zeliako aktibo, tratatutako zeliako eta kontrolen heste biopsietan egindako metilazio azterketa batean, NF κ B bidezidorreko zenbait generen promotoreetan diferentziak ikusi ziren (Fernandez-Jimenez et al., 2014), eta metilazioak EZren patogenesisian parte hartu dezakeela ikusi da. Ikerketa horretan, metilazio mailen alterazioak soilik partzialki itzulgarriak direla ikusi zen GGDn zeuden pazienteetan, eta dieta aldaketaren ondoriozko metilazio alterazioen itzulezintasuna frogatu.

Helburuak

Lan honetan EZren patogenesisian erregulatzaile genetiko eta epigenetikoaren inplikazioa ulertu nahi da, eta horretarako lau helburu nagusi definitu dira:

1. Gliadinarekiko esposizio akutuak eta epe-luzekoak EZ pairatzen duten pazienteen heste meharrean genoma osoko koadierazpen patroiak aldatu ditzakeen zehaztea.
2. EZn ikusitako koadierazpen aldaketan atzean dauden elementu erregulatzaileak identifikatzea; zehazki TFak eta miRNAk, eta hautagai batzuen balioztatze esperimentuak burutzea.
 - a. TF eta miRNA hautagaien, eta beraien itu geneen adierazpena aztertzea EZ duten pazienteen eta kontrolen duodenoko biopsietan.
 - b. Gliadinaren aurreko erantzunaren ondoriozko TF hautagaien nukleorako translokazioa eta ituen promotoreekiko lotura ikertzea, eredu zelularretan.
3. EZren patogenesisian eragina izan dezaketen TADen mugetako balizko alterazioak bilatzea, *in silico* hurbilketak eta eskuragarri dauden datu genomikoak erabiliz, eta TADen fusioa eta haustura horiek daramatzaten eredu zelular egonkorak sortzea, koadierazpena aldarazten duten ikusteko.
4. Gliadinarekiko erantzun akutuaren ondoriozko DNAREN metilazioaren aldaketak ikustea eta EZn, *HLA-B* eta *TAP1* lokusetan metilazioaren parte-hartzea egiaztatzea.

Materialak eta metodoak

1. Materialak

1.1. Pazienteak

1.1.1. Onarpen etikoa

Burutu diren azterketa guztiak PI13/01201 eta PI16/00258 (Espainiako ekonomia eta Lehiako Ministerioak eta Europako Eskualde Garapenerako Funtsak ko-finantziatutako Zientzia-ikerketa, Garapen eta Teknologia Berrikuntzarako 2013-2016 Plan Nazionalaren barruan), eta Eusko Jaurlaritzako Osasun Saileko 2011/111034 ikerketa proiektuen parte dira. Guztiak Gurutzeta Unibertsitate Ospitaleko Etika Batzordeak eta Euskadiko Ikerkuntza Batzorde Etikoak onartutako azterketak dira (CEIC-E13/20, CEIC-E16/46 eta PI2013072 kodeekin, hurrenez hurren). Lagin guztiak errutinazko diagnostikoetako endoskopian bidez jaso dira, emaileen edota beren gurasoen baimen informatua lortu ostean.

1.1.2. Pazienteak eta biopsia laginak

EZ erreklutamenduaren garaian indarrean zegoen ESPGHANen irizpideen arabera diagnostikatu zen (Husby et al., 2012) Gurutzeta Unibertsitate Ospitaleko Gastroenterologia Pediatrikoko Unitatean. Horretarako, anti-gliadina eta EMA edota TGA antigorputzen determinazioak eta heste meharreko biopsia erabili ziren. Lan honetarako, guztira, EZ zuten pazienteen eta indibiduo ez-zeliakoen 92 hesteko biopsia jaso ziren endoskopia bidez (1. taula).

Pazienteak ondorengo 3 taldeetan sailka daitezke:

- EZ aktiboa (29 indibiduo): diagnostikatu berri diren EZdun pazienteak, gaixotasun klinikoki aktiboa dutenak (EZrekin asoziaturiko antigorputzak, hesteetako biloxken atrofia eta kripten hiperplasia) eta dietan glutena hartzen dutenak.

- GGD-EZ (37 indibiduo): gutxienez 2 urtetan GGD jarraitu duten EZdun paziente normalizatuak (sintomarik gabe, antigorputz negatiboak eta heste epitelio osasuntsua).
- Kontrol taldea (26 indibiduo): endoskopiaren momentuan hesteetako inflamaziorik ez zuten indibiduo ez-zeliakoak.

1. taula. Ikerketa honetan parte hartutako pazienteen informazio klinikoa, immunologikoa eta genetikoa. Dx, Diagnosian; TGA, anti-transglutaminasa autoantigorputzak; HLA, Giza Antigeno Leukozitarioa.

Ezaugarriak	
Generoa	Emakume: %73; Gizon: %27
Adina Dx (hilabeteak)	31,7 ± 22,9
TGA Dx	91.66%*
MARSH balioa Dx	3c: %81,25; 3b: %18,75
HLA	DQ2: %82,6; DQ8: %4,4; DQ2/DQ8: %13

*Paziente bat < urte 1, AGA positiboa; Bi paziente, tTG IgG positiboak

1.2. Zelula lerroak eta zelula kultura

Epitelioko giza zelulak *American Type Culture Collection*-etik (ATTC, Manassas, VA, USA) eskuratu dira. Zehazki, koloneko ehunetatik eratorritako C2BBel (Caco-2 zelulen klonak: kat. zk. CRL-2102), HCT116 (kat. zk. CCL-247), eta HCT15 (kat. zk. CCL-225) zelulak, eta giltzurrun enbrionarioen ehunetatik eratorritako HEK293FT (kat. zk. CRL-1573) zelulak erabili dira.

Epitelioko giza zelulak 37°C-tan eta %5 CO₂-an hazi ziren, ingurune heze eta esterilean. Zelula mota bakoitzarentzat hazkuntza medio aproposa prestatu zen (2. taula). DMEM eta DMEM F-12 Lonza-n (Lonza group, Basel, Switzerland; kat. zk. BE12-604F eta BE12-719F) erosi ziren, eta RPMI Medium 1640 (1X) Gibco-n (Gibco, Carlsbad, CA, USA; kat. zk. 21875-034). Medioak penizilina-estreptomizinarekin (Lonza group; kat. zk. DE17-602E), aminoazido ez esentzialekin (NEAA) (Lonza group; kat. zk. BE13-114E) eta inaktibatutako

behi-sero fetalarekin (FBS) (Biochrom, Cambridge, UK; kat. zk. S 0415) suplementatu ziren.

2. taula. Zelula lerroen hazkunde-medioak. DMEM, *Dulbecco's Modified Eagle's* Medioa; RPMI, *Roswell Park Memorial Institute*; FBS, behi-sero fetala; Pen-Strep, penizilina-estreptomizina antibiotikoa; NEAA, aminoazido ez esentzialak.

Zelula lerroak	Medioa	FBS	Pen-Strep	NEAA
C2BBe1	DMEM (<i>high glucose</i>)	%10	%1	0,1 mM
HCT116	DMEM F-12	%10	%1	-
HCT15	RPMI Medium 1640	%10	%1	-
HEK293FT	DMEM F-12	%10	%1	-

Hazkuntza medioak 2-3 egunetik behin aldatu ziren, eta zelulak %0,25 Tripsina-EDTA (Gibco; kat. zk. 25200056) erabiliz plakatik desitsatsi eta plaka berri batean erein ziren %70-80ko konfluentziara iristean. Kriokontserbaziorako, zelulak %5 DMSOdu (Sigma-Aldrich, St. Louis, MO, USA; kat. zk. 67-68-5) hazkuntza medioan berreseki eta hodiak -80°C-ko izozkailuan, isopropanolez betetako kutxan murgildu ziren 24 orduz, -1°C/min izozteko. Ondoren, nitrogeno likidoan gorde ziren. Izoztutako zelulak berrabiarazteko, zelulak 37°C-ko ur bainuan pare bat minutuz desizoztu eta hazkuntza medioan berreseki ziren.

1.3. Zelulen eta biopsien kitzikapena

Zelula lerroen eta biopsien *in vitro* kitzikapenerako, gliadina pepsina eta tripsinaz digeritu zen (PT-G), aurretik deskribatu bezala (Bondar et al., 2014). Laburbilduz, 500 mg gliadina (Sigma-Aldrich; kat. zk. G3375) 5 ml 0,2 M HCl-tan disolbatu eta 5 mg pepsinarekin (Sigma-Aldrich; kat. zk. P6887) inkubatu zen etengabeko agitazioz (37°C, 2 h). Inkubazioa amaituta, pHa 7,4ra doitu zen NaOHrekin, eta nahasketa 5 mg tripsinarekin (Sigma-Aldrich; kat. zk. T9935) inkubatu zen 37°C-tan 4 orduz. Entzimak inaktibatzeke, soluzioa 30 minutuz jarri zen irakiten eta ondoren, 2.000 g-tan zentrifugatu zen 10 minutuz. Azkenik,

gainjalkina 20 µm-ko porodun mintzean zehar iragaziz esterilizatu, alikuotatu eta -80°C-tan gorde zen. Behi-sero albumina (BSA) (Thermo Fisher Scientific; kat. zk. SH30574.03) modu berean digeritu zen (PT-BSA) eta kontrol moduan erabili zen.

C2BBel zelulak 37°C-tan eta %5 CO₂-an inkubatu ziren, eta 1mg/ml PT-G edo PT-BSAdun medioarekin kitzikatu ziren 4 orduz. Biopsien estimulaziorako, lau zati ehun hartu ziren 7 GGD-EZ eta 8 kontroletatik. Indibiduo bakoitzetik bi zati 150 µl RPMI-1640 10X mediotan inkubatu ziren 37°C-tan eta %5 CO₂-an 4 orduz, 250 µg/ml PT-Grekin (kitzikatua) edo gabe (kitzikatu gabea).

1.4. DNA eta RNA erauzketa

DNA eta RNA NucleoSpin Blood eta NucleoSpin RNA kit-ak (Macherey-Nagel, Düren, Germany; kat. zk. 740955.250 eta 740951.250) erabiliz erauzi ziren, hurrenez hurren, horretarako hornitutako protokoloak jarraituz. Laburbilduz, izozturiko biopsiak plastikozko kirtenekin birrindu ziren. Zelula kulturak, aldiz, zentrifugazioz bildu ziren.

DNAren erauzketarako, 25 µl K proteinasa eta 200 µl B3 buffer gehitu zitzaizkien laginei, eta 70°C-tan inkubatu ziren 20-30 minutuz. Laginei 210 µl etanol gehitu zitzaien, eta NucleoSpin Blood zutabeetan kargatu ziren. Laginak 11.000 g-tan zentrifugatu ziren minutu batez, eta garbiketa pausuak burutu ziren. DNA H₂Otan eluitu zen.

RNAren erauzketarako, 350 µl RA1 buffer eta 3,5 µl β-mercaptoethanol gehitu zitzaizkien bai zelula pellet zein apurtutako biopsiei, eta laginak NucleoSpin Filter zutabeetan kargatu ziren. Homogeneizatutako lisatuari 350 µl etanol gehitu zitzaion eta NucleoSpin RNA zutabeetan kargatu zen. Laginak 11.000 g-tan zentrifugatu ziren 30 segundoz eta 350 µl MDB gehitu zitzaien. Laginak 11.000

g-tan zentrifugatu ziren 30 segundoz eta 95 µl DNasa erreazio nahastura gehitu zen zutabeetara. Garbiketa pausak eta gero, RNA H₂Otan eluitu zen.

DNA eta RNA laginen kontzentraioa NanoDrop 1000 espektrometroan (Thermo Fisher Scientific, Boston, MA, USA) neurtu zen eta laginak -80°C-tan gorde ziren erabili arte.

1.5. Datu multzoak

Koadierazpen analisirako, aurretik burututako Human U133 Plus 2.0 Array esperimentuen datuak (Affymetrix, Santa Clara, CA, USA) (Castellanos-Rubio et al., 2008) deskargatu ziren EBI Array Express datu-basetik (<http://www.ebi.ac.uk/microarray-as/ae/>):

- Epe-luzeko esperimntua: E-MEXP-1828 esperimntutik hartu ziren datuak. Esperimntu honetan diagnosiaren momentuan glutena hartzen zuten eta EZ aktiboa zuten 9 pazienteren datuak, eta GGDz bi urtez baino gehiagoz tratatutako 9 pazienterenak daude.
- Esperimntu akutua: E-MEXP-1823 esperimntutik hartu ziren datuak. Esperimntu honetan GGDz tratatutako 10 pazienteren biopsiak bitan zatitu eta RPMI medioan inkubatu ziren, 10 µg/ml gliadinarekin eta gliadina gabe, 4 orduz.

Bestalde, SRP077708 zenbakidun RNA-seq esperimntuko datuak eskuratu ziren NCBIren *Sequence Read Archive*-etik (SRA) (<http://www.ncbi.nlm.nih.gov/sra/>). RNA-seq-a HiScanSQ plataforma batean burutu zen 75 nukleotidoko mutur parekatuko sekuentziarioz.

- Epitelioko zeluletan aberastutako frakzioa: epitelioko zelulek EpCAM (CD326) epitelio zelulen adhesio molekula adierazten dute euren gainazalean. EZ aktiboa duten 10 pazienteren eta zeliakoak ez diren 12 kontrolen datuak daude eskuragarri.

- Immunitate zeluletan aberastutako frakzioa: leukozitoek *PTPRC* geneak kodetutako CD45 antigenoa daramate. EZ aktiboa duten 7 pazienteren eta zeliakoak ez diren 5 kontrolen datuak daude eskuragarri.

EZn DNA metilazioa ikertzeko, aurretik egindako Illumina Infinium HumanMethylation450 *microarray*-en datuak deskargatu ziren *Gene Expression Omnibus* (GEO) datu-basetik (<https://www.ncbi.nlm.nih.gov/geo>); GSE847445 esperimenduak EZ aktiboa duten 10 pazienteren eta zeliakoak ez diren 10 kontrolen duodenoko biopsietatik separatutako epitelio zeluletan eta immunitate zeluletan aberastutako frakzioen datuak biltzen ditu.

Tesi dokumentua zein tesi honetako eranskinak (1-9 [eranskinak](#)) eta hauen goiburuak <https://labur.eus/SvteM> helbidean deskargatu daitezke.

2. Metodoak

2.1. Genoma osoko koadierazpena EZn

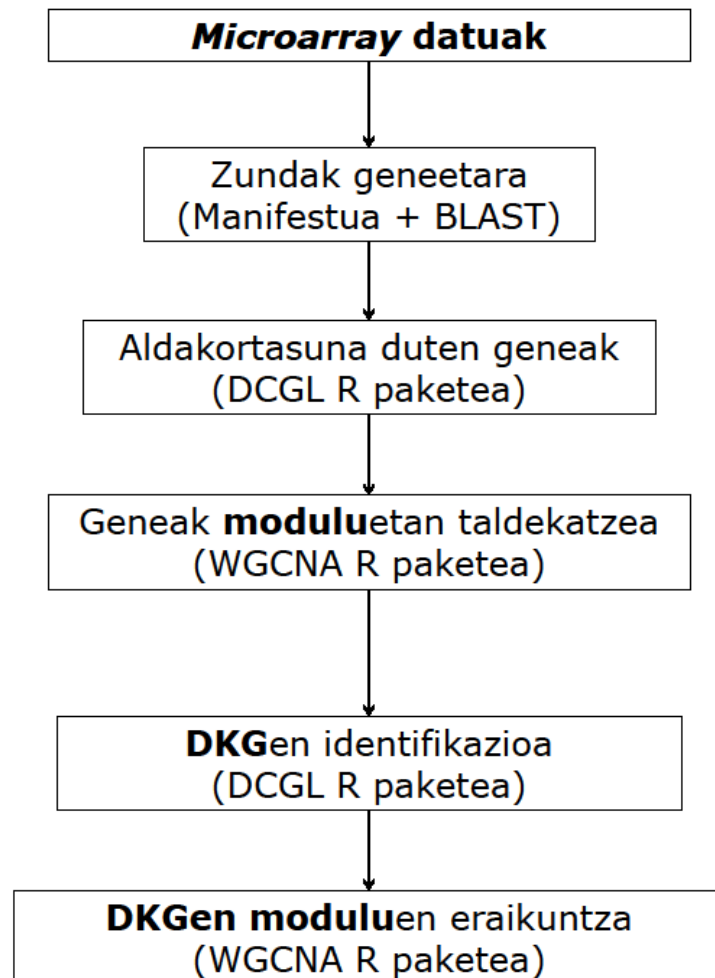
2.1.1. TF eta miRNAen identifikazioa koadierazpen aldaketetan

2.1.1.1. Koadierazpen analisia

Koadierazpen analisiak *microarray* datuetan egin ziren. Adierazpen *array*-eko zunden kodeak geneen izenetara itzuli ziren *array*-aren manifestua edo BLAST erabiliz (Altschul et al., 1997), zunden sekuentziak hg19 erreferentziako giza genomarekin lerrokatuz. Zunda bat baino gehiagok ordezkaturako geneen kasuan, zunda guztien intentsitatearen mediana hartu zen genearen adierazpen maila modura.

Koadierazpen analisisetarako gene informatiboenak lehenesteko, medianak baino adierazpen aldakortasun handiagoa zuten geneak soilik hautatu ziren. Horretarako, DCGL R paketearen “*varianceBasedfilter*” funtzioa erabili zen

(Yang et al., 2013). Koadierazpen moduluak *Weighted Correlation Network Analysis* (WGCNA) R paketea erabiliz definitu ziren (Langfelder eta Horvath, 2008). “Eskalarik gabeko topologia ereduaren doitze” analisiaren balio onenarekin “potentzia” parametroa ezarri zen. Gainera, gliadinaren eraginez sortutako koadierazpen aldaketak ere aztertu ziren, diferentzialki koadierazitako geneak (DKG) identifikatzeko; hau da, epe luzeko esperimentuan zein esperimentu akutuan, lagin taldeen arteko konparaketetan gainontzeko geneekin zituzten erlazioak aldatzen dituzten geneak. DKGak DCGL paketea erabiliz identifikatu ziren, eta p-balioa $< 0,05$ zutenak onartu ziren. Azkenik, DKG hauek baldintza desberdinetan sortzen zituzten koadierazpen moduluak identifikatu ziren WGCNA paketea erabiliz. Lan honetan erabilitako prozedura 7. irudian laburbildu da.



7. irudia. Microarray datuen analisirako prozedura.

2.1.1.2. Erregulatzailer hautagaien aukeraketa

Koadierazpen moduluetakoko DKGak genoma osoarekin alderatu ziren gainordezkaturako miRNA eta transkripzio faktore lotura guneak (ingelesez *transcription factor binding sites* edo TFBS) detektatzeko. Horretarako, Babelomics v4.3-n eskuragarri dagoen FatiGO erreminta erabili zen (<http://v4.babelomics.org>) (Medina et al., 2010). Orokorrean, erremintaren parametro lehenetsiak erabili ziren (FDR-z zuzendutako $p < 0,05$). miRNAen kasuan, miRBase Target datu-basea (Griffiths-Jones et al., 2006) erabili zen miRNA eta horien itu geneen erlazioak identifikatzeko. TFei dagokienez, TFBSen datu-base den TRANSFAC (Wingender, 2000) erabili zen TF bakoitzeko espezifikoak diren motiboak, geneen promotoreekin konparatzeko. TFBS zehatz baterako aberastuak zeuden DKG taldeak, zein modulu original osoak, FatiGO erreminta erabiliz genoma osoarekiko konparatu ziren, modu esanguratsuan gainordezkaturata zeuden *Gene Ontology* (GO) terminoak identifikatzeko.

2.1.2. Hautagaien egiaztatze esperimentalak

2.1.2.1. Geneen adierazpen analisia

mRNAren transkripzio-maila denbora errealeko PCR kuantitatiboaren (qRT-PCR) bidez kuantifikatu zen duodenoko biopsietan.

2.1.2.1.1. Gene hautagaiak eta entseguak

Balioztatze esperimentalak burutzeko, aberaste analisisian esangarriak izan ziren TF eta miRNA batzuk hautatu ziren ondorengo irizpideen arabera:

- 1) Gutxienez hiru modulu desberdinetan errepikatuta azaltzea.
- 2) Aurretiko literatura.

Irizpide hauen aplikazioaren ondorioz 5 TF hautatu ziren: ELK1, NFKB1, HOXA5, CREB1 eta IRF1.

miRNAen kasuan, irizpide berdinen aplikazioaren ondorioz, 9 miRNA aukeratu ziren. miRNA hautagaien pri-miRNAk hautatu ziren adierazpen genikoa neurtzeko (3. taula). Horietako biren miRNA helduak ere aztertu ziren.

3. taula. miRNA heldu hautagaiak eta horien pri-miRNAk. Begizten sekuentzien koordinatu genomikoak (GRCh37/hg19) eta miRBase-ko kodeak azaltzen dira taulan.

miRNA primarioa	Begizten sekuentziak		miRNA heldua
	Kokapena (hg19)	miRBase kodeak	
hsa-mir-33a	chr22:42296948-42297016	MI0000091	hsa-miR-33a
hsa-mir-92a-1	chr13:92003568-92003645	MI0000093	hsa-miR-92a
hsa-let-7b	chr22:46509566-46509648	MI0000063	hsa-let-7b-3p
hsa-mir-503	chrX:133680358-133680428	MI0003188	hsa-miR-503
hsa-mir-655	chr14:101515887-101515983	MI0003677	hsa-miR-655
hsa-mir-17	chr13:92002859-92002942	MI0000071	hsa-miR-18a-3p
hsa-mir-26b	chr2:219267369-219267445	MI0000084	hsa-miR-26b
hsa-mir-520b	chr19:54204481-54204541	MI0003155	hsa-miR-520b
hsa-mir-520e	chr19:54178965-54179051	MI0003143	hsa-miR-520e

TF hautagaien itu geneak ENCODE datuak (<https://www.encodeproject.org/>) eta California Santa Cruz Unibertsitateko (UCSC) genoma arakatzailerak (<https://genome.ucsc.edu/>) erabiliz aukeratu ziren, ondorengo irizpideak jarraituz:

- 1) TF jakin baterako aberastuta zeuden DKG multzoetan *in silico* iragarritako itu geneak, ondorengoei lehenetasuna emanaz:
 - a. Promotorean TFaren lotura gune kontserbatuak dituzten itu geneak.
 - b. Promotorean TFaren lotura erakutsi dutenak, kromatinaren immunoprezipitazio eta sekuentziazio (ChIP-seq) datuen arabera.
 - c. Promotorean kromatina aktiboaren marka den H3K27Ac guneak dituztenak.
 - d. Literaturaren arabera EZren patogenesisirako adierazgarriak direnak.

- 2) DKGen artean egon ez arren, TF zehatz baten ondo definitutako itu geneak. Itu geneak, bere adierazpena kontrolatzen dituzten TFen kopuruaren arabera sailkatu ziren (UCSC genoma arakatzailerako gizaki, sagu eta arratoien lotura gune kontserbatuen tauletan oinarrituta). TF gutxien erregulaziopean zeuden geneak aukeratu ziren, modu horretan TF hautagai bakoitzeko espezifikokoagoak izan zitezten.

Aukeratutako geneak 4. taulan zerrendatu dira, adierazpen analisietan erabilitako hasle eta zunda multzoekin batera. Adierazpen azterketa bakoitzeko bi hasle eta FAMekin markatutako MGB zunda batez osaturiko TaqMan entsegu espezifikoa erabili zen (Thermo Fisher Scientific). DNA genomikoaren detekzioa saihesteko, aukera zegoen guztietan exoi-exoi lotura zeharkatzen duten zundak aukeratu ziren.

4. taula. TFen, itu geneen, miRNAen eta pri-miRNAen adierazpen analisietan erabilitako TaqMan entseguak. Kontrol endogeno modura, *TBP* (gene proteina-kodetzailerak eta pri-miRNAen kasuan) eta *RNU48* (miRNAen kasuan) erabili ziren.

TFak	TaqMan entseguen Id-a	Kokapena (hg19)
<i>ELK1</i>	Hs00901847_m1	chrX:47494920-47509887
<i>NFKB1</i>	Hs00765730_m1	chr4:103422486-103538459
<i>HOXA5</i>	Hs00430330_m1	chr7:27180671-27183287
<i>CREB1</i>	Hs00231713_m1	chr2:208394616-208468155
<i>IRF1</i>	Hs00971960_m1	chr5:131817301-131826490
Itu geneak		
<i>AKTIP</i>	Hs01591423_m1	chr16:53524952-53537163
<i>NAMPT</i>	Hs00237184_m1	chr7:105888731-105925638
<i>TPK1</i>	Hs01558699_m1	chr7:144149034-144533146
<i>ISG15</i>	Hs00192713_m1	chr1:948803-949920
<i>HIST1H4C</i>	Hs00543883_s1*	chr6:26104104-26104538
<i>CRTAM</i>	Hs00219699_m1	chr11:122709208-122743347
<i>PLLP</i>	Hs00762550_s1*	chr16:57290009-57318599
<i>RFX5</i>	Hs00230841_m1	chr1:151313116-151319727
<i>NKG7</i>	Hs01120688_g1*	chr19:51874866-51875969
<i>RAB17</i>	Hs00940833_m1	chr2:238482965-238499736
<i>CISD2</i>	Hs00391903_m1	chr4:103790135-103810399
<i>HDAC4</i>	Hs00195814_m1	chr2:239969864-240322643
<i>WDR43</i>	Hs01064086_m1	chr2:29117509-29171088
<i>CXCL11</i>	Hs04187682_g1*	chr4:76955843-76962568

<i>BATF2</i>	Hs00912737_m1	chr11:64755415-64764517
<i>WNT11</i>	Hs00182986_m1	chr11:75897369-75917576
<i>GSTA4</i>	Hs01119249_m1	chr6:52842751-52860176
<i>TFF1</i>	Hs00907239_m1	chr21:43782391-43786703
<i>PLAUR</i>	Hs00958880_m1	chr19:44152732-44174502
<i>TBP</i>	Hs00427620_m1	chr6:170863421-170881957
pri-miRNak		
hsa-mir-33a	Hs03293451_pri	chr22:42296948-42297016
hsa-mir-520b	Hs03295424_pri	chr19:54204481-54204541
hsa-mir-26b	Hs03302654_pri	chr2:219267369-219267445
hsa-let-7b	Hs03302548_pri	chr22:46509566-46509648
hsa-mir-655	Hs03304873_pri	chr14:101515887-101515983
hsa-mir-520e	Hs03303928_pri	chr19:54178965-54179051
hsa-mir-92a-1	Hs03302603_pri	chr13:92003568-92003645
hsa-mir-503	Hs03304160_pri	chrX:133680358-133680428
hsa-mir-17	Hs03295901_pri	chr13:92002859-92002942
miRNA helduak		
hsa-miR-26b	000407	chr2:219267369-219267445
hsa-miR-18a-3p	002423	chr13:92002859-92002942
<i>RNU48</i>	001006	chr6:31803040-31803103

*Exoi bakarria estaltzen dute.

2.1.2.1.2. Alderantzizko transkripzioa edo cDNAREN sintesia

RNA osoa cDNA bihurtu zen iScript cDNA Synthesis Kit-a (Bio-Rad, Hercules, CA, USA; kat. zk. 1708891) erabiliz. RNA (48 ng), 4 µl iScript Reaction Mix eta 1 µl alderantzizko transkriptasa nahastu ziren 20 µl-ko bolumenean eta erreakzioak ondorengo baldintzetan inkubatu ziren: 25°C-tan 5 minutuz, 42°C-tan 30 minutuz eta 85°C-tan 5 minutuz. cDNA laginak -20°C-tan gorde ziren erabili arte.

miRNA helduen kasuan, TaqMan microRNA Reverse Transcription Kit-a (Applied Biosystems, Waltham, MA; kat. zk. 4366596) eta TaqMan MicroRNA entseguak (miRNA-espezifikokoak) erabili ziren miRNA bakoitzaren cDNA espezifikoa sortzeko. RNA (300 ng), 6 µl RT hasleen nahastura, 0,3 µl dNTP (100 mM), 3 µl MultiScribe Reverse Transcriptase, 1,5 µl 10X RT Buffer eta 0,19 µl RNasa Inhibitzaile nahastu ziren 12 µl-ko bolumenean, eta ondorengo

balditzetan inkubatu ziren erreakzioak: 16°C-tan 30 minutuz, 42°C-tan 30 minutuz eta 85°C-tan 5 minutuz. cDNA laginak -20°C-tan gorde ziren erabili arte.

2.1.2.1.3. Adierazpen analisisia: qPCR

qPCR analisisetarako bi sistema erabili ziren: *Fluidigm BioMark dynamic array* (Fluidigm Corporation, San Francisco, CA, USA) eta *Eco Real-Time PCR* (Illumina, San Diego, CA, USA). Adierazpen analisiak bikoiztuta egin ziren lehenengo kasuan, eta hirukoiztuta bigarreanean. Kontrol endogenoen adierazpena (*TBP* TFs, pri-miRNA eta TFen itu geneen kasuan; eta *RNU48* miRNA helduen kasuan) aldi berean kuantifikatu zen, lagin bakoitzean hasieran jarritako RNA kantitatearen kontrol modura. Gene bakoitzaren adierazpen erlatiboa Ct zehatzaren metodoa erabilita kalkulatu zen, lehenago deskribatu moduan (Martín-Pagola et al., 2004). Adierazpen emaitzak gene endogenoarekiko eta kontrol laginen batz bestearrekiko normalizatu ziren, hurrengo formularen bidez:

$$\text{Adierazpen erlatiboa} = 2^{-\Delta\Delta Ct}$$

Non $\Delta\Delta Ct$ lagin esperimental eta kontrolen arteko ΔCt balioen diferentzia den, eta ΔCt gene hautagaien eta gene endogenoaren arteko Ct balioen diferentzia den.

2.1.2.1.3.1. Fluidigm BioMark dynamic array sistema

TF, pri-miRNA eta TFen itu geneen adierazpen azterketa Euskal Herriko Unibertsitateko (UPV/EHU) Gene Adierazpenen Unitatean burutu zen. Lehenik eta behin, gene bakoitzaren cDNAren preanplifikazioa burutu zen PreAmp Master Mix (Applied Biosystems; kat. zk. 4391128) erabiliz. Laburbilduz, 2,5 μ l PreAmp Master Mix, 1,25 μ l entseguen multzoa (0,2X) eta 1,25 μ l cDNA

nahastu ziren 5 µl-ko bolumenean. Ziklo-baldintza hauek erabili ziren: 95°C 10 minutuz, eta 14 ziklo 95°C-tan 15 segundoz eta 60°C-tan 4 minutuz. Preanplifikatutako cDNA 1:5 ratioan diluitu zen TE bufferrean. Ondoren, qPCRa burutu zen eta horretarako TaqMan Fast Advanced Master Mix (Applied Biosystems; kat. zk. 4444557), 2X Assay Loading Reagent (Fluidigm Corporation; kat. zk. 85000736) eta 20X GE Sample Loading Reagent (Fluidigm Corporation; kat. zk. 85000746) erabili ziren. Bost mikrolitro 10X Assay Premix (2,5 µl 20X TaqMan GE entsegu eta 2.5 µl 2X Assay Loading Reagent nahastura) eta Sample premix (2,5 µl 2X TaqMan Fast Advanced Master Mix, 0,25 µl 20X GE Sample Loading Reagent eta 2,25 µl anplifikatutako cDNA nahastura) erabili ziren putzuko. Erabilitako ziklo-baldintzak honako hauek izan ziren: 95°C 1 minutuz, eta PCRa (35 ziklo) 95°C 5 segundoz eta 60°C 20 segundoz.

2.1.2.1.3.2. *Eco Real-Time PCR* sistema

miRNA helduen kuantifikaziorako TaqMan MicroRNA entseguak erabili ziren. Bost mikrolitro TaqMan Universal PCR master mix (Thermo Fisher Scientific; kat. zk. 4440042), 4,5 µl H₂O, 0,5 µl TaqMan MicroRNA entsegu eta 0,7 µl RT produktu erabili ziren putzuko. Ziklo-baldintzak honako hauek izan ziren: polimerasaren aktibazioa 95°C-tan 10 min eta PCRa (40 ziklo) 95°C-tan 15 seg eta 60°C-tan 1 min.

2.1.2.2. TFen kokapen zelularra

TF hautagaien kokapena aztertu zen C2BBe1 zeluletan, *Material eta Metodoetako 1.3.* atalean azaldu bezala, 4 orduz PT-G edo PT-BSA-rekin inkubatutako kulturetan, hain zuzen ere.

2.1.2.2.1. Immunofluoreszentzia azterketak

C2BBe1 zelulak beirazko estalkietan hazi ziren eta PT-G edo PT-BSArekin tratatu ziren 4 orduz. Zelulak PBSrekin garbitu ziren eta %4ko formaldehido/PBSarekin finkatu. Ondoren, %0,5 Triton erabiliz permeabilizatu ziren, eta PBSarekin hiru garbiketa egin ziren. Zelulak Image-it FX signal enhancer (Invitrogen, Carlsbad, California, USA; kat. zk. A31629) erabiliz blokeatu ziren eta ondoren untxi-IgG antigorputz primarioekin inkubatu ziren: anti-ELK1 (1:500 diluzioa; kat. zk. ab32106), anti-NFκB p105/p50 (1:400; kat. zk. ab7971), anti-CREB1 (1:1000; kat. zk. ab31387) eta anti-IRF1 (1:1000; kat. zk. ab26109) (ABCAM, Cambridge, UK). Estalkiak PBSrekin garbitu ziren eta untxiaren aurkako ahuntz-IgG antigorputz sekundarioarekin inkubatu ziren (Invitrogen, kat. zk. A31627).

Ondoren, prestakinak PBSrekin garbitu eta lotura interzelularren markatzeilea den E-Cadherinaren kontrako sagu-antigorputz primarioarekin inkubatu ziren 1:100 diluzioan (BD Bioscience, San Jose, CA, USA, kat. zk. 610182). Azkenik, saguaren aurkako ahuntz-IgG antigorputz sekundarioarekin inkubatu ziren (Invitrogen, kat. zk. A31621). Prestakinak nukleoan tindatzailea den DAPI Dun VECTASHIELD Antifade Mounting Medium-ean (Vector Laboratories, Burlingame, California, USA; kat. zk. H-1200) muntatu ziren, eta Nikon Eclipse Ti fluoreszentzia mikroskopioan aztertu.

2.1.2.2.2. Nukleo eta zitoplasmako proteinen erauzketa

PT-G edo PT-BSArekin 4 orduz inkubatutako C2BBe1 zelulen nukleoko eta zitoplasmako proteinak Nuclear Extract (Active Motif, Carlsbad, CA, USA; kat. zk. 40010) kit komertzialarekin erauzi ziren. Laburbilduz, zelulak PBS/Fosfatasa inhibitzaileekin garbitu eta poliki-poliki arraspatuz plaketatik altxatu ziren. Zelulak 4°C-tan, 200 g-tan eta 5 minutuz zentrifugatu eta jaso ziren. Frakzio zitoplasmatikoa lortzeko, zelulak 500 µl 1X Hypotonic bufferean berreseki ziren

eta izotzetan inkubatu ziren 15 minutuz zelulak hanpatzeko. Gero, 25 µl detergente gehitu zen. Behin zelulak lisatuta, laginak 4°C-tan zentrifugatu ziren 14.000 g-tan eta 30 segundoz. Gainjalkina (frakzio zitoplasmatikoa) -80°C-tan gorde zen erabili arte, eta pikorra frakzio nuklearra lortzeko erabili zen. Pikorra 50 µl Complete Lysis bufferean berreseki zen eta 150 rpm-ko agitazioan mantendu zen 30 minutuz. Lagina 4°C-tan, 14.000 g-tan zentrifugatu zen 10 minutuz. Gainjalkina (frakzio nuklearra) -80°C-tan gorde zen erabili arte.

Proteinen kontzentrazioa Pierce BCA Protein Assay Kit-arekin (Thermo Fisher Scientific; kat. zk. 23227) neurtu zen. Hasteko, lanerako soluzioa prestatu zen A eta B errektiboak nahastuta (50:1). BSA anpoilaren edukia zenbait tututan diluitu zen, protokoloaren arabera diluzio serie estandarra lortzeko. Lanerako soluzioaren 20 µl gehitu zitzaion lagin zein estandar bakoitzari, eta 37°C-tan inkubatu ziren 30 minutuz. Behin tutuak hoztuta, absorbantzia neurtu zen 562 nm-tan espektrofotometroan, eta kurba estandarra marraztu zen BSA estandar bakoitzaren absorbantzia zegokion kontzentrazioarekin lotuta. Lagin bakoitzaren proteina kontzentrazioa kurba estandarraren arabera definitu zen.

2.1.2.2.3. Immunoblot analisia

Nukleoko eta zitoplasmako proteina laginak 95°C-tan inkubatu ziren 10 minutuz, %12 SDS-PAGE geletan kargatu, eta Trans-Blot Turbo Transfer (Bio-Rad) erabiliz nitrozelulosazko mintzetara transferitu ziren. Ondoren, mintza blokeatu zen %5 gantzik gabeko esne eta %0,05 Tween 20-dun Tris-salino (TBST) tanpoian, giro tenperaturan eta agitazioarekin, ordubetez. Mintzak 4°C-tan inkubatu ziren gau osoz soluzio berean diluitutako sagu-IgG antigorputz primarioekin: anti-HDAC1 (kontrol nuklearra) (1:10000 diluzioa; ABCAM, kat. zk. ab7028) eta anti- α -tubulin (kontrol zitoplasmikoa) (1:5000 diluzioa, Sigma-Aldrich, kat. zk. T9026); eta untxi-IgG antigorputz primarioekin: anti-ELK1 (1:500; kat. zk. ab32106), anti-NF κ B p105/p50 (1:400; kat. zk. ab7971), anti-CREB1 (1:1000; kat. zk. ab31387) eta anti-IRF1 (1:1000; kat. zk. ab26109).

Ondoren, ordubetez inkubatu ziren HRPrekin konjugatutako saguaren (1:1000) eta untxiaren (1:2000) aurkako ahuntz-IgG antigorputz sekundarioekin (Jackson ImmunoResearch Laboratories, Inc; West Grove, PA, USA, kat zk. 115-035-062 eta 111-035-003). Proteinak SuperSignal West Femto Maximum Sensitivity Substrate-arekin (Thermo Fisher Scientific; kat. zk. 34094) errebelatu eta ChemiDoc MP sisteman behatu ziren.

2.1.2.3. Kromatinaren immunoprezipitazioa

Kromatinaren immunoprezipitazio (ChIP) esperimentuak bi aldiz egin ziren Chromatin Shearing Optimization kit-Low SDS (Diagenode, Seraing, Belgium; kat. zk. AA-001-0100) erabiliz. Laburbilduz, C2BBel zelulak %1 formaldehidotan finkatu ziren, eta ondoren glizina gehitu zen finkapena eteteko. Zelulak zentrifugazioz jaso ziren eta iL1 lisis-bufferean (10 ml iL1 buffer 10 milioi zelulako) berreseki ziren. Zelulak 4°C-tan inkubatu ziren 10 minutuz, eta berriro ere zentrifugatu ziren. Gainjalkina bota eta pikorrari iL2 lisis-buffer gehitu zitzaion (10 ml 10 milioi zelulako). Zelulak 4°C-tan inkubatu ziren 10 minutuz, eta berriro ere zentrifugatu ziren. Gainjalkina bota zen eta azkenik, iS1 Complete Shearing bufferra (1 ml 10 milioi zelulako) gehitu zitzaion zelulei. Kromatina sonikazioz apurtu zen Bioruptor aparatuan (Diagenode; kat. zk. UCD-300 TM), intentsitate altuan, eta minutu bateko (30s on eta 30 s off) 5 ziklo egin ziren bi aldiz.

Immunoprezipitazioa modu automatizatuan burutu zen SX-8G IP-Star Compact (Diagenode; kat. zk. UH-002-0001) ekipoa, immunofluoreszentzian eta Western blotean erabilitako anti-CREB1 eta anti-IRF1 antigorputz berberekin (2,5 µg), untxi-IgG-arekin batera. ChIP erreakzioen aberastea qPCR bidez aztertu zen, hirukoiztuta, Mesa Green MasterMix (Thermo Fisher Scientific) erabiliz. Aukeratutako ituak ikertzeko, promotoreetan kokatutako hasle pareak diseinatu ziren, eta 5. taulan zerrendatu dira.

5. taula. ChIP-qPCR analisirako diseinatutako hasleen sekuentziak (5'-tik 3'-ra).

		Aurreranzko haslea	Alderantzizko haslea
IRF1 ituak	<i>CISD2</i>	GACGAAGTAGAGACAGCAAGAG	GGATACTGTGTGCGATGAGATAA
	<i>HDAC4</i>	CAGCCTTGCCTCACCTC	GACGAGCTCTTCATTAGAAACCA
	<i>WDR43</i>	GGTCACTTACGAGTATGGGAGA	AAGGCACGTACTCCTGGT
	<i>CXCL11</i>	CTCTTTGAGTCATGCACCTTTC	TCACAGTGCTTTCACATTCTTATC
	<i>BATF2</i>	GCCAAGTTTCAGTTTCTCCTAAAG	CGGAAGGCCAGTTCATGTTA
CREB1 ituak	<i>AKTIP</i>	CGGTCCTGCAAATCAAATCAC	CGATACTTCCATGACTGACAGG
	<i>NAMPT</i>	CGTTGCTTAAGTCACTGCTC	CCCTCTCTCCGTTTCCC
	<i>TPK1</i>	GGCAGCAGTCGCACTTA	GTCGATCGCCGTAGCTC
	<i>ISG15</i>	TCCCTGTCTTTCGGTCATTC	CTTCAGTTTCGGTTTCCCTTTC
	<i>HIST1H4C</i>	GGTCCGCCAAGTTTGTATTTAAG	CGACCAGACATGATTCCTATCG

2.1.2.4. miRNA ituen adierazpena EZn

Adierazpen aldaketak zituen pri-miRNA bakoitzari dagokion miRNA helduaren itu geneen bilaketa egin zen miRTarBase datu-basean, eta ebidentzia sendoa zutenak hartu genituen kontuan (*reporter assay* eta Western blot). Itu gene horien adierazpena, gure laborategian burututako heste frakzioetako RNA-seq datuetan aztertu zen.

2.1.3. Analisi estatistikoak

- Gene bikoteen gene-adierazpenaren arteko korrelazio maila kalkulatzeko eta emaitzak grafikoki adierazteko, corrgram R paketea erabili genuen (Friendly, 2002).
- Taldeen arteko adierazpen ezberdintasunak Wilcoxon (EZ aktiboa vs. GGD-EZ) eta Mann Whitney (kontrolak vs. bi gaixo taldeak) testak erabiliz aztertu ziren. Horretarako GraphPad Prism 5.0 software-a (GraphPad Software Inc, La Jolla, CA) erabili zen.

- Western blot eta ChIP-qPCR esperimentuetan alde bakarreko Student t-testak eta t-test pareatuak burutu ziren lagin taldeak konparatzeko, hurrenez hurren. Horretarako GraphPad Prism 5.0 software-a erabili zen.

2.2. TADak EZn

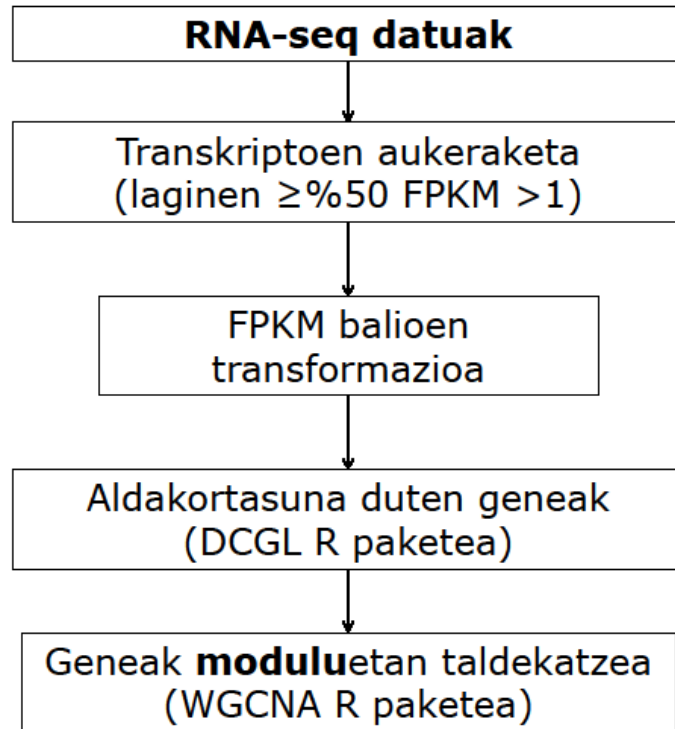
2.2.1. Kromatinaren 3D egitura aldaketak EZn

2.2.1.1. Koadierazpen analisia

Koadierazpen analisia <http://www.ncbi.nlm.nih.gov/sra/> orrialdetik, SRP077708 zenbakia zuen RNA-seq esperimentutik hartutako datuetan burutu zen. Zehazki, 10 EZdun pazienteren eta zeliakoak ez diren 12 kontrolen epitelioko zeluletan aberastutako frakzioen datuak erabili ziren.

Honetan ere WGCNA R paketea erabili zen. Kasu honetan, datuen ezaugarriak zirela eta, ezin izan zen transkripto guztietatik emaitzarik lortu. Laginen erdian FPKM (Fragments Per Kilobase Million) >1 zuten transkriptoak hartu ziren kontuan. FPKM balioak $\log_{10}(\text{FPKM}+1)$ balioetara transformatu ziren, paketearen egileek aholkatu moduan. Medianak baino adierazpen aldakortasun handiagoa zuten geneak soilik hautatu ziren koadierazpen analisietarako. Horretarako, DCGL R paketearen “*varianceBasedfilter*” funtzioa erabili zen eta lehenetsitako prozedura jarraitu zen. Analisi bakoitzean, “eskalarik gabeko topologia ereduaren doitze” analisiaren balio onenarekin “potentzia” parametroa ezarri zen.

Lan honetan erabilitako prozedura 8. irudian laburbildu da.



8. irudia. RNA-seq datuen analisirako erabilitako prozedura.

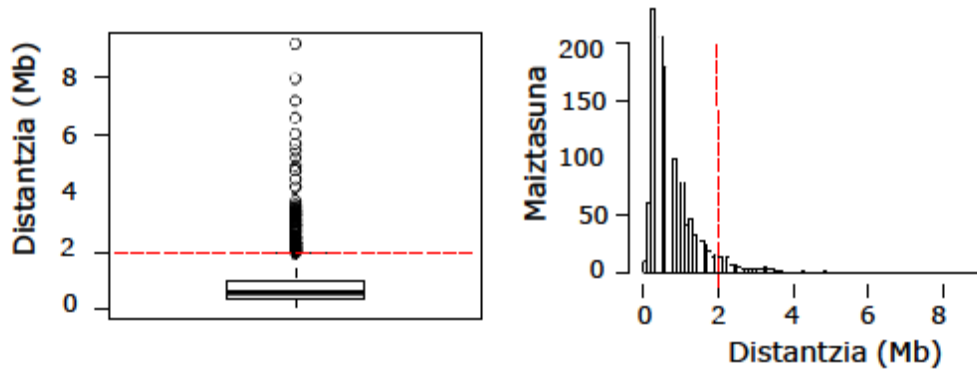
2.2.1.2. Ezaugarri genomikoen gainezarpena

BedTools programaren (Quinlan eta Hall, 2010) *Intersect* erreminta erabili zen kontserbatutako TADen koordinatuak (Dixon et al., 2012) eta RNA-seq datuetan identifikatutako koadierazpen modulu-tako geneen artean gainezarpinak aurkitzeko.

Egiaztatze esperimentalerako eskualde genomikoak ondorengo irizpideen arabera aukeratu ziren:

- TADen haustura: kontrolatan koadierazten ziren baina EZ aktiboan koadierazten ez ziren geneekin gainezarritako TADak.
- TADen fusioa: kontrolatan elkarren alboan dauden TADetan kokatzen diren eta koadierazten ez diren geneak, aldi berean EZ aktiboan koadierazten direnak. TADen arteko distantzia BedTools erabiliz

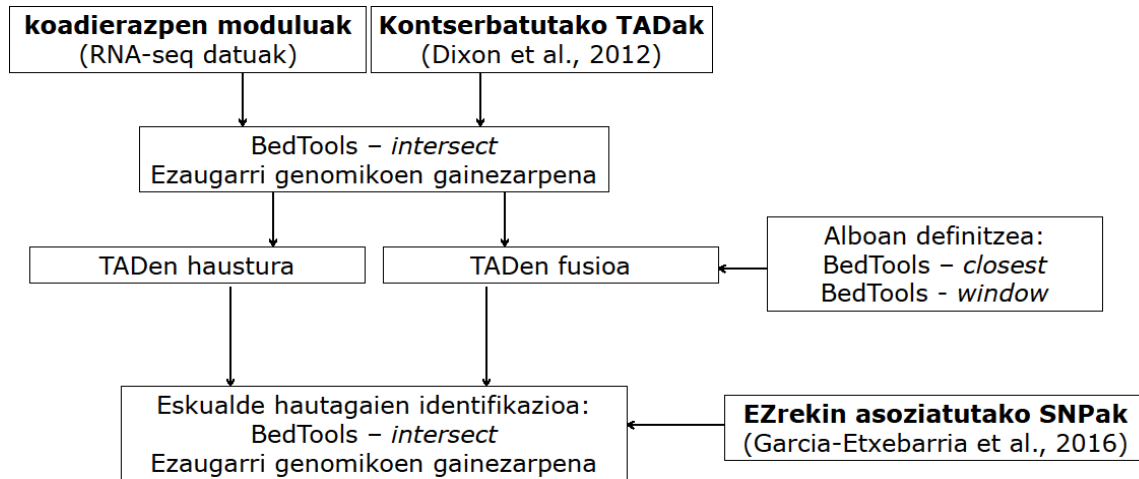
kalkulatu zen. Distantzia maximoa 2 Mb-tan ezarri zen, muturreko datuak kendu eta gero (9. irudia). Distantzia horretan oinarrituta, BedTools-en *window* metodoa erabili zen elkarren alboan zeuden TADak aurkitzeko.



9. irudia. Elkarren alboko TADen arteko distantzia kalkulatzeko erabilitako kaxa-diagrama eta histograma. TADen arteko distantzia ikus daiteke grafiko bietan. Marra gorriak muturreko datuak kendu eta geroko distantzia maximoa adierazten du, 2 Mb.

Azkenik, aukeratutako eskualdeen koordinatuak EZrekin asoziatutako SNPekin (Garcia-Etxebarria et al., 2016) (p balioa $< 0,05$) gainezarri ziren BedTools-en *intersect* metodoa erabilita. Irizpide bakoitzeko (haustura eta fusioa) hautagai bat aukeratu zen, hg19 chr22:29040000-29360000 eremuko *HSCB-XBPI* eskualdea eta hg19 chr20:33480000-34320000 eremuko *PROCR-ROMO1* eskualdea, hain zuzen.

Lan honetan erabilitako prozedura 10. irudian laburbildu da.



10. irudia. TADen analisirako erabilitako metodologia.

2.2.2. Hautagaien egiaztatze esperimentalak

2.2.2.1. Kromatinaren eskuragarritasuna

HCT116, HCT15 eta HEK293FT zelula kulturak HBSSrekin garbitu eta zelulak tripsinaz askatu ziren. Zelulak 1.500 rpm-tan zentrifugatu ziren 5 minutuz, eta 1 ml PBStan, 1 ml C1 lisis-bufferretan (1,28 M sukrosa, 40 mM Tris HCl, 20 mM MgCl₂, 4% Triton X-100) eta 3 ml H₂O-tan (+ proteinase inhibitor) berreseki ziren. Zelulak izotzetan inkubatu ziren 15 minutuz eta nukleoak zentrifugazioz jaso ziren, 4°C 15 min 2.500 rpm. 120.000 nukleo 100 µl garbiketa bufferretan (10 mM Tris, pH 7,4, 60 mM KCl, 15 mM NaCl, 5 mM MgCl₂ eta 300 mM sukrosa) berreseki ziren. DNasa (10 µl) (Macherey-Nagel; kat. zk. CAS 9003-98-9) eta 90 µl DNasa erreakzio buffer gehitu ziren lagin bakoitzeko. Beste lagin bat DNasa erreakzio bufferrean inkubatu zen, DNasarik gabe, kontrol moduan. Laginak 37°C-tan inkubatu ziren 20 minutuz, eta erreakzioak stop bufferrarekin eten ziren. *DNA Material eta Metodoetako 1.4.* atalean azaldu bezala erauzi zen.

Azkenik, laginak behar bezala digeritu zirela ziurtatzeko eta eskualde hautagaietan kromatinaren eskuragarritasuna aztertzeko, qPCRak egin ziren DNasarekin tratatutako laginetan eta digeritu gabeko kontroletan. Horretarako,

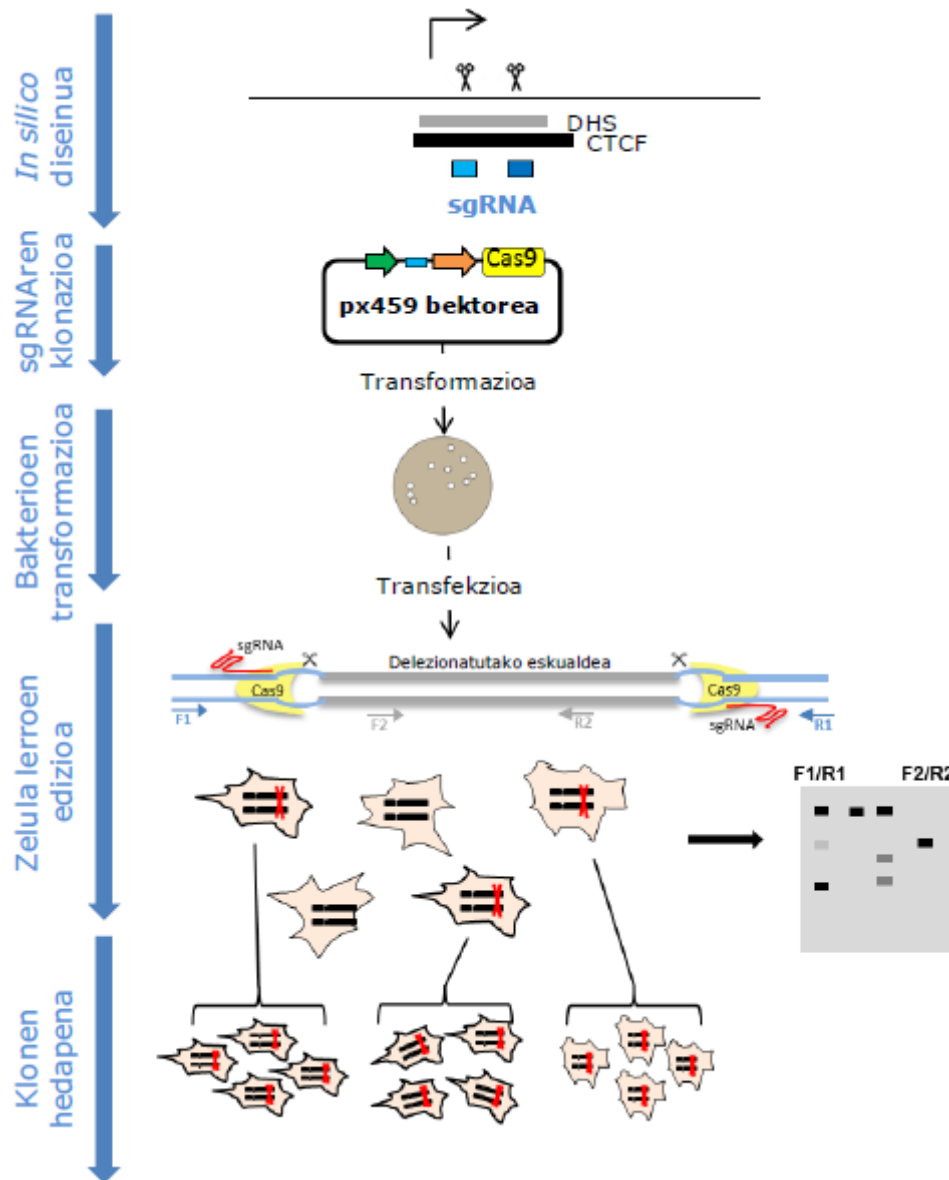
kromatinaren eremu ireki eta itxi konstitutiboak (kontrolak) eta eskualde hautagaiak anplifikatzeko hasleak erabili ziren. Hasleak Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) programarekin diseinatu ziren (6. taula).

6. taula. Kromatinaren eskuragarritasuna aztertzeko diseinatutako hasleen sekuentziak (5'-tik 3'-ra). *TBP* eta *GAPDH* kromatina itxi eta irekiaren kontrol modura erabili ziren, hurrenez hurren.

	Aurreranzko haslea	Alderantzizko haslea
<i>HSCB-XBP1</i> eskualdea	TCCCAAAGTGCTGGGATTAC	AATACTGCCACCCAGTGACC
<i>PROCR-ROMO1</i> eskualdea	ACTGTTCGCCCTTAAGTTCCTC	GCTTTTCCAGCCTCCTGTAG
<i>TBP</i> genea	TTGGCAGGCCTACAGTTTTTC	AAACTGGTCAGCCTTCTTGC
<i>GAPDH</i> promotorea	AAGGTGAAGGTCGGAGTCAAC	CCCATACGACTGCAAAGACC

2.2.2.2. DNasa I-ekiko hipersentikortasun eremuen ezabatzea

CTCF lotura-guneak zituzten bi DNasa I hipersentikortasun eremu (ingelesez *DNase I hypersensitive site* edo DHS) ezabatzeko CRISPR-Cas9 teknika erabili zen giza epitelio-zelula lerro desberdinetan, eta *HSCB-XBP1* eta *PROCR-ROMO1* eskualdeetako hg19 chr22:29186082-29186523 eta hg19 chr20:34026814-34027191 zati genomikoak delezionatu ziren, hurrenez hurren. Prozeduraren irudikapen eskematikoa 11. irudian ikus daiteke.



11. irudia. DHS ezabatzeko zelula epitelialetan erabilitako CRISPR-Cas9 prozeduraren irudi eskematikoa. sgRNAak (laukizuzen urdinak) eta genotipaketarako hasleak *in silico* diseinatu ziren. sgRNA gidak sgRNA gunea eta Cas9 dituen bektore batean klonatu ziren. Plasmidoa zeluletara transfektatu eta insertoaren barneratzea baieztatu zen. Azkenik, transfektatutako zelulak hedatu ziren. Ran et al., 2013 publikaziotik moldatua.

2.2.2.2.1. sgRNA diseinua

Esperokoa da DHS eremuen delezioak ondorio bortitzak sortzea. Zehazki, CTCF lotura-eremuak aukeratu ziren TADen mugetan eragiteko. DHS eremu

bakoitzetik 200 bp ur-gora eta ur-beherako eremuak hautatu ziren Design Tool³² (<http://crispr.mit.edu>) softwarearen bidez sgRNA espezifikoak diseinatzeko. Programak Cas9 bidez mozteko 20 bp-ko sekuentzia optimoak aukeratu zituen, eta sekuentzia horietarako sgRNA oligoak diseinatu ziren (7. taula).

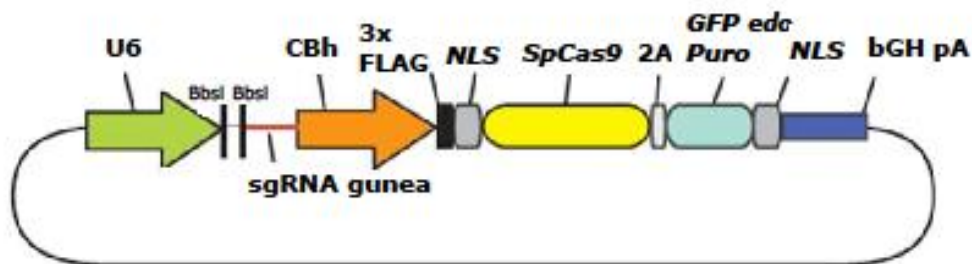
7. taula. CRISPR-Cas9 bidez DHSak ezabatzeke diseinatutako sgRNA oli-goak.

		Lokusa (hg19)	Sekuentzia	PAM*	Sekuentzia osagarria
HSCB-XBPI eskualdea	sgRNA1	chr22:-29186082	CACCGCGGAGTCTCGCTGTTTCGCC	AGG	AAACGGCGAAACAGCGAGACTCCGC
	sgRNA2	chr22:-29186523	CACCGCAGGAGAAATCGCCTGAACCC	AGG	AAACGGGTTTCAGGCCGATTCTCCTGC
PROCR-ROMOI eskualdea	sgRNA1	chr20:-34026814	CACCGTCAATGGCATTGGGGGTA	GGG	AAACTACCCCAAAATGCCATTGAC
	sgRNA2	chr20:-34027191	CACCGACAAACGTAATAAATCAGCC	AGG	AAACGGCTGATTTTTTACGTTTGTGTC

*Protospacer adjacent motif: ituetara lotzeko beharrezko sekuentzia.

2.2.2.2.2. sgRNA klonazioa

sgRNA osagarrien fosforilazio eta hibridaziorako oligo bakoitzaren 1 μ l (100 μ M), 1 μ l 10X T4 DNA ligazio buffer eta 0,5 μ l T4 PNK Polinukleotido kinasa (New England Biolabs, Ipswich, MA, USA; kat. zk. B0202S eta M0201L) nahastu (10 μ l guztira) eta ondorengo baldintzetan eduki zen: 37°C-tan 30 min, 95°C-tan 5 min eta -5°C/min 25°C-tara iritsi arte. Hibridazioa eta gero, sgRNA-duplexa px459 plasmidoan (Addgene, Cambridge, MA, USA) txertatu zen, plasmidoaren sgRNA lotzeko gunean (12. irudia). Intsertoak bektorean txertatu ahal izateko plasmidoa digeritu zen lehenbizi. Horretarako, 2 μ g bektore, 2 μ l *BbsI* entzima eta 2 μ l Fast Digest Green (10X) buffer (Thermo Fisher Scientific; kat. zk. FD1014 eta B72) nahastu ziren 20 μ l-ko bolumenean eta erreakzioa 37°C-tan inkubatu zen ordubetez. Digestioa elektroforesi bidez (150 V – 15 min, %1,5 agarosazko gela) egiaztatu zen. Ondoren, digeritutako plasmidoa geletik purifikatu zen Nucleospin Gel and PCR clean up kit-arekin (Macherey-Nagel; kat. zk. 740609.250). Digeritutako bektorea (50 ng) eta oligonukleotidoak (1 μ l) T4 DNA ligasa (New England Biolabs; kat. zk. M0202S) bidez lotu ziren ordubetez giro tenperaturan inkubatuta.



12. irudia. sgRNA gunea eta Cas9 dituen px459 plasmidoaren irudi eskematikoa. Oligoek plasmidoaren *BbsI* eremuetan lotzeko ertz osagarriak dituzte. *BbsI* entzimaren bidezko plasmidoaren digestioak oligoen txertaketa ahalbidetzen du. Ran et al., 2013 publikaziotik moldatua.

2.2.2.2.3. Transformazioa eta hautespena

Lortutako plasmido errekonbinantea *E. coli* DH5 α bakterio kompetenteetan (Thermo Fisher Scientific; kat. zk. 18265-017) transformatu zen bero kolpez, fabrikatzailearen protokoloa jarraituta. Transformatutako bakterioen hautespen positiborako, bakterioak 50 $\mu\text{g}/\mu\text{l}$ anpizilinadun LB agarrean erein eta gau osoan zehar 37°C-tan inkubatu ziren. Hazitako kolonietako batzuk aukeratu eta anpizilinadun LB medioan inkubatu ziren gau osoan zehar errekonbinanteak hazteko. Ondoren, plasmidoa erauzi zen bakterioetatik NucleoSpin Plasmid EasyPure (Macherey-Nagel; kat. zk. 740727) kitarekin. sgRNAren klonazioa baieztatzeko, 500 ng plasmido DNA digeritu zen 1 μl *BbsI* eta 0,5 μl *AgeI* entzimekin (Thermo Fisher Scientific; kat. zk. ER1461), 2 μl Fast Digest Green (10X) buffer zuen 20 μl -ko erreakzioan, 37°C-tan, ordubetez. Klonen aukeraketarako elektroforesia burutu zen.

2.2.2.2.4. Zelula lerroen edizioa

HCT116 eta HCT15 zelulak 600 ng plasmido errekonbinatzailearekin transfektatu ziren XTremHP DNA reagent (Invitrogen; kat. zk. 06366236001) erabiliz. HEK293FT zelulak 300 ng plasmido errekonbinatzailearekin transfektatu ziren eta Lipofectamine 2000 (Thermo Fisher Scientific; kat. zk. 11668-027) erabili zen kasu horretan. Transfekzioa 24 putzuko plaketan egin zen, 100.000 zelula putzuko dentsitatean. Gau osoko inkubazioaren ostean, transfektutako zelulen hautespen positiborako, zelulak 4 $\mu\text{g}/\text{ml}$ puromizinarekin inkubatu ziren 48 ordutan zehar. Delezio-mutazioa zuten zelulak identifikatzeko, transfektatutako zelulen parte bat lisis-bufferrarekin (50 mM Tris pH8, 50 mM KCl, 2,5 mM EDTA, 0,45% NP-40, 0,45% Tween 20 eta 10 μl 10 $\mu\text{g}/\mu\text{l}$ Proteinasa K) apurtu zen 2 ordutan eta 60°C-tan. Gero, laginak 95°C-tan inkubatu ziren 10 minutuz proteasak inaktibatzeke. Ondoren, delezionatutako eskualdearen alboetara kokatutako hasleekin PCR bidezko anplifikazioa burutu zen, eta delezio-mutazioa zuten zelulak gel-elektroforesian identifikatu ziren.

2.2.2.2.5. Klonen hedapena

Delezio-mutazioa zeramaten zelulen hautespenerako, zelulak diluitu eta 96 putzuko plaketan isolatu ziren dentsitate baxuan (bost putzuko zelula bat). Isolatutako zelulak hazten utzi ziren kolonia osasuntsuak sortu arte, eta aurreko atalean azaldu bezala, delezio-mutazio analisiak egin ziren mutazioa zuten klonak aukeratzeko.

2.2.2.3. Geneen adierazpen analisia

PROCR-ROMOI eskualdearen haustura aztertzeko, *PROCR* eta *ROMOI* geneen adierazpena eta koadierazpena neurtu zen editatutako zeluletan eta jatorrizko zelula epitelialetan. Adierazpen analisiak egiteko SYBR green detekzioa erabili zen, eta hasleak PrimerBank-en (<https://pga.mgh.harvard.edu/primerbank>) aukeratu ziren (8. taula).

RNA *Material eta Metodoetako 1.4.* atalean azaldu bezala erauzi zen editatutako eta jatorrizko zeluletatik, eta cDNA *Material eta Metodoetako 2.1.2.1.2.* atalean azaldu bezala sintetizatu zen.

qPCR analisiak bikoiztuta burutu ziren, eta *HPRT* gene endogenoaren adierazpena aldi berean kuantifikatu zen jarritako RNA kantitatearen kontrol modura. *Eco Real-Time PCR* sistema erabili zen eta gene bakoitzaren adierazpen erlatiboa *Material eta Metodoetako 2.1.2.1.3.* atalean deskribatu bezala kalkulatu zen. Laburbilduz, iTaq universal SYBR Green Supermix (Bio-Rad; kat. zk. 172-5121) erabili zen eta erreakzioak honako baldintzetan inkubatu ziren: lehenik 95°C-tan 30 segundoz, gero PCRa (40 ziklo) 95°C-tan 15 segundoz eta 60°C-tan 1 minutuz, eta ondoren 95°C-tan 15 segundoz, 55°C-tan 15 segundoz eta 95°C-tan 15 segundoz disoziazio-kurben analisia burutzeko.

8. taula. *PROCR* eta *ROMOI* geneen adierazpen analisirako diseinatutako hasleak (5'-tik 3'-ra). *HPRT* kontrol endogeno modura erabili zen.

Genea	PrimerBank ID	Amplikoiaaren luzera (bp)	Aurreranzko haslea	Alderantzizko haslea
<i>PROCR</i>	34335271c1	80	CCTACAACCCGCACTCGGTATG	CGCGGAAAATATGTTTCTGCACA
<i>ROMOI</i>	115430214c1	107	AAGCTGCTTCGACCCGTGTC	CCC GCATTCCCGATCCTTGAG
<i>HPRT</i>	164518913c2	190	ACCAGTCAACAGGGGACATAA	CTTCGTGGGGTCCCTTTTCACC

2.2.2.4. Zelula lerroen karakterizazioa

HCT116, HCT15 eta HEK293FT zelula lerroetan rs6060369, rs224371 eta rs2104417 SNPen genotipazioa burutu zen. SNP hauek *PROCR-ROMO1* eskualdeko bi TADen artean kokatzen diren EZri asoziatutako SNP haplotipoen ordezkari direlako hautatu ziren.

SNPak aurretik diseinatutako rhAmp SNP entseguen (Integrated DNA Technologies, Coralville, IA, USA) genotipatu ziren (9. taula). Aleloen arteko diskriminazioa alelo-espezifikoak ziren bi aurreranzko hasleen lehiakortasunaren bidez lortu zen. Hasle bat FAM-ekin markatuta zegoen, eta bestea Yakima Yellow-rekin (YY) .

9. taula. Genotipazio esperimentuan erabilitako entseguak.

SNPa	Entseguaren izena	Kokapena (hg19)
rs6060369	Hs.GT.rs6060369.C.1	chr20: 33907160
rs224371	Hs.GT.rs224371.A.1	chr20: 34074830
rs2104417	Hs.GT.rs2104417.A.1	chr20: 34127870

DNAREN amplifikazioa *Eco Real-Time PCR* sisteman burutu zen 10 µl-ko erreazioetan. Erreakzioan 40 ng DNA, rhAmp Genotyping Mix (rhAmp Genotyping Master Mix eta rhAmp Reporter Mix w/Reference) (Integrated DNA Technologies; kat. zk. 1076015 eta 1076021) eta rhAmp SNP entseguak (Integrated DNA Technologies) erabili ziren. Zikloak ondorengo baldintzetan burutu ziren: 95°C-tan 10 minutuz, eta 45 ziklo 95°C-tan 10 segundoz, 60°C-tan 30 segundoz eta 68°C-tan 20 segundoz.

2.2.3. Analisi estatistikoak

Kromatinaren eskuragarritasuna digeritutako eta digeritu gabeko kromatinaren Ct balioen arteko diferentziaren arabera kuantifikatu zen, eta eskualde hautagaiak

kromatina itxiaren kontrolarekin (*TBP*) konparatu ziren alde bakarreko Student t-testak erabiliz. GraphPad Prism 5.0 softwarea erabili zen.

Zelula kulturen arteko adierazpen eta koadierazpen ezberdintasunak Mann Whitney testaren eta Spearman korrelazio testaren bidez burutu ziren, hurrenez hurren. Horretarako GraphPad Prism 5.0 softwarea erabili zen.

2.3. Metilazio patroien aldaketa akutuak EZn

DNAren metilazioa *Material eta Metodoetako 1.3.* atalean deskribatu bezala, gliadinarekin eta gliadina gabe inkubatutako 8 kontrolen eta 7 GGD-EZ pazienteren biopsia laginetan aztertu zen. DNA *Material eta Metodoetako 1.4.* atalean azaldu bezala erauzi zen.

2.3.1. Bisulfito bidezko konbertsioa

Metilatu gabeko zitosinak urazilo bihurtzeko EZ DNA Methylation-Lightning Kit (Zymo Research, Irvine, CA, USA; kat. zk. D5030) erabili zen. Laburbilduz, 130 µl Lightning Conversion Reagent gehitu zitzaion 20 µl DNA lagin bakoitzari (300 ng), eta ondorengo baldintzetan inkubatu zen: 98°C-tan 8 minutuz eta 54°C-tan 60 minutuz. Laginak M-Binding bufferdun Zymo-Spin IC zutabeetan kargatu ziren, eta zentrifugatu ondoren, garbiketa bufferra gehitu zen zutabeetara. Ondoren, L-Desulphonation bufferra gehitu zen eta 15 minutuz inkubatu zen giro tenperaturan. Konbertitutako DNA 20 µl M-Elution bufferretan eluitu zen.

2.3.2. Hautatutako eskualdeen amplifikazioa, kuantifikazioa, purifikazioa eta normalizazioa

Sekuentziatu aurretik, bisulfitoarekin konbertitutako DNAREN PCR amplifikazioa burutu zen eskualde hautagaietan. Eskualde horiek ondorengo irizpideen arabera aukeratu ziren:

- Gure taldeak egindako lan batean (Scientific Reports, errebisiopean) identifikatu ziren metilazio aldaketak zituzten eskualdeak (ingelesez *differentially methylated regions* edo DMRs) eta *microarray* esperimentu akutuan 4 orduko gliadinaren kitzikapenaren ondorioz adierazpen aldaketa erakusten zutenak: *HDAC4* genearen gorputza hartzen duen hg19 chr2:240271171-240271276 eskualdea.
- Lan berdinean frakzio epitelialean identifikatutako DMR nagusiak: *TRIM15* genearen 5'UTRa, lehenengo exoia, eta genearen gorputza hartzen dituen hg19 chr6:30131458-30132471 eskualdea; *TAP1* genearen promotorea hartzen duen hg19 chr6:32819858-32820249 eskualdea eta *HLA-B* genearen promotorea hartzen duen hg19 chr6:31322766-31323506 eskualdea.
- *Microarray* esperimentu akutuan 4 orduko gliadinaren kitzikapenaren ondorioz adierazpen aldaketa handienak erakusten zituzten geneak eta literaturaren arabera diferentzialki metilatutako posizioak zituztenak: *SLC46A1* genearen promotorea hartzen duen hg19 chr17:26732864-26733385 eskualdea (Diop-Bove et al., 2009), eta *DFNA5* genearen promotorea hartzen duen hg19 chr7:24796542-24797487 eskualdea (Kim et al., 2008).

Hasleak MethPrimer 2.0 plataforman (<http://www.urogene.org/methprimer2/>) diseinatu ziren: horrek hautatutako sekuentziaren *in-silico* bisulfito konbertsioa egiten du. *SLC46A1* eta *DFNA5* geneen promotoreen kasuan, hasleak eskuz diseinatu ziren; izan ere, plataforma ez zen bi eskualde horietarako hasle egokiak

diseinatzeko gai izan. *GAPDH* hasle estandarrekin eta bisulfitoarekin tratatutako DNarentzako hasle espezifikoekin amplifikatu zen konbertsioaren efizientzia ebaluatzeko (10. taula).

Anplifikazioak PyroMark PCR Kit (Qiagen, Hilden, Germany; kat. zk. 978703 VF 40) erabiliz egin ziren, 40 µl-ko bolumen finalean 15 ng bisulfitoarekin konbertitutako DNA, 20 µl PyroMark PCR Master Mix 2X, 4 µl CoralLoad Concentrate 10X, 2,4 µl MgCl₂ (25 mM), eta hasle bakoitzetik (10 µM) 0,8 µl nahastuta. Erabilitako ziklo-baldintzak hurrengo hauek izan ziren: 95°C-tan 15 minutu, eta 50 ziklo 94°C-tan 30 segundo, 56°C-tan 30 segundo eta 72°C-tan 30 segundo. Azkenik, 72°C-tan 10 minutuz inkubatu ziren erreakzioak.

Anplifikazio produktuak elektroforesi bidez aztertu ziren, eta xingolen intentsitatea Image Lab v5.2.1. programarekin neurtu zen. Indibiduo beraren PCR produktuak kantitate berdinean bateratu ziren, lagin bakoitzaren aplikoi multzo orekatua lortzeko. Sekuentziatu aurretik, aplikoi multzoak NucleoSpin Gel and PCR Clean-up (Macherey-Nagel; kat. zk. 740609.250) erabiliz purifikatu ziren.

Azkenik, multzo bakoitzaren DNA kontzentrazioa fluoreszentzia bidez (QuBit) neurtu zen eta 0,2 ng/µl-tara normalizatu zen. DNA liburutegiak Nextera XT kitarekin (Illumina) eratu ziren, eta Illumina MiSeq sisteman sekuentziatu ziren Miseq Reagent kit v3 (600 ziklo, 25M *read*) (Illumina) erabiliz, Euskal Herriko Unibertsitateko (UPV/EHU) Sekuentziazio eta Genotipo Azterketen Unitatean.

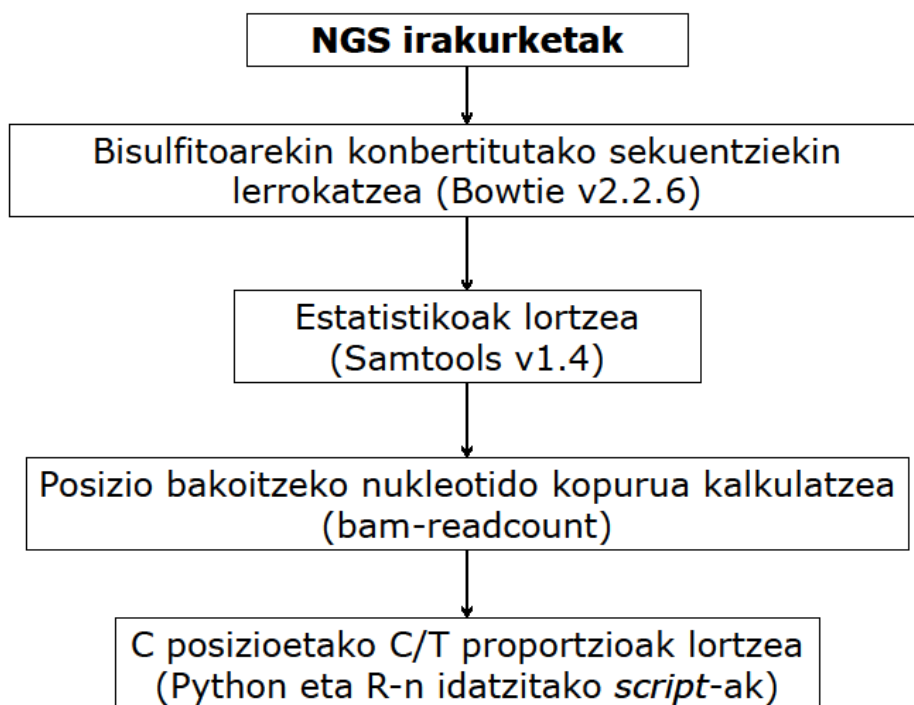
10. taula. Aukeratutako eskualdeen amplifikaziorako hasleak (5'-tik 3'-ra).

	Aurreranzko haslea	Alderantzizko haslea	Anplikoia- luzera (bp)
<i>HDAC4</i>	AATTTAATAAATGATGAAAAAGTAGGA	AAAACAAAAACCCCTTATACCCAA	293
<i>TRIM15</i>	TGTTAAGAGGAGGAGTAGGTAGAGATTT	CCTATCTACCTTCAATCTAAAAATACC	362
<i>TAPI</i>	TAGGGAATAGATTGAAGGTTTTAGG	CAATCTAACTAAAACTAACCTACTTAAACT	290
<i>HLA-B</i>	AAATTTTTAGTGGGATAAGAAAAAT	CCAAAAATAAACAACTATAATAATACCTTC	364
<i>SLC46A1</i>	TTGTAGGATTAAGGTAAAGTTGG	CACITTTACAAAATAAAAATCATCCC	368
<i>DFNA5</i>	AGGGTGGTTTAGAGAGAAA	CTCTCTAAAACCTTCTAAAAAATC	346
<i>GAPDH</i> modifikatu gabea	CTCTTGCTACTCTGCTCTGG	GCTAAGTTTAGCCCTGCCTGG	189
<i>GAPDH</i> modifikatua	GTATTTGTTGATGGGTTAAGG	ATAAAAAACAAATCCCCTACCC	150

2.3.3. Metilazio azterketa belaunaldi berriko sekuentzioaziaz erabiliz

Belaunaldi berriko sekuentzioazioan (ingelesez *Next-generation sequencing* edo NGS) lortutako irakurketak Meth Primer 2.0 plataforman lortutako bisulfitoarekin konbertitutako sekuentziekin lerrokatu ziren Bowtie v2.2.6 (Langmead eta Salzberg, 2012) erabiliz. Gero, Samtools v1.4 (Li et al., 2009) erabili zen lerrokatzearen arrakastaren estatistikoak lortzeko, eta bam-readcount (<https://github.com/genome/bam-readcount>) posizio bakoitzeko nukleotidoen kopurua kalkulatzeko. CpGen eta ez-CpGen C posizioetako C/T proportzioak lortzeko, Python eta R-n idatzitako *script*-ak erabili ziren. Ez-CpGen C posizioetan C/T proportzioa %5 baino txikiagoa zuten aplikioiak besterik ez ziren kontuan hartu analisisetarako.

Lan honetan erabilitako prozedura 13. irudian laburbildu da.



13. irudia. Bisulfito NGS analisisirako prozedura.

2.3.4. Analisi estatistikoak

Gliadinarik gabe inkubatutako biopsien metilazio mailak, GGD-EZ *vs.* kontrolen artean, zein gliadinarekin *vs.* gliadinarik gabe inkubatutako kontrolen biopsien arteko metilazio aldaketak, eta gliadinarekin inkubatutako *vs.* gliadinarik gabe inkubatutako GGD-EZen biopsien arteko metilazio aldaketak Mann Whitney testen bidez aztertu ziren R-n.

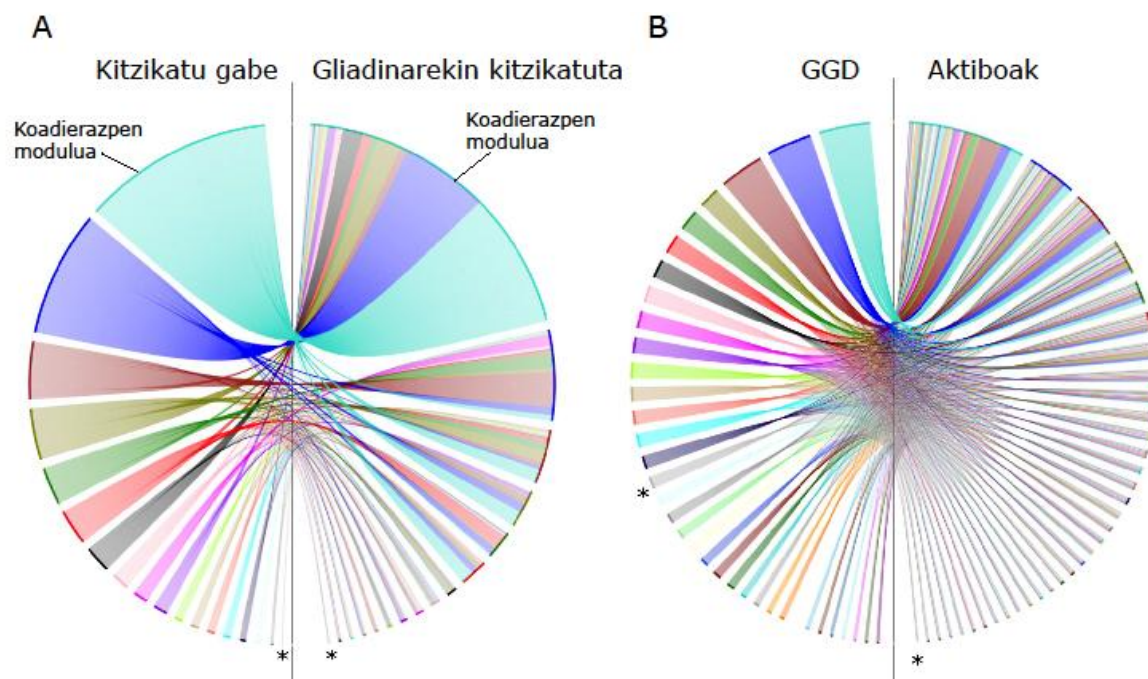
Emaizak

1. Genoma osoko koadierazpena EZn

Koadierazpen analisien bidez, elementu erregulatzaile komunak partekatzen dituzten kotranskribatutako gene-taldeak identifikatu ziren.

1.1. Gliadinaren kitzikapenak eragindako koadierazpen aldaketak EZn

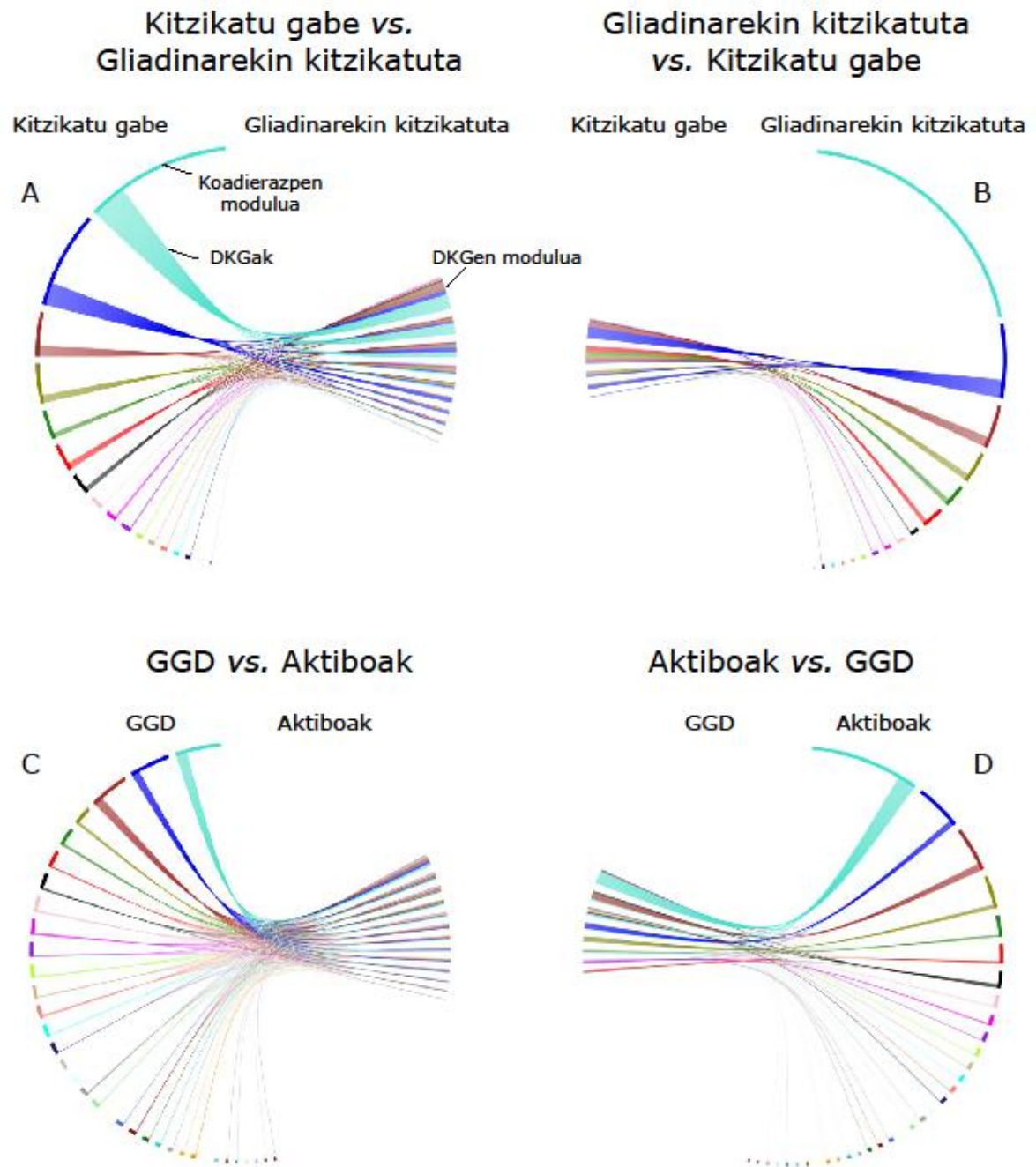
Adierazpen *array*-eko 19.851 gene proteina-kodetzaileen artean, guztien medianak baino aldakortasun handiagoa zuten geneak erabili ziren koadierazpen moduluak eratzeko, esperimentu bakoitzean. Esperimentu akutuan, 6.863 gene 18 modulutan taldekatu ziren kitzikatu gabeko biopsietan (1. eranskina), eta 16 modulutan gliadinarekin kitzikatutako biopsietan (14A irudia) (2. eranskina). Epe-luzeko esperimentuan, 8.342 gene 35 modulutan taldekatu ziren GGD-EZ laginetan (3. eranskina), eta beste 35 modulutan lagin aktiboen kasuan (14B irudia) (4. eranskina).



14. irudia. Genoma osoko koadierazpena A) esperimentu akutuan eta B) epe-luzeko esperimentuan. Koadierazpen moduluak irudikatu dira kitzikatu gabeko laginetan, gliadinarekin kitzikatutako laginetan, GGD laginetan eta lagin aktiboetan. Kolore bakoitzak koadierazpen

modulu bat irudikatzen du, eta izartxoak koadierazpen moduluetan taldekatu ez diren geneak seinalatzen ditu.

Ondoren, esperimentu bakoitzean diferentzialki koadierazitako geneen edo DKGen identifikazioari ekin genion (15. irudia). Esperimentu akutuan, kitzikatu gabeko eta gliadinarekin kitzikatutako biopsietan, geneen %21,71 eta %9,93an koadierazpena aldatuta zegoen, hurrenez hurren, talde bakoitza bere osagarriarekin konparatzean (5 eta 6. eranskinak). Epe-luzeko esperimentuan, GGD laginetan eta lagin aktiboetan geneen %14,68 eta %12,60an koadierazpena aldatuta zegoen hurrenez hurren, talde bakoitza bere osagarriarekin konparatzean (7 eta 8. eranskinak). Gainera, DKGek ere koadierazpen moduluak eratzen zituztela ikusi genuen (15. irudia, 5-8. eranskinak).



15. irudia. DKGak A,B) esperimentu akutuan eta C,D) epe-luzeko esperimentuan. Koadierazpen moduluak eta DKGak (ezkerretara A eta C-n, eskuinetara B eta D-n), eta DKGek eraturako koadierazpen modulu berriak (eskuinetara A eta C-n, ezkerretara B eta D-n) irudikatu dira konparaketa bakoitzeko. Termino bakoitzaren adibide bat erakutsi da irudian.

1.2. Koadierazpen aldaketetan eragina duten elementu erregulatzailen identifikazioa

DKGak eta DKGen moduluak sakonago aztertu ziren koadierazpen aldaketekin erlazionatuta egon zitezkeen elementu erregulatzaille posibleen aberastea zegoen argitzeko.

TFei dagokienez, esperimentu akutuan hiru modulu aurkitu ziren TFBSetan aberastuta (11. taula): kitzikatu gabeko biopsietan *Magenta* moduluko DKGak HOXA5en lotura-guneetan aberastuta zeuden (5. eranskina); gliadinarekin kitzikatutako biopsietan *Purple* moduluko DKGak HOXA5, IRF1 eta NFkB1en lotura-guneetan zeuden aberastuta (6. eranskina); eta DKGen *Yellow* modulua HOXA5en lotura-guneetan zegoen aberastuta gliadinarekin inkubatutako biopsien kasuan (6. eranskina). Epe-luzeko esperimentuan (11. taula) GGD-EZ laginetan, *Brown* moduluko DKGak GFI1, FOXI1, ELK1, CREB1, GABPA, HOXA5, MYF eta RORA TFen lotura-guneetan zeuden aberastuta (7. eranskina).

miRNAen kasuan, esperimentu akutuan hiru moduluk erakutsi zuten miRNA mailako aberastea (12. taula): kitzikatu gabeko biopsietan *Turquoise* moduluko DKGek (5. eranskina) eta DKGen *Brown* eta *Turquoise* moduluek (5. eranskina). Epe-luzeko esperimentuan, bederatzi moduluk erakutsi zuten miRNA mailako aberastea (12. taula): GGD laginetan *Black* eta *Brown* moduluetakoko DKGek (7. eranskina) eta DKGen *Green* eta *Red* moduluek (7. eranskina); eta lagin aktiboetan *Black*, *Blue* eta *Pink* moduluetakoko DKGek (8. eranskina), eta DKGen *Brown* eta *Turquoise* moduluek (8. eranskina).

11. taula. DKGetan eta DKGen moduluetan eginiko TF aberaste-analisiaren emaitza esangarrien laburpena, esperimentu eta baldintza bakoitzean ($P < 0,05$).

Esperimentua	Konparaketa	Behaketa	Modulua	TF	Funtzioa (OMIM edota RefSeq-en arabera)
Akutua	Kitzikatu gabe vs. Gliadinarekin kitzikatuta	DKGak	Magenta	HOXA5	Tumorigenesia eta p53-ren erregulazioa.
		DKGak	Purple	HOXA5 IRF1 NFKB1	Tumorigenesia eta p53-ren erregulazioa. α , β eta γ interferoiek eragindako geneen aktibazioa. Apoptosia eta tumore supresioa. Gene proinflamatorioen erregulazioa.
	Gliadinarekin kitzikatuta vs. Kitzikatu gabe	DGGen moduluak	Yellow	HOXA5	Tumorigenesia eta p53-ren erregulazioa.
Epe-luzea	GGD vs. Aktiboak	DKGak	Brown	GF11	Histona modifikazioen kontrola, CD4 T zelulen proliferazioa.
				FOX11	Koklea eta bestibuluaren garapena eta enbriogenesia.
				ELK1	Muskulu lisoaren diferentziazioa eta proliferazioa. ras-raf-MAPK bidezidorrerako itu nuklearra.
				CREB1	cAMP bidezidorreko geneen transkripzioaren indukzioa homonen estimulazioaren aurrean.
				GABPA	Zitokromo oxidasaren adierazpenaren erregulazioa eta mitokondriaren funtzioaren kontrol nuklearra.
HOXA5	Tumorigenesia eta p53-ren erregulazioa.				
MYF	Muskuluen garapena.				
RORA	Hormona erantzunaren elementuen lotura eta gene adierazpena.				

12. taula. DKGetan eta DKGen moduluetan eginiko miRNA aberaste-analisiaren emaitza esangarrien laburpena, esperimentu eta baldintza bakoitzean ($P < 0,05$).

Esperimentua	Konparaketa	Behaketa	Modulua	miRNA
Akutua	Kitzikatu gabe vs. Gliadinarekin kitzikatuta	DKGak	Turquoise	hsa-miR-369-3p, hsa-miR-937, hsa-miR-655, hsa-miR-324-3p, hsa-miR-891b, hsa-miR-323-5p, hsa-miR-18a-3p
		DKGen moduluak	Brown	hsa-miR-662
	GGD vs. Aktiboak	DKGak	Turquoise	hsa-miR-520h
			Black	hsa-miR-33a
Epe-luzea	Aktiboak vs. GGD	DKGak	Brown	hsa-miR-92a, hsa-let-7b-3p, hsa-miR-507
			Green	hsa-miR-551b-3p
		DKGen moduluak	Red	hsa-miR-15b, hsa-miR-92a-2-3p, hsa-miR-663, hsa-miR-886-5p, hsa-miR-933, hsa-miR-15a, hsa-miR-16, hsa-miR-195, hsa-miR-615-5p, hsa-miR-650, hsa-miR-922, hsa-miR-503, hsa-miR-518c-3p, hsa-miR-423-3p
		Black	hsa-miR-26a, hsa-miR-26b	
	DKGak	Blue	hsa-let-7f, hsa-miR-613, hsa-miR-520b, hsa-miR-520c-3p, hsa-miR-518b, hsa-miR-518c, hsa-miR-518d-3p	
		Pink	hsa-miR-886-5p, hsa-miR-486-5p	
		DKGen moduluak	Brown	hsa-miR-302b, hsa-miR-302c, hsa-miR-518c, hsa-miR-520a-3p, hsa-miR-520b, hsa-miR-520c-3p, hsa-miR-520d-3p, hsa-miR-520e, hsa-miR-134, hsa-miR-933
			Turquoise	hsa-miR-674

1.3. Erregulatzaille hautagaien eta TFen itu geneen aukeraketa ondorengo analisietarako

Balioztatze esperimentalerako erregulatzaille hautagaien aukeraketa terminoen errepikakortasunaren eta literaturaren arabera egin zen *Material eta Metodoetan* azaldu bezala.

Identifikatutako hautagai batzuk aurretik inflamazioarekin erlazionatu izan dira, IRF1 (Guo et al., 2010) eta CREB1 (Wen et al., 2010) esaterako. Beste batzuk EZrekin asoziatutako bidezidorrarekin erlazionatu dira, hala nola, Toll-like hartzaileen bidezidorrarekin, hsa-miR-92a (Lai et al., 2013) eta hsa-let-7b-3p (Teng et al., 2013), adibidez; NFκB bidezidorrarekin, hsa-miR-33a (Kuo et al., 2013), hsa-miR-503 (Zhou et al., 2013), hsa-miR-18a-3p (Trenkmann et al., 2013), hsa-miR-26b (Zhao et al., 2014), hsa-miR-520e (Zhang et al., 2012) eta NFKB1 esaterako (Fernandez-Jimenez et al., 2014); eta lotura hertsiekin, ELK1 kasu (Al-Sadi et al., 2013). Beste hautagai batzuk, hsa-miR-655 eta hsa-miR-520b hain zuzen, *MAGI2* (Kitamura et al., 2014) eta *MICA* (Yadav et al., 2008) erregulatzen dituzte hurrenez hurren, EZ jasateko arriskuarekin asoziatuta dauden bi gene (Martin-Pagola et al., 2003; Wapenaar et al., 2016). Bestalde, HOXA5en itu geneen aberastea ikusi zen esperimentu akutuko zein epe-luzeko esperimentuko zenbait modulutan.

TFen itu geneen aukeraketari dagokionez, ondorengo itu geneak aukeratu ziren *Material eta Metodoetan* definitutako irizpideen arabera (13. taula):

13. taula. Ondongo esperimentuetarako aukeratutako TFen itu geneak, azaldutako irizpideen arabera.

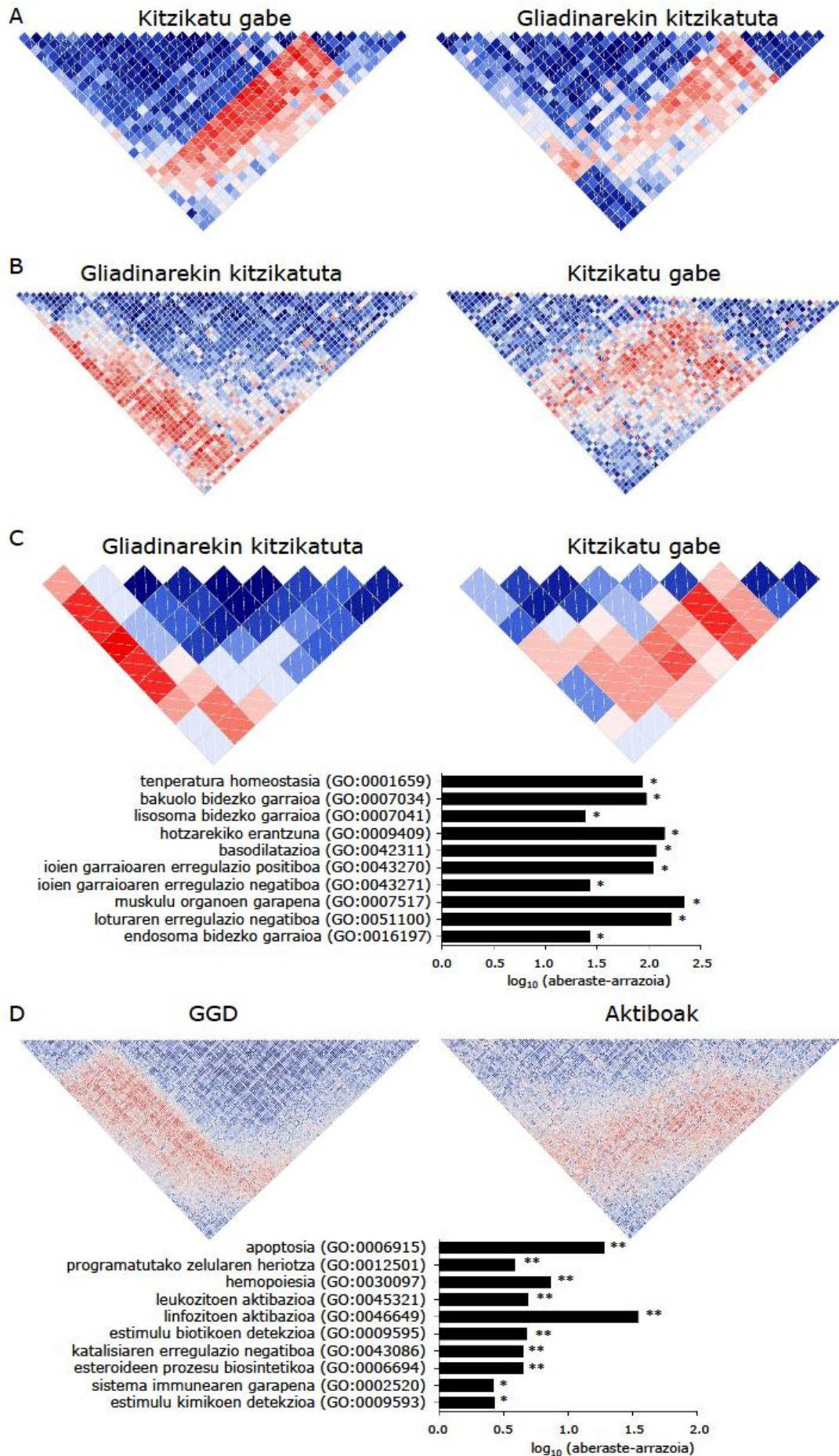
	DKG multzoetako itu geneak				DKG multzoetan ez dauden itu geneak
	Kontserbatuta ko lotura-guneak	ChIP-seq	H3K27Ac	Literatura	
IRF1 ituak	<i>CISD2</i>	x	x		Espezifikotasuna*
	<i>HDAC4</i>	x	x		
	<i>WDR43</i>	x	x		
	<i>CXCL11</i>				9
	<i>BATF2</i>				11
CREB1 ituak	<i>AKTIP</i>		x	Beste gaixotasun batzuk (Liang et al., 2012)	
	<i>NAMPT</i>		x	Inflamazioa (Garten et al., 2011)	
	<i>TPK1</i>		x	Beste gaixotasun batzuk (Banka et al., 2014)	
	<i>ISG15</i>				
	<i>HIST1H4C</i>				4
ELK1 ituak	<i>CRTAM</i>		x	T zelulen diferentziazioa (Takeuchi et al., 2016)	
	<i>PLLP</i>	x	x	Epitelio zelulen diferentziazioa (Rodríguez-Fraticelli et al., 2015)	
	<i>RFX5</i>	x	x	MHCII geneen adierazpenaren aktibazioa (Garvie et al., 2007)	
	<i>NKG7</i>				
	<i>RAB17</i>				
	<i>WNT11</i>		x		10
NFKB1 ituak	<i>GSTA4</i>				7
	<i>TFPI</i>				
	<i>PLAUR</i>				
					10

*Itu gene bakoitza erregulatzen duen TF kopurua. TF gutxienean erregulaziopean zeuden ituak espezifikogoztat jo eta aukeratu ziren.

1.4. Transkripzio faktoreak

1.4.1. Moduluetakoko funtzio biologikoak

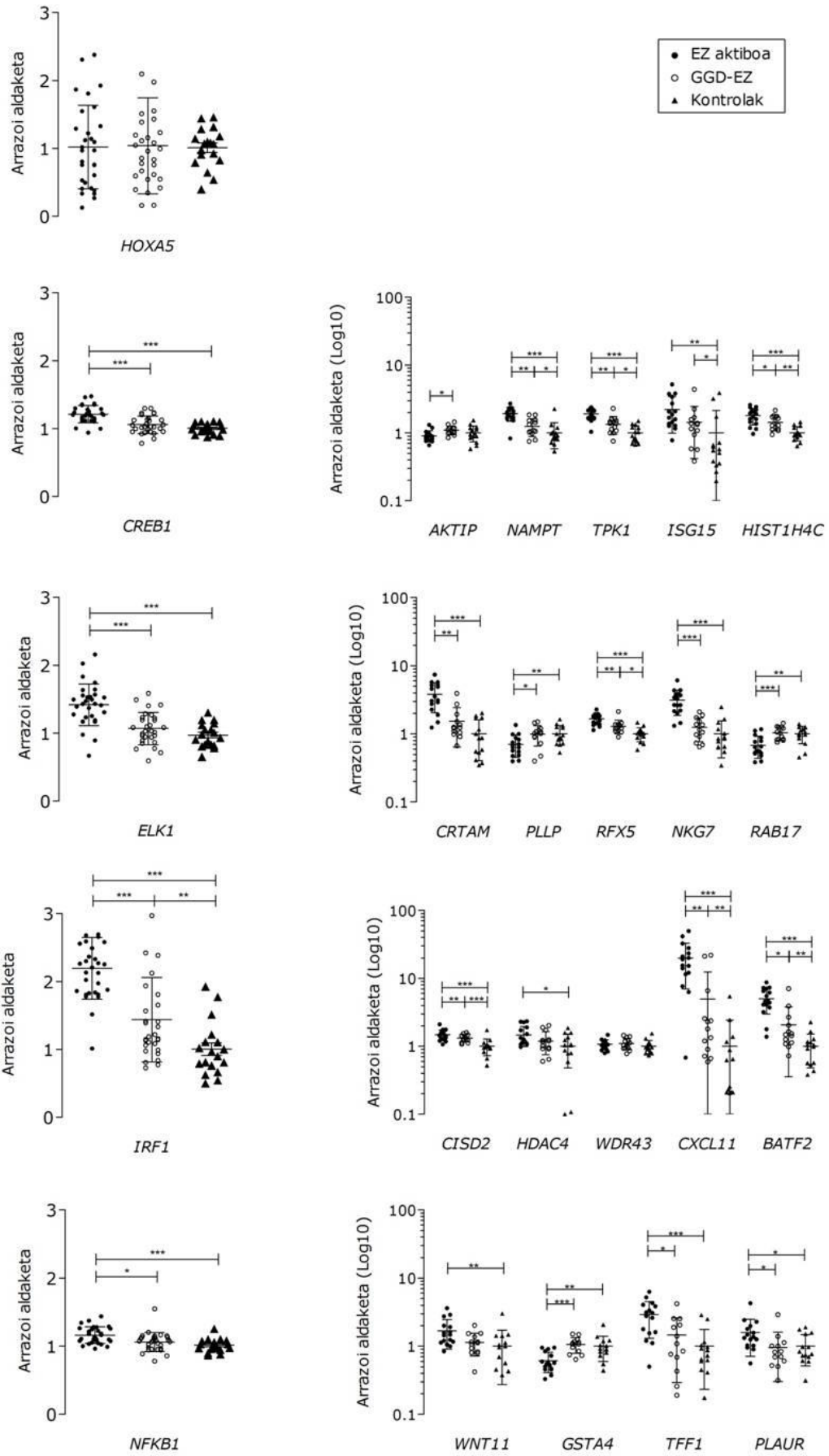
TFBS aberastea zuten DKGek eta DKGen moduluek funtzio edo esanahi biologikoren bat partekatzen zuten aztertzeko, GO terminoen aberaste-analisiak egin ziren. *Magenta* moduluko DKGetan (16A irudia) eta DKGen *Yellow* moduluan (16B irudia) ez zen aberasterik aurkitu, *Purple* (16C irudia) eta *Brown* (16D irudia) moduluko DKGek gaixotasunarekin erlazionatutako zenbait termino aurkeztu zuten bitartean; hala nola, lisosoma bidezko garraioa (Maiuri et al., 2010) (GO:0007041), apoptosia (GO:0006915), linfuzitoen aktibazioa (GO:0046649) eta sistema immunearen garapena (GO:0002520). DKGak modulu originaletatik kentzean, ez zen GO termino esangarririk aurkitu geratzen ziren geneetan.



16. irudia. Koadierazpen eta GO terminoen azterketa DKGetan. DKGen koadierazpen patroien aldaketak A) *Magenta*, B) *Yellow*, C) *Purple* eta D) *Brown* moduluetan. Lauki txiki bakoitzean bi gene jakinen adierazpen mailen arteko Spearman korrelazioaren koefizientea (ρ) irudikatu da (gorritik-urdinerako eskalak korrelazio positibotik negatiborako gradientea adierazten du). Azpian, DKG horiek genoma osoarekin konparatzean lortutako 10 GO termino nagusiak agertzen dira. GO terminoak Y ardatzean daude. X ardatzean $\log_{10}(\text{aberaste-arrazoia})$ azaltzen da, hau da, testatutako geneetatik GO terminoarekin anotatutako gene proportzioaren eta termino horrekin anotatutako genoma osoko geneen proportzioaren arteko ratioa (** $P < 0,01$ eta * $P < 0,05$).

1.4.2. TF hautagaien eta haien itu geneen adierazpena EZn

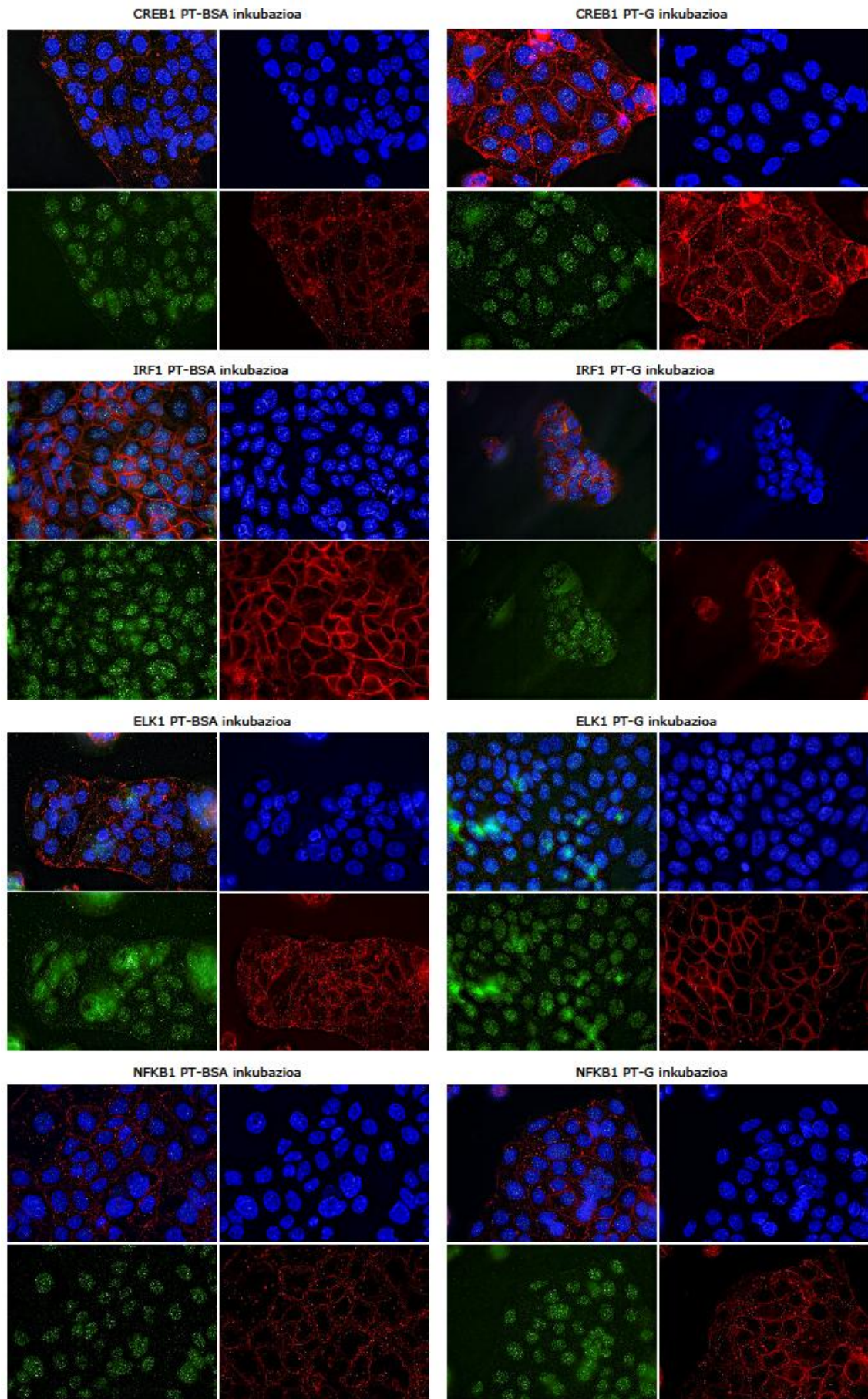
Pazienteen eta kontrolen lagin talde independente batean egindako adierazpen azterketek *IRF1*, *ELK1*, *NFKB1* eta *CREB1* gaixotasun aktiboan modu esangarrian gainadierazita zeudela erakutsi zuten, GGD-EZ zein kontrolekin konparatuta. TF horien 19 itu geneetatik, 17k zuten adierazpena aldatuta gaixotasun aktiboan ($P < 0,05$) (17. irudia).



17. irudia. Adierazpenaren emaitzak TF hautagaietan (EZ aktiboa, n = 30; GGD-EZ, n = 29; kontrolak, n = 18) eta haien itu geneetan (EZ aktiboa n = 16, GGD-EZ n = 13, kontrolak n = 14). Irudian, laginek kontrolen batz bestearikiko duten adierazpen erlatiboaren batz bestekoa eta desbiderapen estandarra ikus daitezke (*** P < 0,001; ** P < 0,01 eta * P < 0,05).

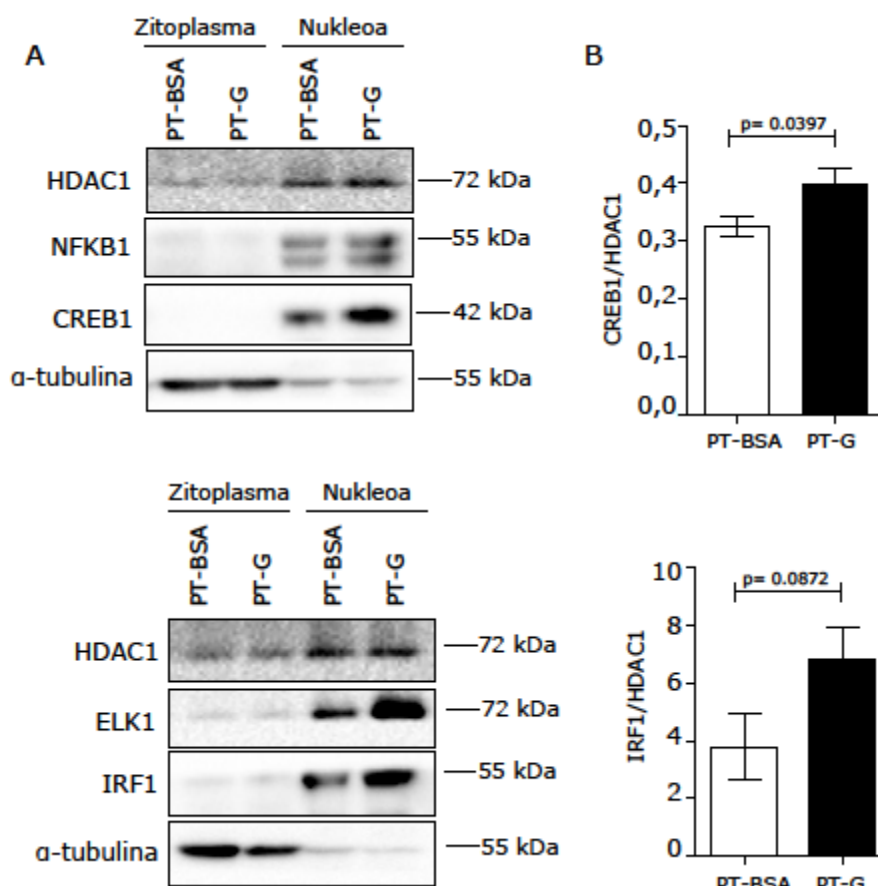
1.4.3. TF hautagaien kokapena lerro zelularretan

TF aktiboek eragin biologikoa izan dezaten nukleoan sartu behar dutenez, IRF1, NFkB1, ELK1 eta CREB1 TFen kokapena aztertu genuen PT-G edo PT-BSA-rekin 4 orduz inkubatutako C2BBe1 zeluletan. Immunofluoreszentzia esperimentuek (18. irudia) erakutsi zuten, 4 TFek kokapen nuklearra zuten bai PT-Grekin zein PT-BSA-rekin tratatutako zeluletan, baina ez zen ezberdintasun esangarririk aurkitu bi baldintzen artean.



18. irudia. Immunofluoreszentzia-tindaketa 4 orduz PT-G/PT-BSA tratamenduan egondako C2BBel zeluletan. CREB1, IRF1, ELK1 eta NFkB1 (berdez), DAPI (urdinez) eta E-Kadherina (gorriz).

Ondoren, IRF1, NFkB1, ELK1 eta CREB1 TFen kokapena aztertu genuen Western blot bidez, 4 orduz PT-G edo PT-BSAekin inkubatutako C2BBel zelulen frakzio nuklear eta zitoplasmatikoan; horrela, zelularen nukleoan PT-G kitzikapenaren ondoriozko TF emendioa gertatzen zen ikusteko asmoz. NFkB1 eta ELK1 kasuetan, ez zen ezberdintasunik ikusi baldintzen artean, baina PT-Gk nukleoan CREB1 eta IRF1 TFen emendioa eragiten zuela ikusi genuen (19. irudia).



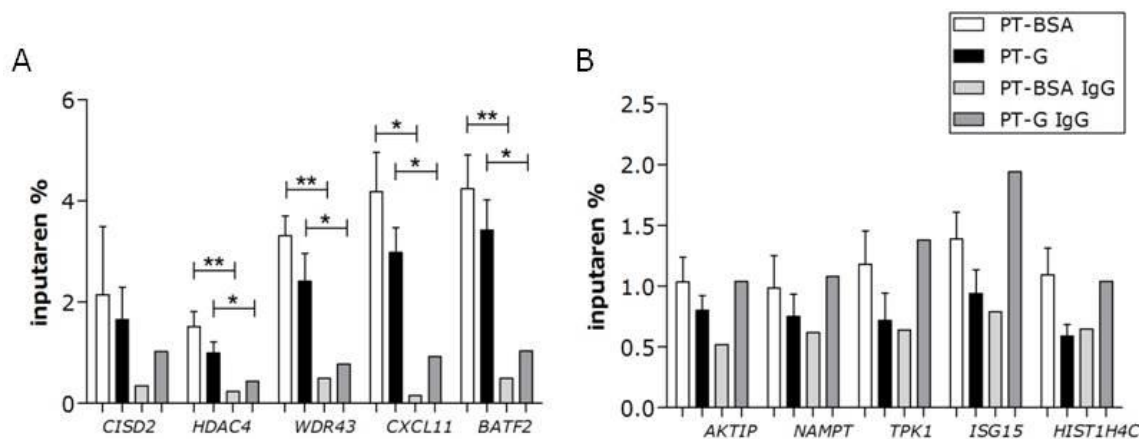
19. irudia. Aukeratutako TFen *in vitro* karakterizazioa. A) NFkB1, CREB1, ELK1 eta IRF1 TFen nukleorako translokazioaren azterketa Western blot bidez burutu zen C2BBel zelulak PT-G edo PT-BSAekin kitzikatutako 3 esperimentu independentetan. HDAC1 eta α-tubulina nukleoaren eta zitoplasmararen kontrol moduan erabili ziren, hurrenez hurren. B) Xingolen intentsitatea kuantifikatu eta HDAC1i zegokion xinglarekiko normalizatu ziren. Bataz besteko

baliok t-testen bidez konparatu ziren. Irudian batz bestekoa eta batz bestekoaren errore tipikoa marraztu dira.

1.4.4. TF hautagaien itu geneekiko lotura

ChIP esperimentuak egin ziren 4 orduz PT-G edo PT-BSAekin inkubatutako C2BBel zelulekin, gliadinarekin kitzikatzean IRF1 eta CREB1 TFen eta haien itu geneen arteko loturan aldaketarik gertatzen zen ikusteko.

Koadierazpen modulu batetik hartu ziren *HDAC4* eta *WDRA3*, IRF1en ituak zirela konfirmatu ahal izan genuen *in vitro*, baina PT-G eta PT-BSA baldintzen artean ez zen desberdintasun esangarririk topatu. CREB1en kasuan, lotura inespezifikoa bezain ahula izan zen ituekiko lotura espezifikoa eta beraz, ezin izan genuen ondoriorik atera (20. irudia).

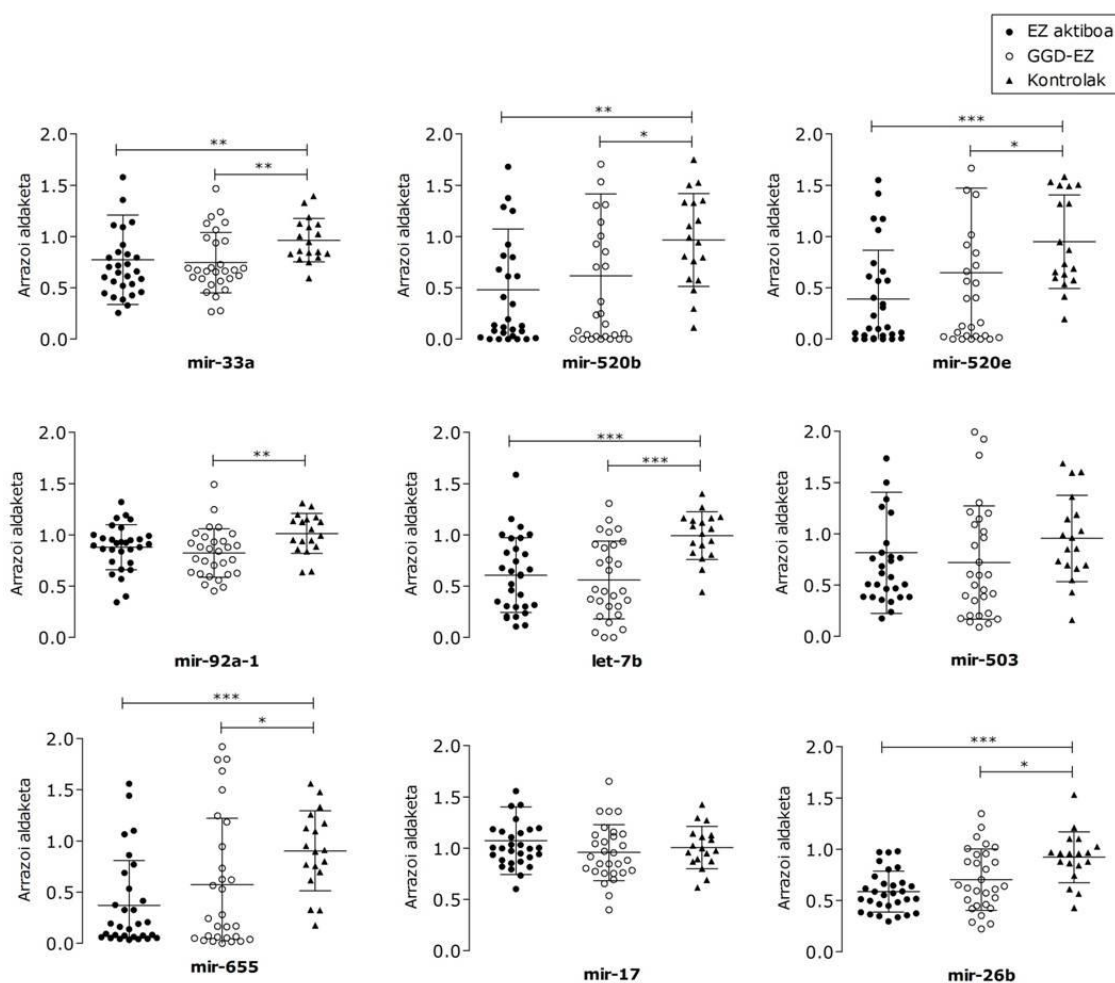


20. irudia. A) IRF1 eta B) CREB1 TFen itu geneen ChIP-qPCR balioztatzea. ChIP esperimentuak anti-TF antigorputzak erabiliz egin ziren 4 orduz PT-G edo PT-BSAekin inkubatutako C2BBel zeluletan. ChIParen eta unxi-IgGren (kontrol negatiboa) eluzioen batz bestekoa eta desbiderapen estandarra inputarekiko ehunekotan adierazi dira. **P < 0,01 eta *P < 0,05.

1.5. mikroRNAk

1.5.1. miRNA hautagaien adierazpena EZn

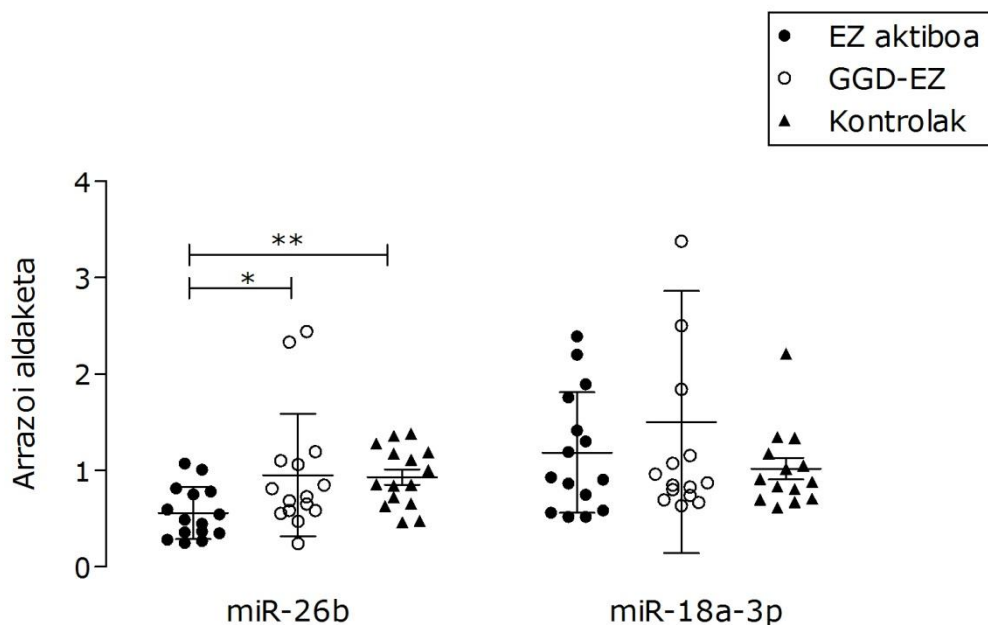
Bederatzi miRNA hautagaien pri-miRNAk kuantifikatu genituen adierazpen geniko modura. Sei pri-miRNA azpiadierazita zeuden EZ aktiboan kontrolekin konparatuta, pri-miRNA hauek azpiadierazita mantendu ziren GGD-EZ laginetan (21. irudia).



21. irudia. pri-miRNAen adierazpen analisia EZ aktibo (n = 29), GGD-EZ (n= 29), eta kontrolen (n = 18) duodenoko biopsietan. Irudian, laginek kontrolen bataz bestearekiko duten adierazpen erlatiboaren bataz bestekoa eta desbiderapen estandarra ikus daitezke (***) P < 0,001; ** P < 0,01 eta * P < 0,05).

Bi miRNA helduen adierazpena ere aztertu zen pri-miRNAen emaitzekin bat zetozen ikusteko. has-miR-18-3p miRNA helduak ez zuen aldaketarik erakutsi,

bere pri-miRNAk (has-mir-17) bezala. Bestalde, has-miR-26b miRNA heldua azpiadierazita zegoen EZ aktiboan kontrol ($P = 0,0025$) eta GGD-EZ ($P = 0,0125$) taldeekin konparatuta, bere pri-miRNA bezala (22. irudia).

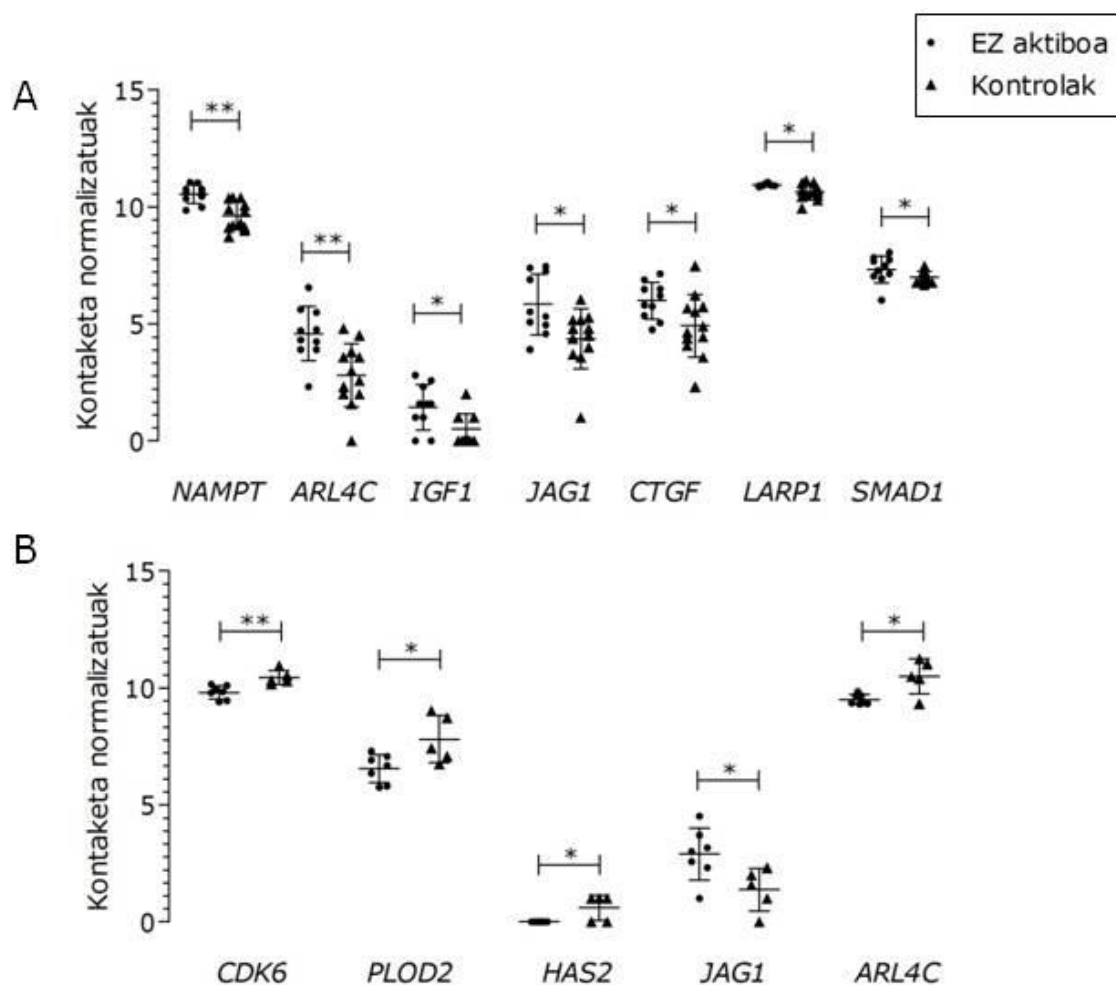


22. irudia. Bi miRNA helduen adierazpen analisia (EZ aktibo $n = 15$, GGD-EZ $n = 15$, kontrolak $n = 15$) duodenoko biopsietan. Irudian, laginek kontrolen batz bestearikiko duten adierazpen erlatiboaren batz bestekoa eta desbiderapen estandarra ikus daitezke (** $P < 0,01$ eta * $P < 0,05$).

1.5.2. miRNAen itu geneen adierazpena EZn

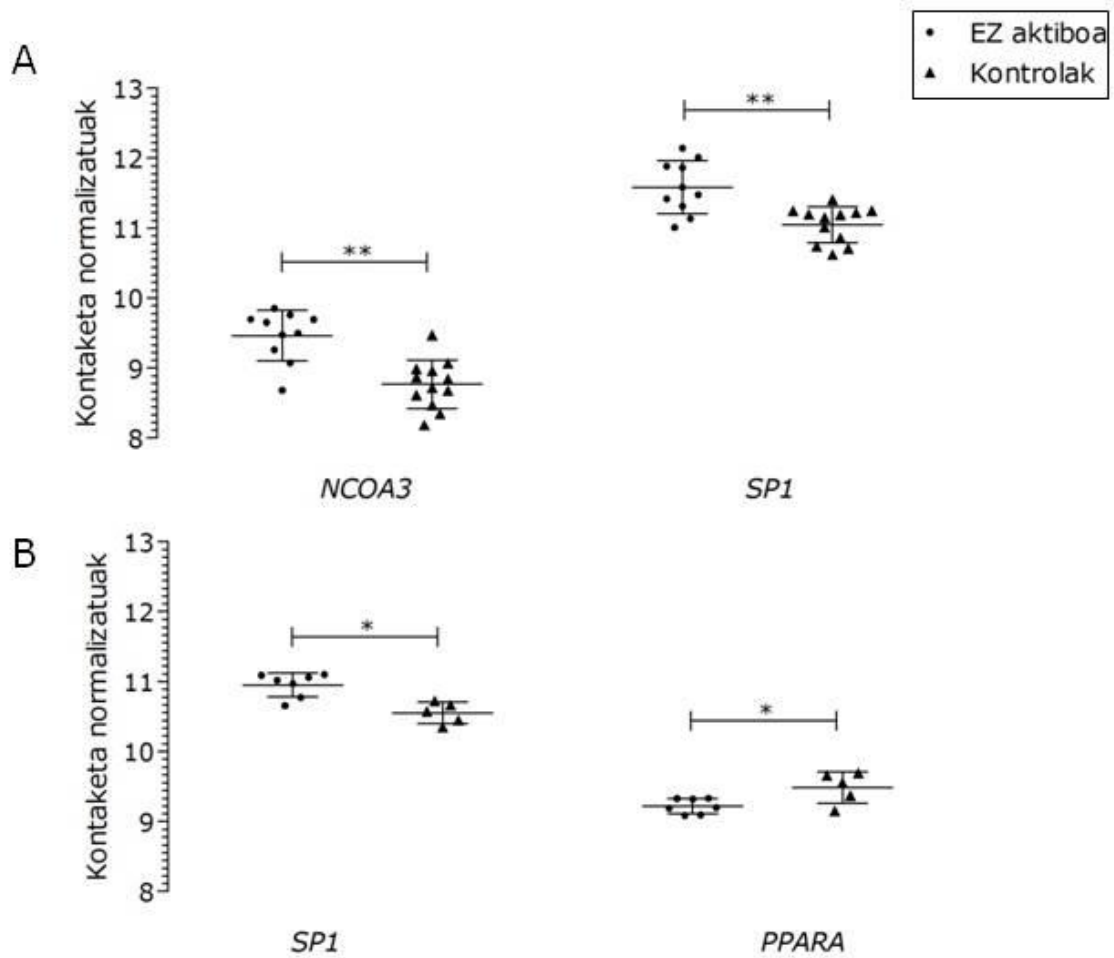
miRNAen itu geneen adierazpena EZ aktibo eta kontrolen epitelioko zeluletan eta immunitate zeluletan aberastutako frakzioen RNA-seq datuetan aztertu zen.

hsa-miR-26b-ren kasuan, MirTarBase datu-basean identifikatutako 35 itu geneetatik zazpik adierazpen aldaketak erakutsi zituzten ($P < 0,05$) epitelioko zeluletan aberastutako frakzioan (23A irudia). Immunitate zeluletan aberastutako frakzioan, aldiz, adierazpen aldaketak bost itu genetan ikusi ziren (23B irudia).



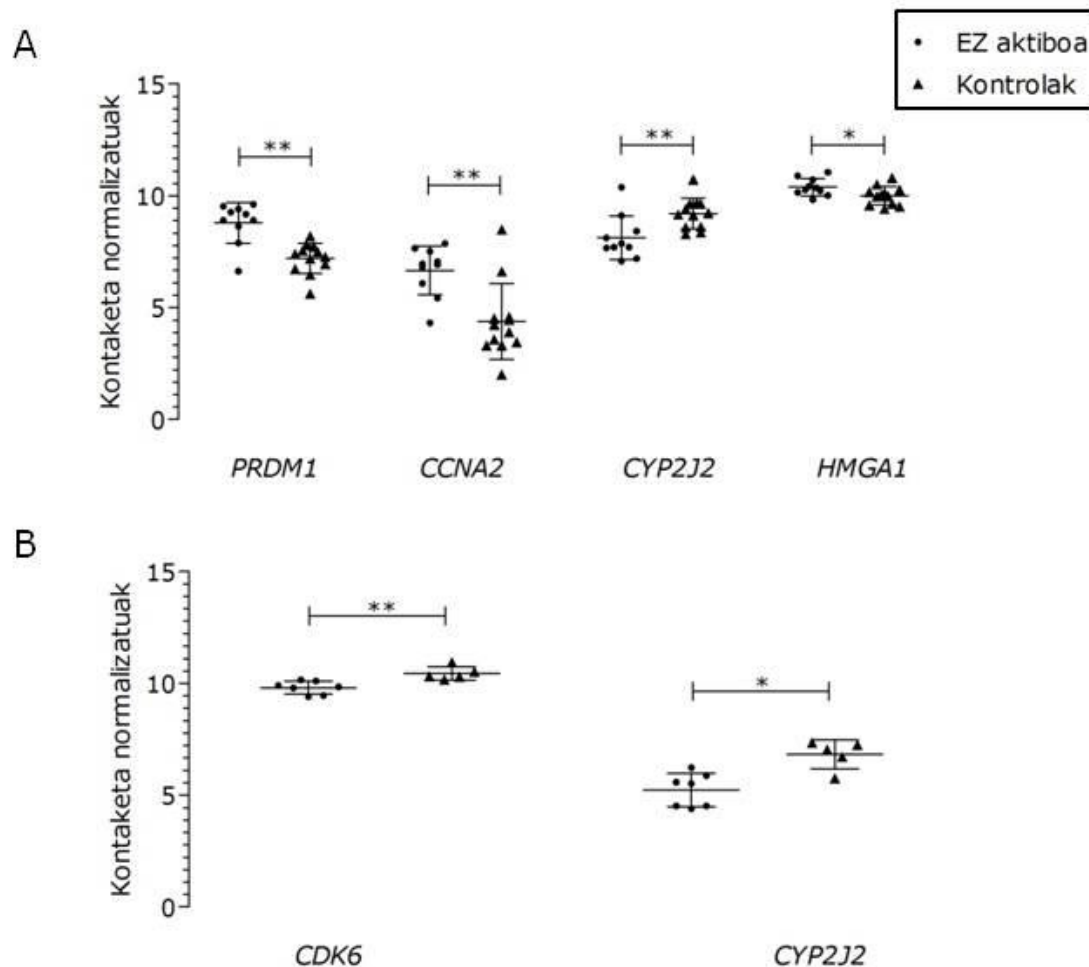
23. irudia. hsa-miR-26b-ren itu geneen adierazpen analisia EZ aktibo eta kontrolen biopsien A) epitelio- eta B) immunitate-zeluletan aberastutako frakzioetan. Irudian kontaketa normalizatuen bataz bestekoa eta desbiderapen estandarra ikus daitezke (** $P < 0,01$ eta * $P < 0,05$).

has-miR-33a-ri dagokionez, datu-basean identifikatutako 34 itu geneetatik bik erakutsi zituzten adierazpen aldaketak EZ aktiboan ($P < 0,05$), epitelioko eta immunitate zeluletan aberastutako frakzioetan (24. irudia).



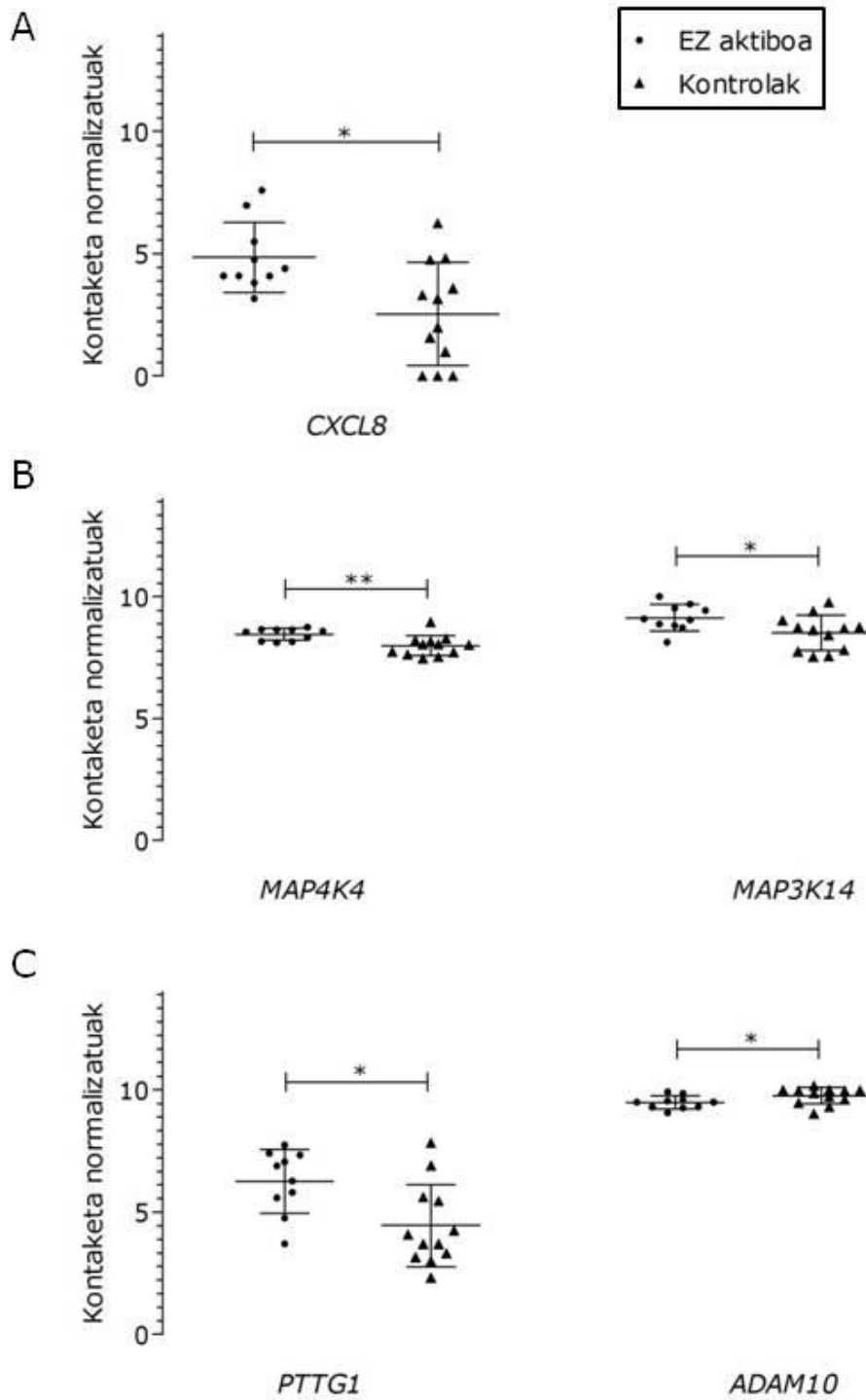
24. irudia. hsa-miR-33a itu geneen adierazpen analisia EZ aktibo eta kontrolen biopsien A) epitelio- eta B) immunitate-zeluletan aberastutako frakzioetan. Irudian kontaketa normalizatuen batz bestekoa eta desbiderapen estandarra ikus daitezke (** $P < 0,01$ eta * $P < 0,05$).

hsa-let-7b-3p-ren kasuan, datu basean identifikatutako 36 itu geneetatik lauk adierazpen aldaketak erakutsi zituzten ($P < 0,05$) EZ aktiboan epitelioko zeluletan aberastutako frakzioan (25A irudia). Immunitate zeluletan aberastutako frakzioan, aldiz, adierazpen aldaketak 2 genetari ikusi ziren (25B irudia).



25. irudia. hsa-let-7b-3p-ren itu geneen adierazpen analisia EZ aktibo eta kontrolen biopsien A) epitelio- eta B) immunitate-zeluletan aberastutako frakzioetan. Irudian kontaketa normalizatuen batz bestekoa eta desbiderapen estandarra ikus daitezke (** $P < 0,01$ eta * $P < 0,05$).

hsa-miR-520b-ren kasuan, 9 ituetatik batean zegoen adierazpena aldatuta, EZ aktiboan epitelioko zeluletan aberastutako frakzioan. hsa-miR-520e eta has-miR-655-ren kasuan, identifikatutako sei ituetatik bi zeuden aldatuta (26. irudia).



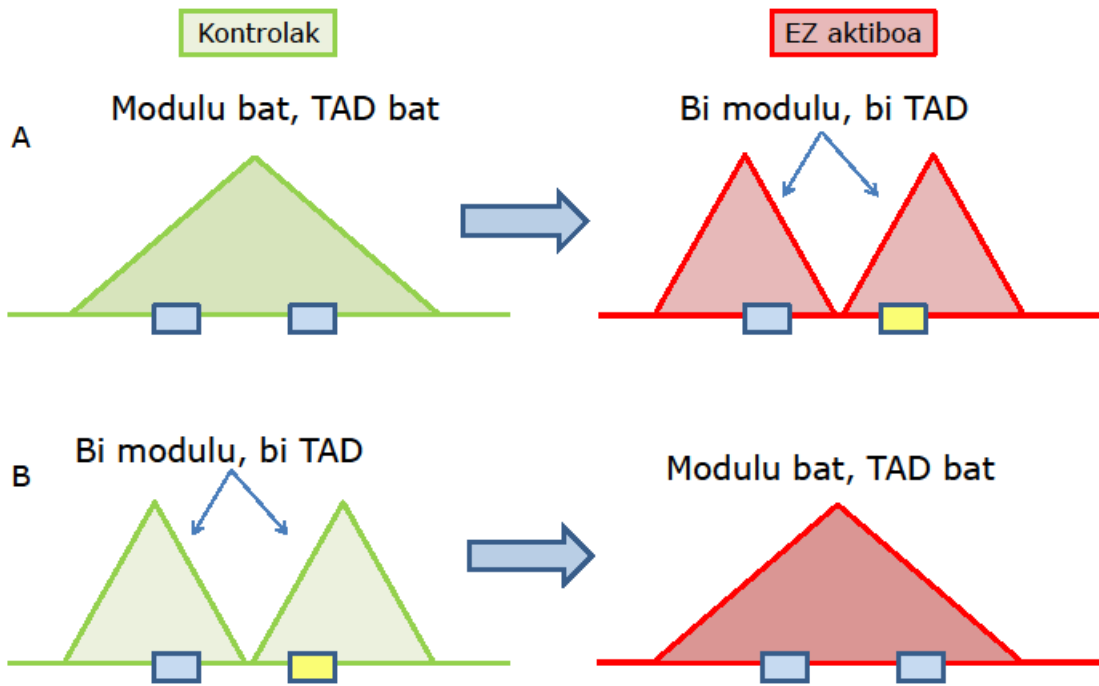
26. irudia. A) hsa-miR-520b, B) hsa-miR-520e eta C) has-miR-655 miRNAen itu geneen adierazpena EZ aktibo eta kontrolen biopsien epitelioko zeluletan aberastutako frakzioetan. Irudian kontaketa normalizatuaren batz bestekoa eta desbiderapen estandarra ikus daitezke (**P < 0,01 eta * P < 0,05).

2. TADak EZn

EZ gaixotasun konplexua da, adierazpen eta koadierazpen patroietan alterazioak erakusten dituena, besteak beste. Hala ere, aldaketen oinarrian dauden mekanismoak zeintzuk diren ez dago argi. TADak adierazpen geniko koordinaturako domeinu funtzionalak dira, eta aipaturako aldaketetan eragin lezakete.

2.1. Eskualde eta gene hautagaien identifikazioa eta karakterizazioa

Elkarrengandik gertu dauden geneen arteko koadierazpen aldaketak TADen egituran egon daitezkeen alterazioekin asoziatu daitezkeela uste dugu (27. irudia). Beraz, epitelioko zeluletan aberastutako frakzioen RNA-seq datuak erabiliz, koadierazpen patroiak eratu genituen, eta koadierazten ziren geneak kontserbatutako TADekin gainezartzen ziren begiratu genuen. Kontroletan, 11 koadierazpen modulutan taldekatutako 2.826 gene aztertu ziren, eta horietatik 739 gene 486 TADetan kokatuta zeuden. EZ aktiboa zuten pazienteen kasuan, 18 koadierazpen modulutako 3.992 gene aztertu ziren, eta horietatik 613 gene 430 TADetan kokatzen ziren.



27. irudia. TAD alterazioen eredu hipotetikoak EZn. TADen A) haustura eta B) fusio ereduak ikus daitezke. Lehenengo kasuan, elkarrengandik gertu dauden bi gene TAD berean kokatzen dira eta kontroletan koadierazi egiten dira, baina ez ordea, EZn, TADen haustura irudikatuta. Bigarren kasuan, elkarrengandik gertu dauden bi gene TAD desberdinetan kokatzen dira eta ez dira kontroletan koadierazten, bai ordea, EZn, TADen fusioa irudikatuta.

Alde batetik, 16 TADen haustura islatu dezaketzen koadierazpen aldaketak identifikatu genituen: koadierazitako geneak TAD bakarrean kokatzen ziren kontroletan, baina bi koadierazpen modulu desberdinetan banatzen ziren gaixotasun aktiboan (27A irudia) (14. taula). Bestetik, TADen fusioa adierazi lezaketzen balizko 30 TAD ere identifikatu ziren: TADak elkarrengandik gertu zeuden (2 Mb baino gutxiagora) eta kontroletan koadierazten ez ziren baina EZ aktiboan koadierazten ziren geneak zituzten (27B irudia) (15. taula).

14. taula. TADen haustura eta koadierazpenaren galera. Ilara bakoitzean TAD bat eta kontroletan koadierazten ziren baina EZ aktiboan koadierazten ez ziren geneak ikus daitezke, TAD barruan dauden EZrekin asoziatutako SNPekin batera. Horien artean, p balio txikiak dutenak zerrendatu dira, parentesian TAD bakoitzerako SNP kopurua.

TAD koordenatuak (hg19)	geneak	SNP nagusia	p balio nominala
chr1:178680000-178760000	<i>RALGPS2; RP11-428K3.1</i>		
chr6:260400000-261600000	<i>HIST1H2BB; XLOC_059355</i>	rs3857546 (22)	4,35e-09
chr6:526000000-531200000	<i>GSTA2; GSTA1</i>		
chr7:165600000-169600000	<i>AC073333.1; AGR2</i>		
chr8:717200000-718000000	<i>XLOC_065435; XLOC_065445</i>		
chr8:860400000-864400000	<i>E2F5; CA3; CA2</i>		
chr9:131400000-131480000	<i>SET; XLOC_070707</i>		
chr10:819200000-820800000	<i>XLOC_009508; XLOC_009510</i>	rs12219032 (4)	0,005309
chr11:623200000-624800000	<i>UQCC3; EEFIG</i>		
chr12:132400000-133600000	<i>XLOC_014371; XLOC_014372; XLOC_014373</i>		
chr15:405200000-408000000	<i>DISP2, PHGR1; IVD</i>		
chr16:678800000-679600000	<i>XLOC_025619; XLOC_027035</i>		
chr2:130840000-131240000	<i>MZT2B; CCDC115</i>		
chr22:290400000-293600000	<i>HSCB; XBP1</i>	rs2097461 (3)	0,02041
chr22:300000000-302400000	<i>NF2; UQCR10</i>	imm_22_28471822 (15)	0,0004322
chr22:373600000-375200000	<i>MPST; TST</i>	rs5995385 (1)	0,03635

*XLOC: anotatu gabeko adierazpen seinaleak.

15. taula. TADen fusioa eta koadierazpenaren ezarpena. Ilara bakoitzean TAD bat eta kontroletan koadierazten ziren baina EZ aktiboan koadierazten ez ziren geneak ikus daitezke, TAD barruan dauden EZrekin asoziatutako SNPekin batera. Horien artean, p balio txikiak dutenak zerrendatu dira, parentesia TAD bakoitzerako SNP kopurua.

A + B TADa (hg19)	A TADa (hg19)	A gena	B TADa (hg19)	B gena	SNP nagusia	p balio nominala
chr1:178400000-178760000	chr1:178400000-178520000	<i>XLOC_006089</i>	chr1:178680000-178760000	<i>RALGPS2</i>	rs10207341 (9)	0,0006984
chr2:134280000-135000000	chr2:134280000-134400000	<i>NCKAP5</i>	chr2:134880000-135000000	<i>XLOC_037431</i>		
chr3:23960000-25680000	chr3:23960000-24040000	<i>RPLI5</i>	chr3:25560000-25680000	<i>RARB</i>	rs9867080 (1)	0,03304
chr3:32680000-33960000	chr3:32680000-33160000	<i>CRTPAP</i>	chr3:33800000-33960000	<i>RP11-10C24.1</i>		
chr6:144360000-146360000	chr6:144360000-144440000	<i>SF3B5</i>	chr6:146040000-146360000	<i>RP11-54515.3</i>		
chr6:99760000-101480000	chr6:99760000-100040000	<i>COQ3</i>	chr6:101000000-101480000	<i>ASCC3</i>	rs924974 (2)	0,0136
chr8:41480000-42400000	chr8:41480000-41640000	<i>NKX6-3</i>	chr8:42320000-42400000	<i>RP11-503E24.2</i>		
chr11:33680000-34520000	chr11:33680000-33840000	<i>RP4-541C22.5 CD59 C11orf91</i>	chr11:34320000-34520000	<i>ABTB2</i>		
chr11:34320000-34920000	chr11:34320000-34520000	<i>ABTB2</i>	chr11:34840000-34920000	<i>APIP</i>		
chr11:34320000-36840000	chr11:34320000-34520000	<i>ABTB2</i>	chr11:36440000-36840000	<i>C11orf74</i>	rs4755450 (10)	0,000522
chr14:74600000-74680000	chr14:74600000-74680000	<i>LIN52</i>	chr14:75920000-76440000	<i>C14orf1</i>	rs1569328 (4)	0,0003735
chr15:63320000-65040000	chr15:63320000-63440000	<i>RP11-244F12.3</i>	chr15:64920000-65040000	<i>OAZ2</i>	rs4411464 (2)	0,008319
chr15:71960000-72480000	chr15:71960000-72080000	<i>THSD4 RP11-592N21.2</i>	chr15:72400000-72480000	<i>GRAMD2</i>		
chr17:18840000-19120000	chr17:18840000-19120000	<i>SNORD3A</i>	chr17:20160000-20400000	<i>LGALS9B</i>	rs2703817 (1)	0,03975
chr17:18840000-20400000	chr17:18840000-19120000	<i>SNORD3C</i>	chr17:20160000-20400000	<i>LGALS9B</i>	rs2703817 (1)	0,03975
chr17:3480000-3720000	chr17:3480000-3720000	<i>RP11-235E17.4</i>	chr17:4680000-5080000	<i>RP5-1050D4.4</i>	rs9906760 (8)	0,001804
chr17:40520000-42160000	chr17:40520000-40680000	<i>STAT3</i>	chr17:41720000-42160000	<i>TMEM101</i>	imm_17_38056077 (77)	7,51E-02
chr17:40920000-42160000	chr17:40920000-41080000	<i>G6PC</i>	chr17:41720000-42160000	<i>HDAC5</i>	rs382571 (4)	0,005572
chr17:41720000-42160000	chr17:41720000-42160000	<i>HDAC5</i>	chr17:43040000-43320000	<i>PLCD3 MIR6784</i>	rs9903582 (1)	0,0114
chr17:9360000-9800000	chr17:9360000-9800000	<i>USP43</i>	chr17:10440000-10800000	<i>XLOC_027658</i>	rs9903582 (1)	0,0114
chr17:9360000-9800000	chr17:9360000-9800000	<i>USP43</i>	chr17:10440000-10800000	<i>TMEM220-AS1</i>		
chr19:12880000-13000000	chr19:12880000-13000000	<i>RNASEH2A</i>	chr19:14040000-14160000	<i>CTB-55O6.13</i>	rs416162 (1)	0,04632
chr19:13800000-13880000	chr19:13800000-13880000	<i>XLOC_032850</i>	chr19:14400000-14520000	<i>ADGRE5</i>	rs416162 (1)	0,04632
chr19:51840000-51960000	chr19:51840000-51960000	<i>XLOC_033809</i>	chr19:53080000-53160000	<i>ZNF83</i>	rs16982743 (3)	0,03813
chr20:32800000-32920000	chr20:32800000-32920000	<i>RP4-785G19.5 AHY CTD-3216D2.5</i>	chr20:33480000-33800000	<i>PROCR</i>	rs6059916 (1)	0,03293
chr20:33480000-33800000	chr20:33480000-33800000	<i>PROCR</i>	chr20:34160000-34320000	<i>ROMO1</i>	rs224436 (11)	0,003579
chr22:41200000-41400000	chr22:41200000-41400000	<i>RBX1</i>	chr22:43400000-43640000	<i>TSP0</i>	rs5758209 (2)	0,004118
chr22:41200000-41400000	chr22:41200000-41400000	<i>RBX1</i>	chr22:43400000-43640000	<i>MCAT</i>	rs5758209 (2)	0,004118
chr22:41800000-41960000	chr22:41800000-41960000	<i>XLOC_045800</i>	chr22:43400000-43640000	<i>BIK</i>		
chr22:41800000-43640000	chr22:41800000-41960000	<i>XLOC_045800</i>	chr22:43400000-43640000	<i>XLOC_045853</i>		

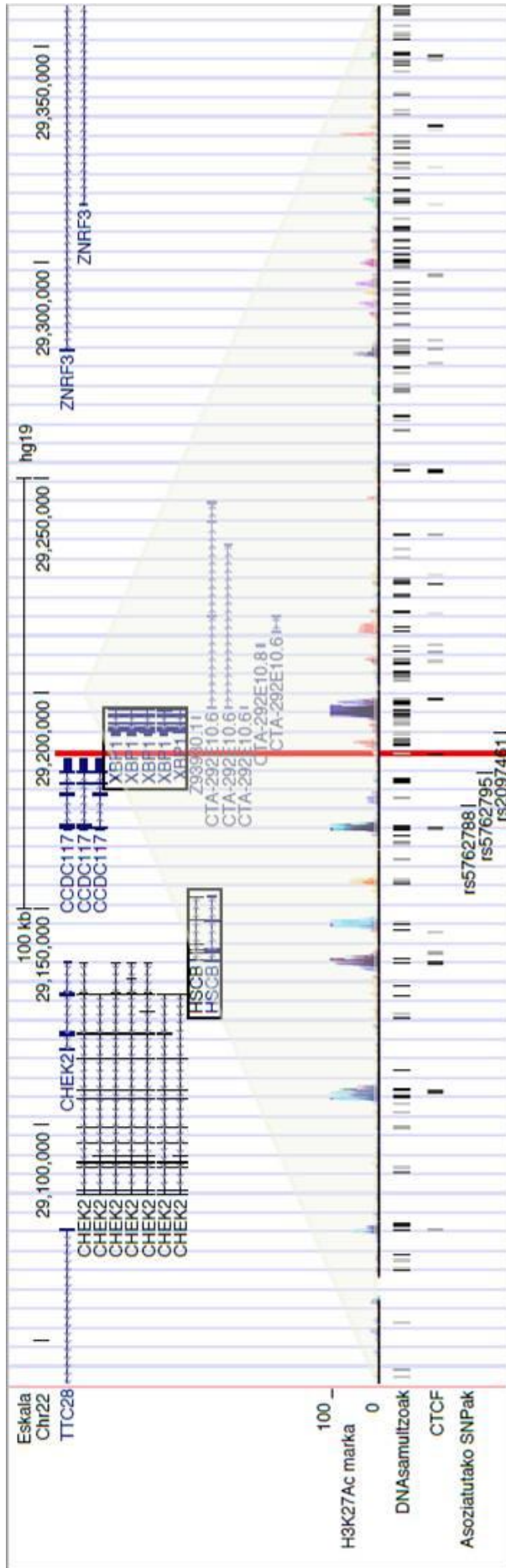
**XLOC*: anotatu gabeko adierazpen seinaleak.

Geroago, TADen egituran eragin eta EZn gertatzen diren koadierazpen aldaketak azaldu ditzaketen EZrekin asoziatutako SNPak (p balio nominala $< 0,05$) bilatu genituen. Horretarako, TADen eta EZri asoziatutako SNPen koordinatuak gurutzatu genituen. TAD zehatz batekin lerrokatzen ziren *Immunchipeko* 434 SNP identifikatu genituen. SNPok 13 eskualdetan banatzen ziren, eta 434tik 45 zeuden EZrekin asoziatuta; beraz, gaixotasunarekin erlazionatutako TADen hausturak eragin zitzaizkiren (14. taula). Bestetik, TADen fusio posiblearekin lerrokatzen ziren 916 SNP identifikatu ziren. Hauek, 25 eskualdetan banatzen ziren, eta 916tik 142 zeuden EZrekin asoziatuta; beraz, gaixotasunarekin lotutako TADen fusioak eragin zitzaizkiren (15. taula).

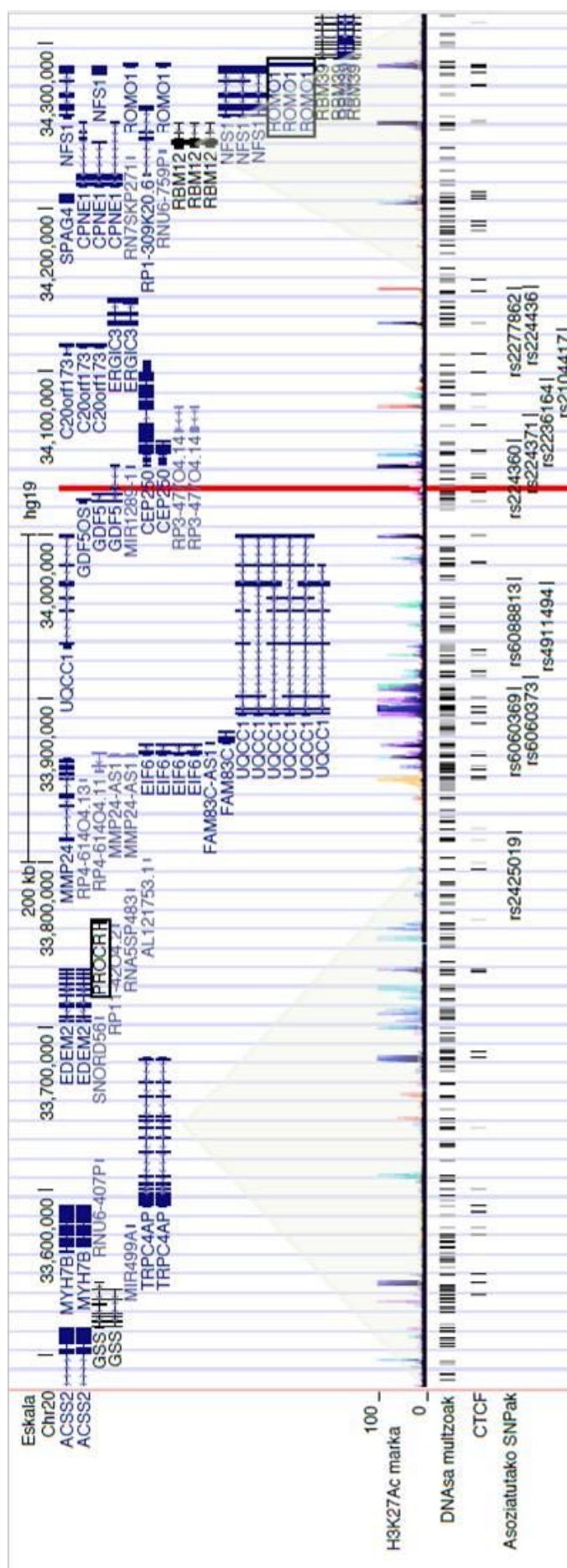
Ikerketan sakontzeko, horietako bi eskualde hautatu ziren. Anotatutako geneen koadierazpen aldaketetan inplikatuak egon litezkeen eta asoziatutako SNP talde batekin lerrokatzen zirenak hautatu ziren: alde batetik, hg19 chr22:29040000-29360000 koordinatuetan dagoen *HSCB-XBPI* eskualdea (TADaren haustura); eta beste aldetik, hg19 chr20:33480000-34320000 koordinatuetan dagoen *PROCR-ROMOI* eskualdea (bi TADen fusioa).

2.2. Identifikatutako eskualdeen kromatinaren eskuragarritasuna

Ondoren, CTCF motiboak zituzten DHSak aukeratu genituen, TADen mugak izan zitezkeen kromatina aktiboaren guneak identifikatzeko. Hurbilketa hau erabiliz, *HSCB-XBPI* eskualde barruko hg19 chr22:29186201-29186490 eta *PROCR-ROMOI* eskualde barruko hg19 chr20:34026861-34027090 eremuak hautatu ziren (28 eta 29. irudiak).

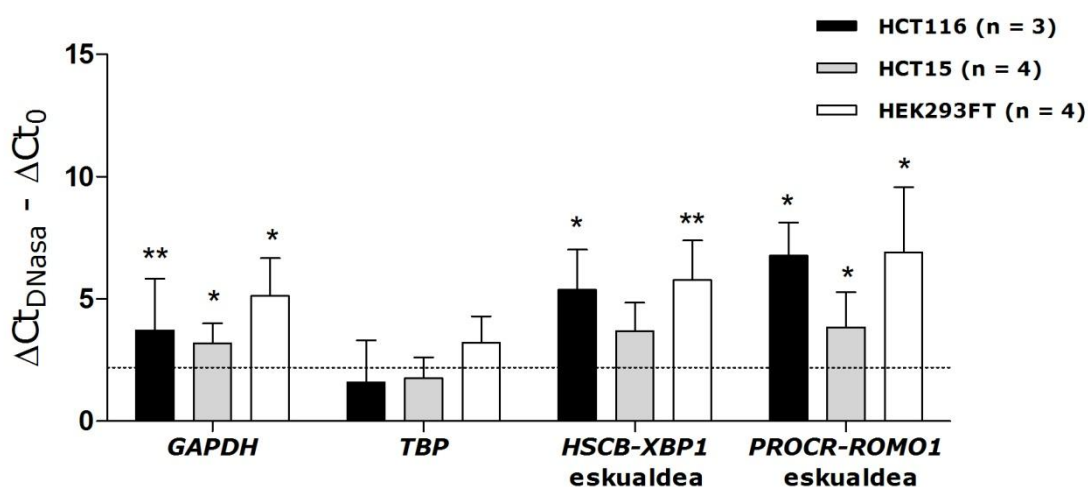


28. irudia. *HSCB-XBPI* eskualde hautagaia (TADaren haustura). Laukizuzen beltzek EZn koadierazpen aldaketak dituzten geneak inguratzen dituzte; TADaren egitura aldarazi dezaketen asoziatutako SNPak ikus daitezke. Marra gorriak CTCF lotura-guneak dituen DHSa adierazten du. Hiruki berde argia kontrolatan dagoen TADa da.



29. irudia. *PROCRA-ROMO1* eskualde hautagaia (TADen fusioa). Laukizuzen beltzek EZn koadierazpen aldaketak dituzten geneak inguratzen dituzte; TADaren egitura aldarazi dezaketen asoziatutako SNPak ikus daitezke. Marra gorriak CTCF lotura-guneak dituen DHSa adierazten du. Hiruki berde argiak kontroletan dauden TADak dira.

In silico identifikatutako DHSak gure eruedetan DNasarentzat eskuragarri zeuden ikertzeko, DNasarekin digeritutako kromatinaren qPCRak egin ziren HCT116, HCT15 eta HEK293FT zelula lerroetan. Bi guneek DNasa eskuragarritasun handia azaldu zuten HCT116 eta HEK293FT zelula lerroetan, eukromatinaren egiturarekin bat etorrira. HCT15 zeluetan *PROCR-ROMO1* eskualdean kokatutako DHSak DNasarekiko eskuragarritasuna erakutsi zuen, ez ordea, *HSCB-XBP1* eskualdean kokatutakoak (30. irudia).



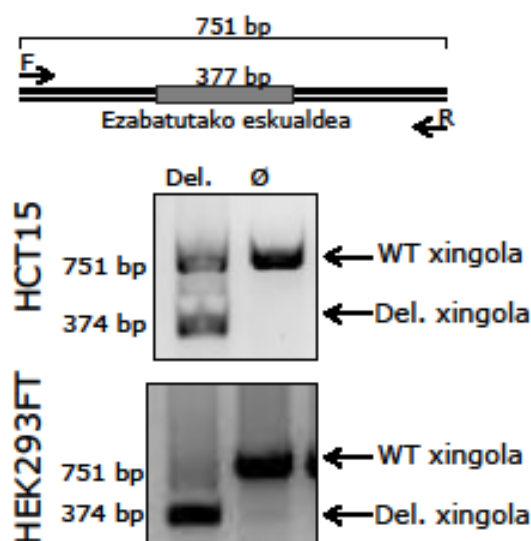
30. irudia. DNasarekiko eskuragarritasuna. HCT116, HCT15 eta HEK293FT zelulen kromatinaren digestioa burutu zen DNasarekin, eta qPCRak burutu ziren. DNasarekiko eskuragarritasuna digeritutako eta digeritu gabeko kromatinaren Ct balioen arteko aldearen arabera kuantifikatu zen. Eskualde hautagaiak kromatina itxiaren kontrolarekin (*TBP*) konparatu ziren zelula lerro bakoitzean. Irudian batz bestekoa eta desbiderapen estandarra ikus daitezke (** $P < 0,01$ eta * $P < 0,05$). *TBP*ren batz besteko balioa marra etenaz adierazi da.

2.3. Eskualde hautagaien edizioa

Identifikatutako eskualdeen funtzioa argitzeko, eredu diren epitelioko zeluletan CRISPR-Cas9 erabili zen CTCF lotura-guneak zituzten DHSak modu egonkor batean hausteko. Horretarako, *HSCB-XBP1* eskualdean 441 bp-ko delezioa egin zen, eta 377 bp-koa *PROCR-ROMO1* eskualdean.

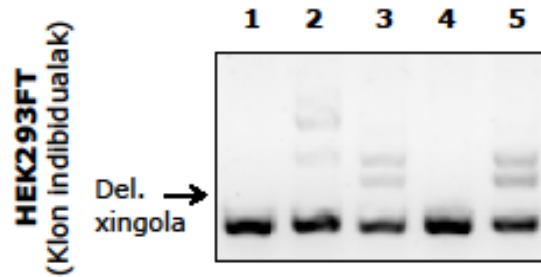
2.3.1. Gene edizioen konfirmazioa

sgRNA egokiak Cas9 bektorean klonatu ziren eta plasmido errekonbinanteak epitelio zeluletan (HCT116, HCT15 eta HEK293FT) transfektatu ziren. Delezio posibleak balioztatzeko eskualdearen amplifikazioa burutu zen ituaren alboetara zeuden hasleak erabiliz. *PROCR-ROMO1* eskualde barneko eremuaren ezabaketa lortu zen HCT16 eta HEK293FT zeluletan (31. irudia).



31. irudia. *PROCR-ROMO1* eskualde barruan CTCF lotura-gunea duen DHSaren ezabaketa giza zeluletan CRISPR-Cas9 teknologia erabiliz. Itu eskualdearen alboetara dauden hasleak (F eta R) erabili ziren delezioa gertatu zen aztertzeko. Del. xingola = delezioa duen xingola; WT xingola = jatorrizko xingola; Ø = editatu gabeko lerroak; Del. = editatutako lerroak.

Genomaren edizioa %100ean eraginkorra izan ez zenez (jatorrizko zelulen xingola ageri da kasu guztietan), editatu gabeko zelulak ere agertu ziren; hau da, populazio mistoa (heterogeneoa) lortu zen. Zelulak banaka isolatu eta bakoitzetik klonak hedatu ziren; ondoren, delezio-mutazioa zeramaten klonak hautatu ahal izateko (32. irudia). HEK293FT zelulen kasuan, zelulen isolamendua aurrera gauzatzea eta lerro klonalak hedatzea lortu zen, eta *PROCR-ROMO1* eskualde barruko eremuaren deleziodun 6 klon homozigoto lortu ziren. Heste zelulen lerroetan, aldiz, ez zen mutazioa zerman klonik lortu.

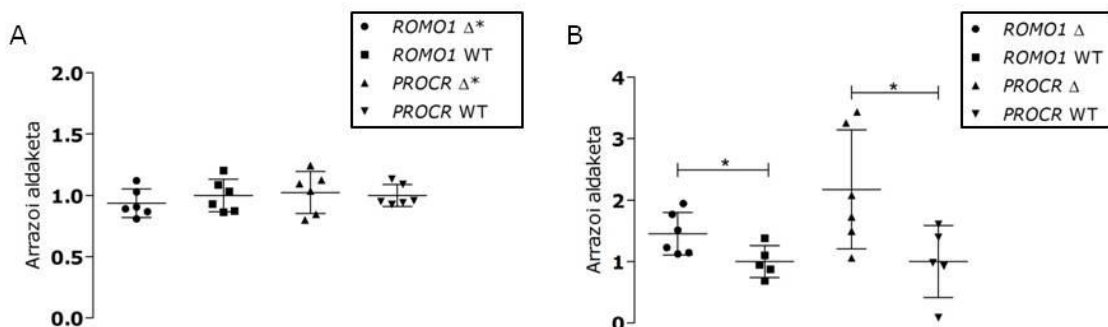


32. irudia. Klon homozigoto (1, 4) eta heterozigotoen (2, 3, 5) adibidea HEK293FT zeluletan CRISPR-Cas9 bidezko delezio gidatua eta gero. Delezioa daramaten lerro klonalak PCR bidez identifikatu ziren, horretarako ituaren alboko eremuetako hasleak erabiliz. Del. xingola = delezioa duen xingola.

2.3.2. Adierazpen eta koadierazpen analisiak

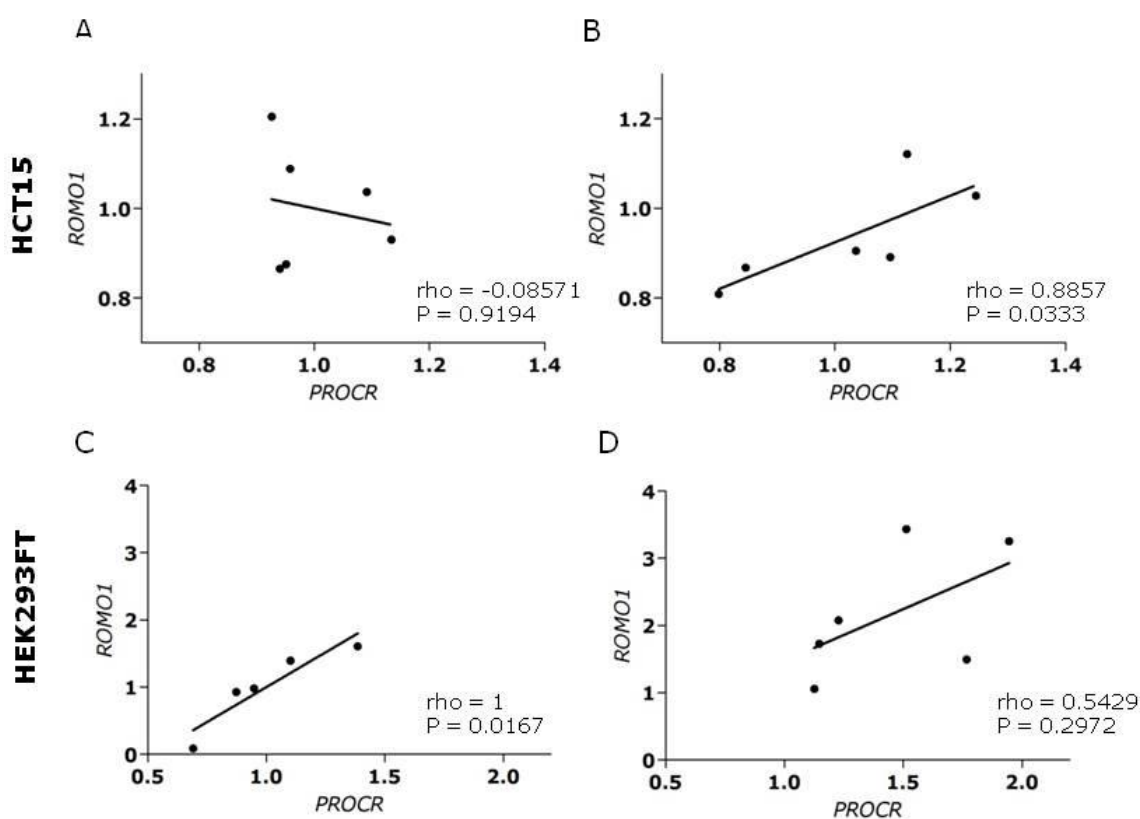
Delezioak ondoko geneen erregulazioan eragiten zuen ikertzeko, adierazpen eta koadierazpen azterketak burutu ziren. Jatorrizko eta editatutako HEK293FT klonak, zein jatorrizko HCT15 eta HCT15 populazio mistoa ikertu ziren.

PROCR eta *ROMO1* geneen adierazpenari dagokionez, ez zen aldaketa esangarririk ikusi HCT15 populazio mistoa jatorrizko zelulekin alderatuta (33A irudia). Hala ere, editatutako HEK293FT zelulak jatorrizko HEK293FT zelulekin alderatzean, *PROCR* eta *ROMO1* geneen adierazpenean aldaketa esangarriak aurkitu ziren (33B irudia).



33. irudia. *PROCR* eta *ROMO1* geneen analisia A) HCT15 eta B) HEK293FT zelula lerroetan (n = 6 baldintza guztietan). WT = jatorrizko zelulak; Δ = editatutako zelulak. *populazio mistoa. Irudian batz bestekoa eta desbiderapen estandarra ikus daitezke (*P < 0.05).

Koadierazpena aztertu zenean, jatorrizko HCT15 zeluletan ez zen *PROCR* eta *ROMO1* geneen arteko koadierazpenik ikusi (34A irudia), baina bai ordea, populazio mistoan (34B irudia). HEK293FT zelulen kasuan, jatorrizko zeluletan *PROCR* eta *ROMO1* geneen arteko koadierazpena ikusi zen (34C irudia), eta ez editatutako zeluletan (34D irudia).



34. irudia. *ROMO1* eta *PROCR* geneen adierazpenaren korrelazioa A) jatorrizko HCT15 zeluletan eta B) populazio mistoan; eta C) jatorrizko eta D) editatutako HEK293FT zeluletan (n = 6 baldintza guztietan). Spearman rho eta aldebiko p balioak erakusten dira.

2.4. Zelula lerroen genotipazioa

PROCR-ROMO1 eskualdeko bi TADen artean dauden eta gaixotasunari asoziatutako haplotipoen ordezkari diren 3 SNP genotipatu ziren HCT15 eta

HEK293FT zelula lerroetan. SNP horien genotipoaren eta bi zelula lerroetan ikusitako koadierazpen aldaketan artean korrelaziorik zegoen aztertu nahi izan genuen. HCT15 zelulak alelo alternatiborako homozigotoak ziren rs6060369 eta rs224371 SNPetan, eta heterozigotoak rs2104417 SNPean. HEK293FT zelulak alelo alternatiborako homozigotoak ziren rs6060369 SNParen kasuan, eta heterozigotoak beste bi SNPetan (16. taula).

16. taula. HCT15 eta HEK293FT zelula lerroen genotipoak.

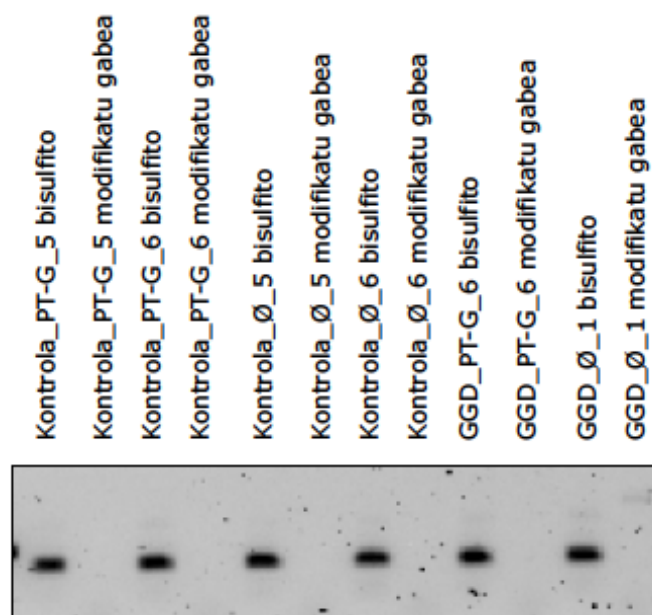
	HCT15	HEK293FT
rs6060369	GG	AG
rs224371	GG	AG
rs2104417	AG	AG

3. Metilazio patroien aldaketa akutuak EZn

DNaren metilazioa gene erregulazioan inplikaturik dagoen mekanismo epigenetiko garrantzitsua da. DNaren metilazio aberrantea zenbait gaixotasun konplexuren ezaugarri da, eta EZn estimulu desberdinen aurreko erantzunean inplikaturik egon liteke.

3.1. Bisulfito konbertsioa eta NGS metilazio-espezifikoak

Inongo konposaturik gabe eta 250 µg/ml gliadinarekin inkubatutako 7 GGD-EZ eta 8 kontrolen biopsia zatietatik, DNA genomikoa erauzi eta bisulfitoarekin tratatu zen. Bisulfito tratamenduaren efizientzia PCR bidezko amplifikazioz egiaztatzen zen; horretarako, bisulfitoz modifikatutako eta modifikatu gabeko DNA amplifikatzeko *GAPDH* hasle espezifikoak erabiliz (35. irudia).



35. irudia. Bisulfito konbertsioaren efizientziaren adibidea. Modifikatu gabeko hasleak erabiliz ez da PCR anplifikaziorik gertatzen, eta horrek laginen konbertsio egokia adierazten du. GGD = glutenik gabeko dietan zeuden EZ pazienteak; Kontrola = EZrik ez duen kontrola; PT-G = kitzikatutako biopsiak; Ø = kitzikatu gabeko biopsiak.

Sei eskualde genomiko (*DFNA5*, *HDAC4*, *HLA-B*, *SLC46A1*, *TRIM15* eta *TAP1* geneen inguruan) PCR bidez anplifikatu ziren bisulfitoarekin tratatutako DNA laginetan. Guztira, sei aplikoiak 2.023 bp eta 125 CpG posizio hartzen dituzte. PCR produktuak elektroforesi bidez aztertu ziren, eta xingolen intentsitatea Image Lab v5.2.1. programarekin kalkulatu zen, ondoren indibiduo bakoitzaren PCR produktuak kontzentrazio berdinetan nahasteko.

Anplikoi multzoak Illumina MiSeq sisteman sekuentziatu ziren. Guztira, 23.431.176 irakurketa lortu ziren, eta horietatik %71,56 konbertitutako erreferentzia sekuentziekin ondo lerrokatzen zen (17. taula) (9. eranskina).

17. taula. Eremu bakoitzeko irakurketak eta erreferentzia sekuentziarekiko lerrotatzea.

	Irakurketa guztiak	Lerrotatzea (%)	Lerrotatze egokia (%)	Irakurketak eremuka					
				DFNA5	HDAC4	HLA-B	SLC46A1	TAPI	TRIM15
Guztira	23431176			30870	5236231	1322009	155580	2138500	7622379
Bataz bestekoa lagineko (+/- SD)	781039,2 (+/- 104855,52)	72,12% (+/- 4,67)	71,56% (+/- 4,71)	1029 (+/- 354,89)	174541,03 (+/- 78410,41)	44066967 (+/-26401,09)	5186 (+/- 2931,23)	71283,33 (+/- 30295,26)	254079,3 (+/- 74969,78)

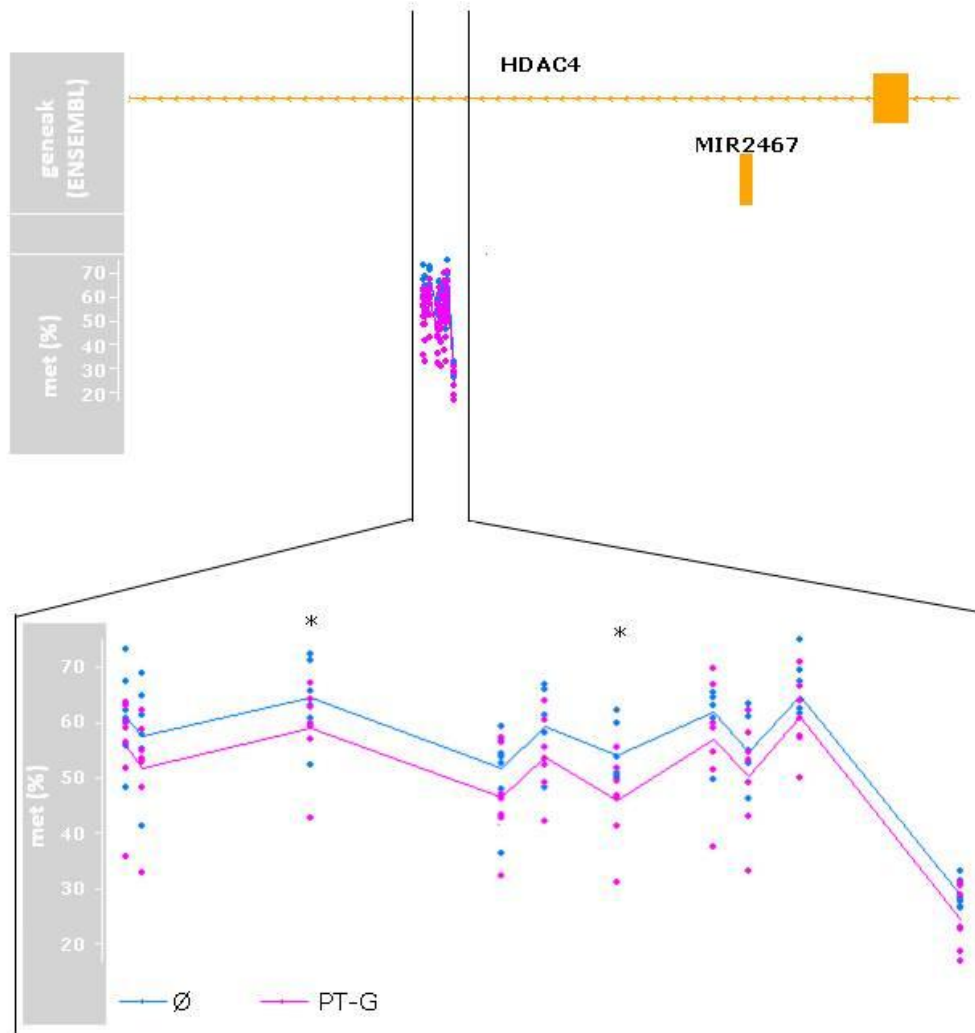
Bisulfitoaren tratamenduaren efizientzia kuantifikatzeko, ez-CpG posizioetako zitosinen konbertsio-proportzioa aztertu zen. Ez-CpG posizioetako zitosina guztiak timina bihurtu ziren eta timina modura irakurri ziren irakurketen %98,12an, lagin guztietan, bisulfito konbertsioaren efizientzia altua nabarmenduta (18. taula).

18. taula. Ez-CpG posizioetako C proportzioa lagin eta eremu bakoitzean. Kurtsiban ez-CpG posizioetan C proportzioa > %5 (balioa > 0,05). Kasu horietan, irakurketak baztertu eta eremu hori ez zen laginean aztertu. GGD = glutenik gabeko dietan zeuden EZ pazienteak; Kontrola = EZrik ez duen kontrola; PT-G = kitzikatutako biopsiak; Ø = kitzikatu gabeko biopsiak.

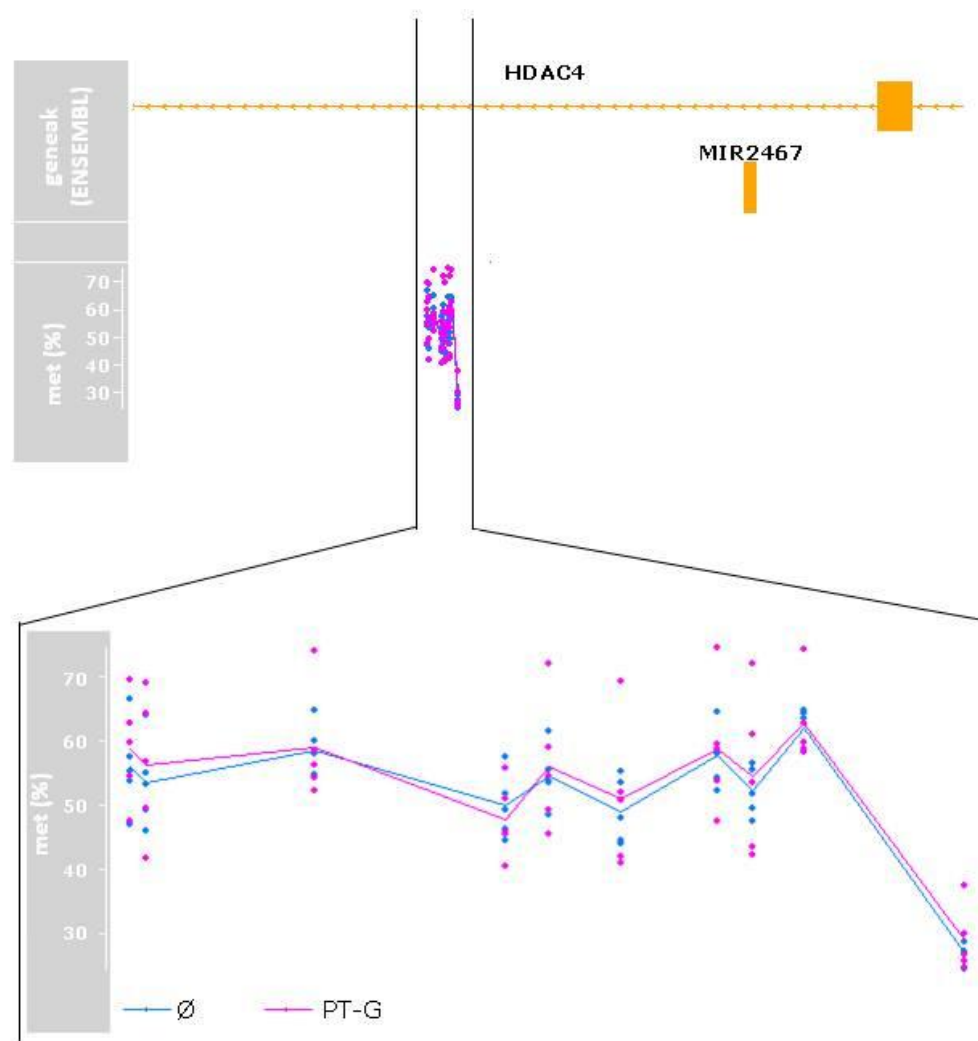
Lagina	Ez-CpG posizioetako C proportzioa					
	<i>DFNA5</i>	<i>HDAC4</i>	<i>HLA-B</i>	<i>SLC46A1</i>	<i>TAP1</i>	<i>TRIM15</i>
Kontrola_Ø_1	0,0091	0,0069	0,0115	0,011	0,0031	0,0035
Kontrola_Ø_2	0,0073	0,1334	0,0259	0,0093	0,004	0,0046
Kontrola_Ø_3	0,0069	0,0081	0,0484	0,0097	0,0118	0,0068
Kontrola_Ø_4	0,0292	0,0047	0,0041	0,0106	0,0026	0,0053
Kontrola_Ø_5	0,0131	0,0078	0,0071	0,0114	0,0069	0,008
Kontrola_Ø_6	0,0159	0,1942	0,0105	0,0092	0,008	0,0062
Kontrola_Ø_7	0,0741	0,0078	0,0326	0,0135	0,0068	0,0093
Kontrola_Ø_8	0,0091	0,01	0,0201	0,0098	0,0072	0,0038
Kontrola_PT-G_1	0,0081	0,006	0,0068	0,0115	0,005	0,0053
Kontrola_PT-G_2	0,1264	0,005	0,0168	0,0121	0,0062	0,0057
Kontrola_PT-G_3	0,0164	0,0058	0,0117	0,0166	0,0046	0,0041
Kontrola_PT-G_4	0,0358	0,01	0,0103	0,0308	0,0112	0,0181
Kontrola_PT-G_5	0,005	0,0128	0,0227	0,0095	0,0089	0,0086
Kontrola_PT-G_6	0,0077	0,008	0,0069	0,0134	0,0055	0,0062
Kontrola_PT-G_7	0,0426	0,008	0,0186	0,02	0,0066	0,0104
Kontrola_PT-G_8	0,0034	0,0019	0,0049	0,0048	0,0022	0,002
GGD_Ø_1	0,0283	0,0064	0,0142	0,0115	0,0062	0,0062
GGD_Ø_2	0,1086	0,0107	0,0228	0,0162	0,0058	0,0056
GGD_Ø_3	0,0108	0,107	0,0052	0,0073	0,0442	0,0018
GGD_Ø_4	0,0756	0,0101	0,0087	0,0149	0,0071	0,0065
GGD_Ø_5	0,0167	0,0094	0,0029	0,0102	0,0065	0,0054
GGD_Ø_6	0,009	0,0142	0,0162	0,0139	0,0069	0,0054
GGD_Ø_7	0,0087	0,1653	0,0083	0,1422	0,0042	0,0017
GGD_PT-G_1	0,1368	0,1155	0,0249	0,009	0,0351	0,0037
GGD_PT-G_2	0,0073	0,0496	0,0221	0,0113	0,0252	0,0037
GGD_PT-G_3	0,0054	0,0066	0,0317	0,014	0,0511	0,0022
GGD_PT-G_4	0,0236	0,0067	0,0175	0,0101	0,0095	0,0062
GGD_PT-G_5	0,0162	0,0671	0,0129	0,0308	0,0059	0,0051
GGD_PT-G_6	0,0116	0,0036	0,0232	0,0115	0,0035	0,0054
GGD_PT-G_7	0,0077	0,0238	0,0025	0,0086	0,0094	0,0062
Bataz bestekoa (+/- SD)	0,0142 (+/-0,0102)	0,0102 (+/- 0,0094)	0,0157 (+/- 0,0104)	0,0128 (+/- 0,0058)	0,0093 (+/- 0,0095)	0,0058 (+/-0,0031)

3.2. Metilazio aldaketak EZn

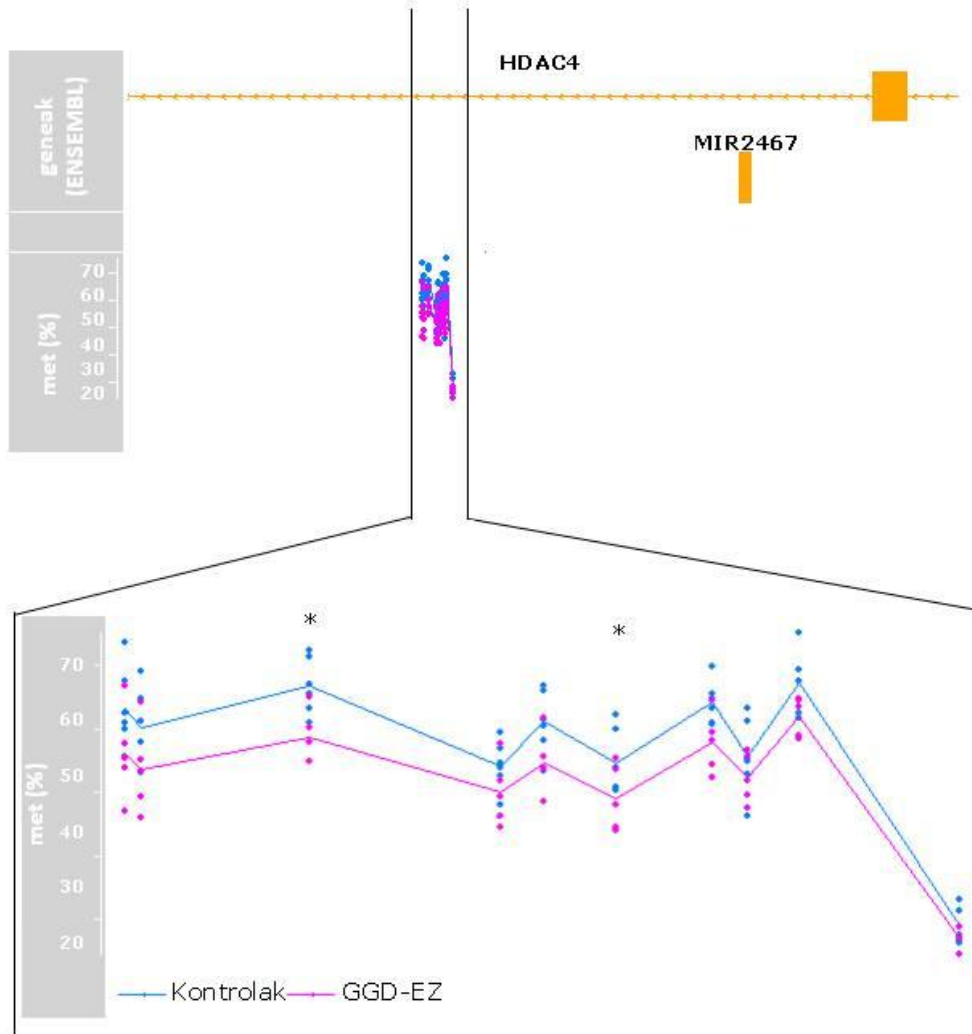
*HDAC4*ren kasuan, gliadinarekin inkubatutako kontrolen biopsietan hipometilazioa ikusi zen (36. irudia), ez ordea GGD-EZn (37. irudia). Hala ere, hipometilazio hori kitzikatu gabeko GGD-EZ biopsietan ere detektatu zen, kitzikatu gabeko kontrolekin konparatzean (38. irudia). Beraz, *HDAC4* eskualdea hipometilatua dago GDD-EZn, baita gliadinaren esposizio akaturik gabe ere. Gainera, emaitzok kontrol ez-zeliakoen hesteak ere gliadinaren aurrean erantzuten dutela erakusten dute.



36. irudia. *HDAC4* eskualdearen metilazio profila kitzikatutako eta kitzikatu gabeko kontrolen biopsietan. Izartxoak diferentzialki metilatutako posizio (DMP) esangarriak adierazten ditu (* $P < 0,05$).

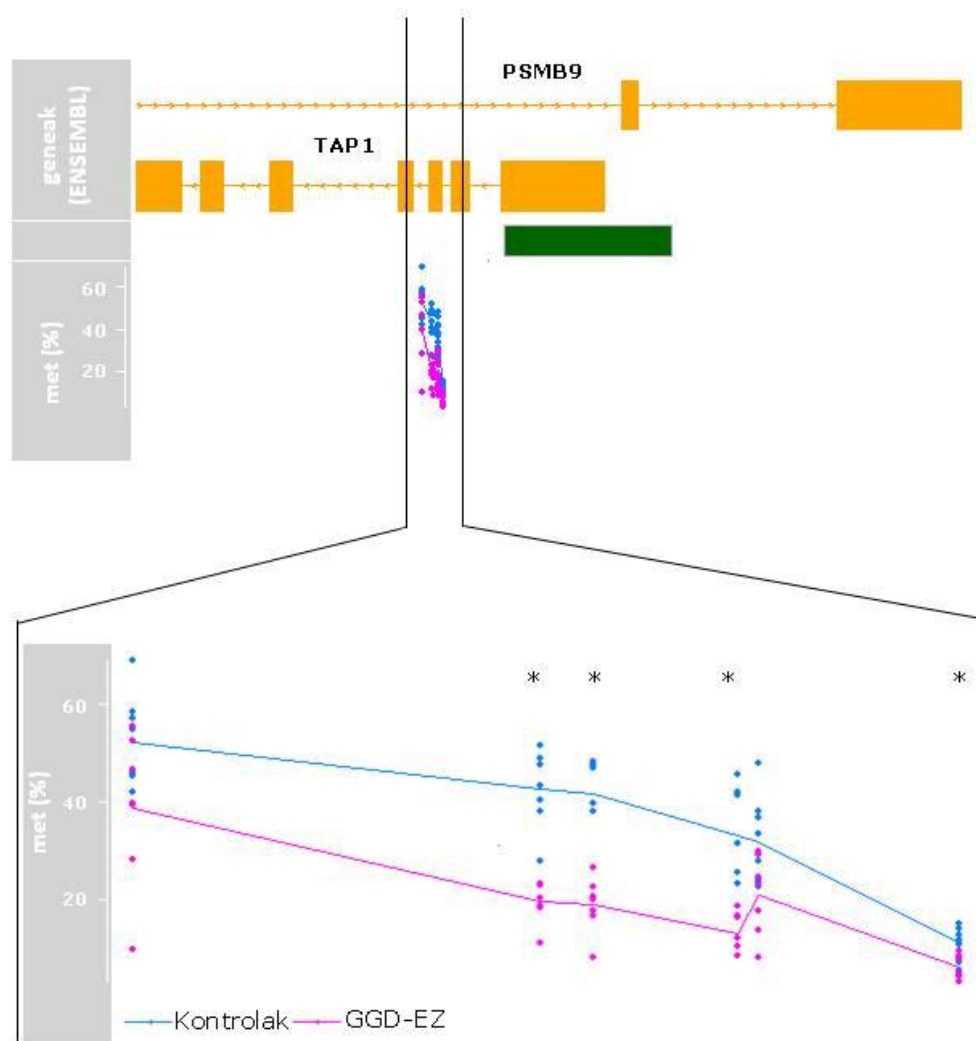


37. irudia. *HDAC4* eskualdearen metilazio profila kitzikatutako eta kitzikatu gabeko GGD-EZ biopsietan.

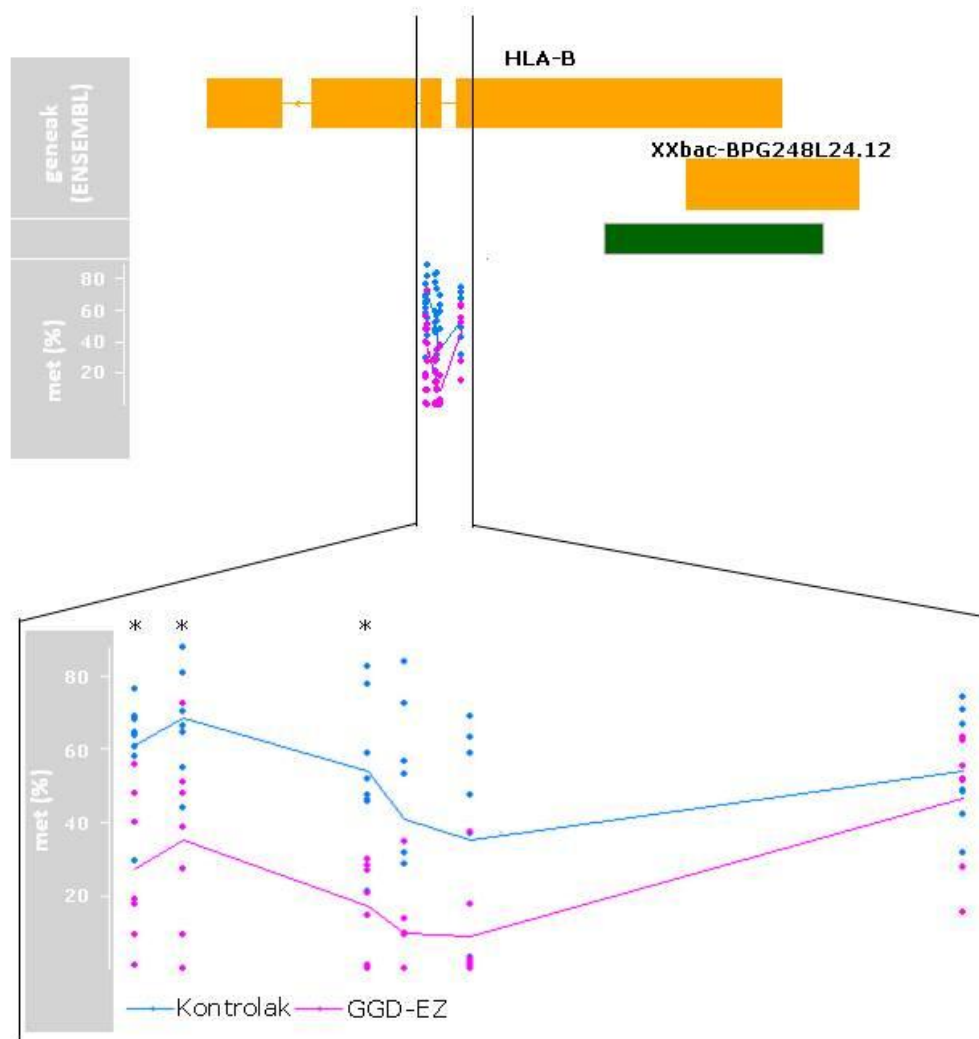


38. irudia. *HDAC4* eskualdearen metilazio profila kitzikatu gabeko kontrol eta GGD-EZ biopsietan. Izartxoak DMP esangarriak adierazten ditu (* $P < 0,05$).

Bestalde, lan honetan lortutako emaitzak aurretik gure laborategian egindako ikerketa bat (Scientific Reports, errebisiopean) balioztatzeko erabili dira. Lan horretan, EZren metilazio patroiak aztertu ziren hesteko epitelio eta immunitate zeluletan aberastutako frakzioetan. Hipometilatutako DMR bat identifikatu zen *TAP1* genean, EZ aktiboen hesteko frakzio biak kontrolen frakzioekin konparatu zirenean, eta beste DMR bat *HLA-B* genean, epitelioko zeluletan aberastutako frakzioan. Tesi honetan, lokus bakoitzeko 6 CpG posizio aztertu dira 7 GGD-EZ eta 8 kontrolako kohorte independente batean, eta emaitzak erreplikatu dira (39 eta 40. irudiak).



39. irudia. *TAP1* eskualdeko metilazio profilak kitzikatu gabeko kontrol eta GGD-EZ biopsiak konparatzean. Izartxoak DMP esangarriak adierazten ditu (* $P < 0,05$).



40. irudia. *HLA-B* eskualdeko metilazio profila kitzikatu gabeko kontrol eta GGD-EZ biopsiak konparatzean. Izartxoak DMP esangarriak adierazten ditu (* $P < 0,05$).

Eztabaida

Eritasun zeliakoa genetikoki suszeptibleak diren indibiduoetan dietako glutenaren ondorioz garatzen den gaixotasun autoimmune konplexua da. HLA-DQ2 eta HLA-DQ8 haplotipoek %40 inguruko ekarpena egiten diote gaixotasunaren arrisku genetikoari, eta ia paziente guztietan daude. Gaixotasunarekiko predisposizio genetikoa identifikatzera zuzendutako esfortzu gehienak, SNPen asoziazio ikerketetan oinarritu dira eta egun, identifikatutako aldakortasun genetikoa herentziaren %50 inguru azaltzeko gai da (Trynka et al., 2011). Hala ere, DNAREN sekuentziaren aldakortasunaz gain, genomaren beste informazio geruza batzuek EZren patogenesisian ere eragin lezakete, baina azken hauek gutxiago aztertu dira.

Zentzu honetan, mekanismo epigenetikoak gaixotasunean garrantzitsuak izan litezke; izan ere, gene adierazpena erregulatzen dute eta kanpoko estimuluekiko sentikorrak dira, ingurumen faktoreen eta faktore genetikoen arteko zubia eginez (Gupta eta Hawkins, 2015).

Tesi honetan, publikoki eskuragarri dauden omika ezberdinen datuak erabili dira EZren patogenesisian eragin lezaketan mekanismo erregulatzailer genetikoen eta epigenetikoak identifikatzeko. Esaterako, aurretik gure taldean egindako adierazpen *microarray* esperimentuen, *ImmunoChip* proiektuaren, publikatutako TADen, eta berriki sortutako metilazio eta adierazpen datuak bildu eta erabili dira EZren garapenean eragin dezaketan elementu erregulatzailer eta eremu genomikoen identifikaziorako.

Genoma osoko koadierazpena EZn

Adierazpen diferentzialaren analisi klasikoez gain, beste hurbilketa transkriptomiko batzuk ere (koadierazpen ikerketak, esaterako) erabilgarriak izan daitezke, eta momentu berean aktibo dauden eta prozesu biologiko berberetan parte-hartzen duten geneak identifikatzeko lagungarriak izan daitezke (de la Fuente, 2010). Koadierazpen analisisien bidez identifikatutako kotranskribatutako

geneek elementu erregulazaile komunak izan ditzakete, eta azken hauek ingurumenekiko elkarrekintzen ondorioz gertatzen diren adierazpen eta koadierazpen aldaketen erantzule izan daitezke, eta gaixotasunaren garapenerako beharrezkoak izan daitezke.

Sare-mailako koadierazpen azterketa ulergarri eta konparatiboak burutzeko zenbait saiakera egin da. Arlo horretako lehenbiziko lanetako bat B hepatitis eta C hepatitis birusek eragindako kartzinoma hepatozelularren arteko ezberdintasun eta antzekotasunak argitzen saiatu zen, batez ere, infekzio biralak ekarritako prozesu inflamatorioan fokatuta (He et al., 2012). Lan horrek frogatu duenez, sareetan oinarritutako Sistemen Biologiaren hurbilketak aurretik erabilitako adierazpen diferentzian eta proteinen ituetan errotutako metodoak baino indartsuagoak dira, modulu onkogeniko eta funtzionalak identifikatzeko.

Koadierazpenean oinarritutako analisiek eskaintzen duten abantailaz baliatzeko, gure taldeak koadierazpena ikertu du EZn, zenbait bidezidor biologikotan, eta gliadina gene talde batzuen koadierazpena eteteko gai dela ikusi du, bai eta erantzun koordinatua sortzeko ere (Fernandez-Jimenez et al., 2014; Plaza-Izurietan et al., 2015). Hala ere, EZn egin diren koadierazpen analisiok gene talde txikietan edo bidezidorretan egin dira, eta burutu den genoma-mailako koadierazpen ikerketa bakarrak odol periferikoa izan du aztergai (Quinn et al., 2015). Hauxe da EZ duten pazienteen duodenoko biopsietan egin den genoma osoko lehenengo koadierazpen analisia, eta gaixotasunaren garapenean eragin garrantzitsua izan dezaketen elementu erregulazaileak identifikatzeko gai izan da.

Tesi honetan, koadierazpena ikertu zen, heste zeliakoan gliadinaren eragin akutua eta epe-luzekoa aztertzen zituzten adierazpen *microarray* experimentuen datuekin (Castellanos-Rubio et al., 2008, 2010). Bi esperimenduetan koadierazitako geneen berrantolaketa gertatzen zela ikusi zen, baita erantzun akutuak koadierazpen aldaketa handiagoak sortzen zituela ere. Aipagarria da gliadinarik gabeko baldintzek, hau da, egoera osasuntsua (edo tratamendua)

imitatzen zutenek, gaixotasuna ordezkatzen zutenek baino koadierazpen modulu aldakorragoak zituztela. Honek, kitzikapen immunogenikoak osasunarekin erlazionatutako koordinazioa desregulatzen duela iradokitzen du, eta gaixotasunean sortutako koadierazpen moduluak iraunkorragoak direla. Gene erregulazioak gaixotasun konplexuen patogenesisian berebiziko garrantzia duela kontuan hartuta (Cookson et al., 2009), koordinazioaren aldaketa horietan parte-hartzen duten elementu erregulatzailerik potentzialen identifikazioa estrategia ona izan daiteke, EZren patogenesisia hobeto ulertzeko eta etorkizunean aplikazio medikoak garatzeko.

Beraz, identifikatu diren diferentzialki koadierazitako geneetan oinarrituta, TFBS eta miRNAentzako aberaste analisiak burutu ziren, koadierazpenaren eraldatzaile potentzialak aurkitzeko. TFek gene adierazpena kontrolatzen dute, eta hauen %10 inguruk giza gaixotasunetan eragina du (Bouhlef et al., 2015). TFBSen aberaste analisiak erakutsi dutenez, TF batzuek koadierazpen aldaketan erantzule izan litezke. Hala ere, DKG multzoek ez zuten emaitza esangarririk eman kasu gehienetan, eta ondorioz, EZn koadierazpena kontrolatzen duten bestelako mekanismoak egon behar dute, TFez gain.

Espero bezala, DKGekin erlazionatutako funtzio biologikoen artean, erantzun immunearekin eta apoptosiarekin erlazionatutako anotazioak aurkitu genituen. Biak EZren patogenesisian garrantzitsuak dira, EZren heste mukosan apoptosia emendatua baitago (Green eta Cellier, 2007; Shalimar et al., 2013). Horretaz gain, EZrekin erlazionatutako beste termino batzuk ere aurkitu ziren. Hauen artean, basodilatazioa, EZdun pazienteetan gutxitua dagoena (Sari et al., 2012), eta lisosoma bidezko garraioa, EZn asaldatuta dagoena (Lebreton et al., 2012), topa ditzakegu. Honek guztiak, gliadinak induzitutako koadierazpen aldaketek aipatutako prozesu biologikoen funtzionamendu egokian eragina izan dezaketela iradoki du.

Jatorrizko gene moduluak hondotzat hartuta, zenbait GO anotazio identifikatu zen. Horietako batzuk EZrekin erlazionatuta zeuden aurretik; hala nola, heste iragazkortasunean eragina duen zelula-zelula atxikitzea (Jauregi-Miguel et al., 2014), bitartekari inflamatorioak (proteina-kinasen kaskada, hain zuzen) (Yao et al., 2018), eta apoptosia (Shalimar et al., 2013). Jatorrizko moduluetatik DKGak kentzean, ez zen GO anotazio esangarririk geratu. Emaitza horiek konfirmatzen dute, identifikatu ziren moduluatarako definitutako funtzioetan, DKGek duten garrantzia. Era berean, erabilitako aberaste analisisien emaitzen zentzu biologikoa indartzen dute.

Lau TFk adierazpen aldaketak azaldu zituzten EZ zuten pazienteetan. Berezko immunitate-sistemaren eta immunitate-sistema egokitzailaren erantzunekin erlazionatuta dauden IRF1 eta CREB1 TFak gainadierazita zeuden paziente aktiboetan (Guo et al., 2010; Wen et al., 2010). Honako hau da EZn *CREB1*en gainadierazpena erakutsi duen lehenengo lana, baina *IRF1*en emendioa EZn aurretik ikusi izan da (Lahdenperä et al., 2011). Bestalde, *NFKB1*en gainadierazpena bat dator NFκB bidezidorrak EZn duen inplikazioarekin (Fernandez-Jimenez et al., 2014). *ELK1*, hesteko iragazkortasunean eragiten duen TFa (Al-Sadi et al., 2013), gainadierazita zegoen EZ aktiboan. Jakina da hesteko iragazkortasuna EZren ezaugarri garrantzitsua dela (Jauregi-Miguel et al., 2014). Honetaz gain, aztertutako TFen ituak diren gene gehienek TFen adierazpen patroia berberak jarraitzen zituzten.

DKGak gliadinaren kitzikapenak sortutako koadierazpen aldaketak aztertuta identifikatu zirenez, gliadinak TF hautagaien proteina mailetan eraginik ote zuen jakin nahi izan genuen. Gliadinak kitzikatutako C2BBe1 zeluletan CREB1 eta IRF1en nukleorako translokazioaren emendioa ikusi genuen. Berriki burututako lan batean ikusi dutenez, IRF1 gainadierazita dago EZ aktiboan, hesteko epitelio zelulak γ -interferoiarekin zein gaixotasunari asoziatutako bakterioekin kitzikatzean (Pietz et al., 2017). Emaitza horiek, IRF1en aktibitatearen emendioa konfirmatzen dute gaixotasuna bultzatzen duten egoeretan, eta gure emaitzak

berresten dituzte. Bestalde, EZn ez dago CREB1i buruzko aurretiko daturik, baina beste gaixotasun batzuekin erlazionatu izan da; koloneko minbizian esaterako, CREB1ek *RRM2* tumore-gidaria aktibatzen duela ikusi da, edo gliobastoman adibidez, CREB1ek TGF β 2ren indukzioaren bitartekari modura parte-hartzen du (Fang et al., 2016; Rodón et al., 2014). Nukleorako translokazioaren emendioak gliadinak itu geneen gainadierazpenean eragin dezakeela iradokitzen du, horretarako TFak aktibatuz, baita zeliakoak ez diren indibiduoetan ere.

Emaitza horiek TF hautagaien eta beraien itu geneen arteko elkarrekintzak ikertzerantz eramanez gintuzten, betiere gliadinaren kitzikapenak izan zezakeen eragina aztertzeko. Helburu horrekin, ChIP-PCR burutu zen C2BBel zeluletan. Gaixotasunean itu gene batzuen adierazpen aldaketak ikusi bagenituen ere, ez genuen ez CREB1 ez IRF1en eta beraien itu geneen arteko elkarrekintzetan aldaketarik ikusi, gliadinarekin *in vitro* kitzikatutako zeluletan. Hala ere, DKG multzoetatik hautatutako itu geneekiko lotura konfirmatu ahal izan genuen IRF1en kasuan. Horrela, koadierazpenari esker, IRF1en itu berriak definitu ditugu, nahiz eta gure eredu mugatua izan, erabilitako zelulak zeliakoa ez den minbizi ehun batetik eratorriak direlako. Testuinguru honetan, lan honek gaixotasun konplexuak aztertzeko, eta kasu honetan EZrako, dauden *in vitro* eredu urritasuna ere azpimarratzen du. Hala ere, C2BBel zelula lerroa da egun dagoen enterozito xurgatzaileen *in vitro* eredurik onena, eta bere erabilera zabala izan da gure intereseko gaixotasunean (Rauhavirta et al., 2011).

Beste alde batetik, zenbait miRNA ere identifikatu ziren aberaste analisietan. miRNAk gene adierazpena kontrolatzeko gai dira post-transkripzio mailan, eta proteina-kodetzaileak diren geneen %60ren itzulpena kontrola dezakete (Esteller, 2011). RNA ez-kodetzaileek, miRNAk barne, gene adierazpena molda dezakete, eta aurretik EZrekin erlazionatu izan dira (Castellanos-Rubio et al., 2016; Felli et al., 2017). Zenbait ikerketa lanek miRNAen asaldurak deskribatu dute EZn (Buoli Comani et al., 2015; Capuano et al., 2011; Magni et al., 2014; Vaira et al.,

2014), eta baliteke gaixotasunaren patogenesia eragina izan eta diagnosirako erabilgarriak izatea.

Balioztatze biologikorako 9 miRNA aukeratu ziren, eta haien adierazpena aztertu zen, EZdun paziente aktibo eta tratatu, zein kontrolen duodenoko biopsietan. miRNAen forma primarioak aztertu ziren; izan ere, RNA kantitate handia behar da miRNA helduak aztertzeko, eta biopsietatik lortu zen RNA mugatua zen. Edozein modutan, bi miRNA helduen adierazpena ere neurtu genuen, eta beste kasu batzuetan ikusi den moduan, miRNA helduek eta pri-miRNAek adierazpen patroi berdinak zituzten (Powrózek et al., 2018). Gainera, pri-miRNAek ituen errepresioan ere eragin dezaketela ikusi izan da (Trujillo et al., 2010; Yue et al., 2011), gure prozedurarekin bat. Azkenik, aurretik aztertu gabeko 6 pri-miRNA EZn azpiadierazita zeudela ikusi genuen, kontrolekin konparatuta.

miRNA bakoitzaren adierazpenean ikusten diren aldaketek zenbait mRNAren eragin dezaketela uste denez (Chen et al., 2016; Erson eta Petty, 2008), azpiadierazitako pri-miRNAen ituen geneen adierazpena ere aztertu zen, gure taldeak burututako RNA-seq esperimentuen datuetan. Hogeita bat ituen genetan mRNA mailen aldaketa esangarria ikusi zen gaixo aktiboetan. Orokorrean, ituen gainadierazita zeuden EZdun indibiduen epitelio zeluletan, eta hori pri-miRNAen azpiadierazpenarekin koherentea zen. Bestalde, frakzio immunean gainadierazitako eta azpiadierazitako geneak aurkitu ziren. Beraz, epitelio zeluletan aberastutako frakzioa miRNAen adierazpen aldaketan aurrean suszeptibleagoa da.

Lau ituen geneak erakutsi zuten adierazpen aldaketak duodenoko zelula frakzio bietan; *JAG1*, *SPI* eta *CYP2J2* geneek patroi bera jarraitzen zuten bi zelula frakzioetan, eta *ARL4C* geneak aldiz, patroi ezberdina zuen bakoitzean. Horren arabera, erregulazioaren zelula-espezifikotasuna kontuan hartzekoa izan daiteke.. Hala ere, miRNAen zenbait ituen geneak mRNA mailen aldaketak erakutsi badituzte

ere, ezin bazter dezakegu ikusitako aldaketak miRNAetatik independenteak diren beste faktore batzuetatik datozenik.

Gure behaketak eta haien atzean dauden mekanismo zelularrak guztiz ulertzeko ikerketa mekanistikoak beharrezkoak izango dira. Hala ere, koadierazpen ikerketak adierazpen diferentzialen azterketen osagarri izan daitezke, bai eta giza gaixotasunetan garrantzitsuak diren elementu erregulatzailen identifikaziorako erabilgarri ere, gure lanak erakutsi duen moduan. Gainera, gure metodologia publikoki eskuragarri dauden genoma osoko datuetan erabil daiteke, eta esperimentuen emaitzak berriro aprobetxatu. Hala ere, erabilitako hurbilketa DKGak, TFak eta miRNAk identifikatzeko erabilgarria izan bada ere, *in vitro* lortutako emaitzak ez dira ikusitako koadierazpen alterazio guztiak azaltzeko gai izan.

TADak EZn

Gene erregulazioa kromosomen egituraren eraginpean lotua dago eta TADak kromatinaren 3 dimentsioko konformazioan parte hartzen duten elementu garrantzitsuak dira (Lupiáñez et al., 2016; Sexton et al., 2007). TADak kromatinaren unitate funtzionalak dira, kromatina-kromatina interakzioetan aberatsak dira, eta zelula mota ezberdinen eta ugaztun espezie ezberdinen artean kontserbatuak daude (Dixon et al., 2012). CTCF eremuak domeinu horien alboetan kokatzen dira, horrela TADen mugak definitzen dituzte eta kromatinaren begizten eraketan eragina dute (Ong eta Corces, 2014).

TADen mugetako alterazioak zenbait prozesu patogenikotan deskribatu izan dira. Gorputzadarren hiru malformazio mota TADen mugetako hausturarekin asoziatu dira; bertan, delezioek, duplikazioek eta inbertsioek *WNT6/IHH/EPHA4/PAX3* lokusa barneratzen duen TADaren egitura aldatzen dute (Lupiáñez et al., 2015). TADen haustura minbizi zeluletan ere ikusi izan da sarritan. Glioman, esaterako, kohesina eta CTCF lotura-guneetan gertatzen den hipermetilazioak CTCFekiko

lotura murrizten du, eta TADen mugen inaktibazioa eragin. Horrek *PDGFRA* moduko minbizi-gidarien aktibazioa dakar, jatorrizko TADetik kanpoko areagotzaileen eskuhartzeagatik (Flavahan et al., 2016).

Gainera, TADak distantzia luzeetako erregulazioarako kromosomen unitate estruktural erabakigarriak dira (Dekker eta Heard, 2015). TADak gene adierazpenarekin erlazionatzen zituzten lehenengo ikerketetako batean, TADak modu koordinatuan erregulatzen ziren gene taldeekin ere lerrokatzen zirela ikusi zuten, eta TAD bereko geneen adierazpen mailak TAD ezberdinetakoenak baino korrelatuago daudela erakutsi (Nora et al., 2012).

Aurretik azaldu den bezala, koadierazitako geneek gaixotasunaren etapa ezberdinetan ikusten diren adierazpen eta koadierazpen aldaketen erantzule diren elementu erregulatzailer komunak izan ditzakete. Zentzu horretan, TADek eginkizun garrantzitsua izan dezakete EZn; izan ere, gene erregulazioan garrantzitsuak direla uste da. Doktorego tesi honetan, koadierazpen aldaketen atzean egon daitezkeen TADen haustura eta fusioak identifikatzeko hurbilketa berria proposatu da. Laburbilduz, gene adierazpenaren koordinazioan eragina izan lezaketen eskualde erregulatzailer hautagaiak aukeratu dira, lan honetan lortutako koadierazpen emaitzak eta publikatutako TADen eta EZri asoziatutako SNPen datuak erabiliz.

TADen egituran gertatzen diren asaldurek, EZn ikusitako koadierazpen aldaketak azal litzaketela, parte batean behintzat, hipotetizatu genuen. Beraz, koadierazpen patroiak definitu ziren RNA-seq datuak erabilita, eta kontserbatutako TADekin lerrokatzen ziren koadierazitako geneak identifikatu (Dixon et al., 2012). Indibiduo osasuntsuetan, koadierazitako 739 gene 486 TADetan kokatu ziren, eta EZ aktiboaren kasuan, 613 gene 430 TADetan. EZn eragina izan zezaketen TADak lehenesteko, gaixotasunari asoziatutako SNPak barneratzen zituzten kasuak aukeratu genituen (Garcia-Etxebarria et al., 2016). Identifikatutako eskualde guztietatik, EZri asoziatutako SNPak zeramatenak hautatu genituen,

elkarren ondoan zeuden TADen artean edo TADen barruan. Arrisku genetikoa emendatzen duten aleloek domeinuak apurtu ditzaketela proposatzen dugu, TADen arteko mugak mutatu eta elkarren ondoan dauden bi TADen fusioa eragin, edo TAD bat haustu, bi domeinu berri sortzeko.

Kromatinaren 3D egituraren aldaketetan intsertzioak, delezioak edo inbertsioak ikusi dira, orokorrean. Hala ere, TADen mugetan eta CTCF lotura-guneetan aldaketak sortzeko gai diren nukleotido bakarreko aldaketak ere aurkitu dira. Esate baterako, CTCF/kohesina lotura-guneetan ohiko mutazioak topatu izan dira, eta minbiziaren genoma ez-kodetzaileko mutazio puntual garrantzitsuak dira (Katainen et al., 2015). Beste lan batean, *NOTCH1* genea barneratzen zuen TADaren mugako CTCF lotura-gunean mutazioak identifikatu zituzten obarioko minbizian. Nukleotido aldaketa horiek *NOTCH1*en desregulazioa dakarte, seguruenik TADaren hausturaren ondorioz gertatzen den areagotzaileekiko elkarrekintzaren asalduragaitik (Ji et al., 2016). Berrikiago egindako lan batean, SNP batek eragindako CTCF lotura-gunearen apurketa gripe larriarekin asoziatu da. Laburbilduz, autoreek rs34481144 SNP-aren genotipoak CTCFren lotura baldintzatzen duela ikusi zuten, eta hori *IFITM3* lokusaren inguruan zeuden infekzio biralaren kontrako erantzunean parte hartzen duten geneen adierazpenarekin korrelazionatuta zegoen (Allen et al., 2017).

Tesi honetan, balioztatze biologikoetarako bi eskualde aukeratu genituen; *HSCB-XBPI* eskualdea (TAD baten haustura) eta *PROCR-ROMO1* eskualdea (bi TADen fusioa). Hautatutako bi eskualdeek EZrekin asoziatutako SNP bat baino gehiago zeramaten, eta eskualde bakoitzean gertu zeuden bi geneen artean ikusitako koadierazpen aldaketetan inplikatuak egon litezke.

Horietako hiru gene EZrekin asoziatu izan dira aurretik: *PROCR* Th17 erantzunaren erregulatzaile negatiboa da, eta Th17 zelulen gainazalean adierazten da. *PROCR*ren adierazpenaren murrizpenak Th17ren patogenizitatearen emendioa eta entzefalomiелitis autoimmune experimentalaren (ingelesez

experimental autoimmune encephalomyelitis edo EAE) areagotzea dakartza (Kishi et al., 2016). Azken hori gure taldeak publikatutako *microarray* datuetan ikusitako EZ aktibodun eta GGD-EZ pazienteen arteko adierazpen desberdintasunekin bat dator (Castellanos-Rubio et al., 2008). Gainera, aurretik ere, Th17 erantzuna EZn emendatuta dagoela ikusi da (Castellanos-Rubio et al., 2009; Cicerone et al., 2015).

ROMO1 mitokondiako oxigeno espezie erreaktiboan (ROS) askapenerako erregulatzaile nagusia da (Bae et al., 2011). Minbizi zelulen hedapenerako beharrezkoa dela deskribatu da, eta koloneko minbizian pronostiko txarrarekin asoziatu da (Kim et al., 2017). *ROMO1*ek NFκBren aktibazioan eragina duela ere ikusi da, eta bidezidor hau aktibatuta dago EZn (Fernandez-Jimenez et al., 2014). *ROMO1*en gainadierazpenak p65en translokazio nuklearra sustatzen du, eta NFκBren aktibazio konstitutiboaren erregulatzailea izan liteke tumore zeluletan (Chung et al., 2014). Beraz, *ROMO1* itu terapeutiko posiblea izan liteke NFκB desregulatuta duten gaixotasunetan (Lee et al., 2015).

Amaitzeko, *XBPI* berezko erantzun immunean parte hartzen duen transkripzio faktorea da (Jheng et al., 2012). Gainera, Toll-moduko hartzaileak (ingelesez *toll-like receptor* edo TLR) eta erretikulu endoplasmaticoaren estresak eragindako aktibazioaren osteko IL-6 moduko hantura faktoreen ekoizpenerako beharrezkoa dela aurkitu zuten (Gargalovic et al., 2006; Martinon et al., 2010; Savic et al., 2014; Toosi et al., 2012). Horretaz gain, beste gaixotasun batzuetan ere parte hartzen du; hala nola, 2 motako diabetesean (Özcan et al., 2006), minbizian (Jin et al., 2016), eta hesteetako hanturazko gaixotasunean (ingelesez *inflammatory bowel disease* edo IBD) (Kaser et al., 2008).

Bi eremu hautagaiak lerro zelularretan aztertu ziren DNAsarentzat eskuragarriak ziren eta ondorioz, kromatina irekia zegoen ikusteko. Gainera, eraldaketa egonkorrak dituen lerro zelular eredu bat ere prestatu genuen, ikerketa gehiago egin ahal izateko. Aurretik azaldu den bezala, C2BBel lerro zelularra EZn

sarritan erabili da *in vitro* eredu modura; baina, zelula hauek editatzeko gure laborategian aurretik egindako saiakerek, horretarako zailtasunak azaleratu zituzten. Beraz, HCT116, HCT15 eta HEK293FT epitelio zelulak erabili genituen hurrengo esperimentuetarako. HCT116 eta HCT15 koloneko minbizia ikertzeko erabiltzen diren kolonetik eratorritako zelula lerroak direlako hautatu ziren (Bessa et al., 2018; Gong et al., 2018). Gainera, HCT116 zelula lerroa EZn asaldatuta dauden lotura hertsia eta hesteko iragazkortasuna (Jauregi-Miguel et al., 2014) ikertzeko ere erabili izan dira (Kolodziej et al., 2011). Bestetik, HEK293FT zelula lerroa giza giltzurrun zelula enbrionarioetatik eratorria da, eta askotan erabili da CRISPR-Cas9 teknikarekin genomak aldaketak sortzeko (He et al., 2016).

Ondoren, TADen muga izan zitezkeen CTCF lotura-guneak identifikatu nahi izan genituen (Ghirlando eta Felsenfeld, 2016). DHS eremuen barruan zeuden CTCF lotura-guneak aukeratu genituen, eremu horiek proteinak lotzeko eskuragarri dauden kromatinaren eskualdeak baitira (Gross eta Garrard, 1988). Kromatina DNAsaz digeritu eta gero burututakoqPCRk erakutsi zuenez, aukeratutako bi eremuak DNAsarentzat eskuragarritasun handiko guneak zirela ikusi genuen aztertutako lerro zelularretan. Orokorrean, CTCFren lotura transkripzio aldetik aktiboak diren lokusekin erlazionatu izan da (Batlle-López et al., 2015), baina TADen mugetako CTCF lotura-guneen DNAsa I-ekiko sentikortasun seinaleak TADen barnekoenak baino ahulagoak direla ere ikusi izan da (Hong eta Kim, 2017).

Epitelioko lerro zelularretan CTCF lotura-guneetan aldaketa egonkor eta zehatzak sortzeko CRISPR-Cas9 teknologia erabili genuen. Teknika horrek editatutako zeluletan epe-luzeko ikerketak egitea ahalbidetzen du, eta fenotipoak sakon aztertzeko aukera ematen du. Ikerketa batzuetan CTCF lotura-guneak ezabatu dituzte CRISPR-Cas9 teknika erabiliz, eta zenbait kasutan, CTCF lotura-gune baten delezioa nahikoa zen TADen mugan aldaketak sortzeko (Lupiáñez et al., 2015), ez ordea beste zenbait kasutan, CTCF loturaz gain, beste mekanismo

batzuk ere parte hartzen dutela TADen mugen ezarpenean agerian utzita (Barutcu et al., 2018).

Lan honetan, HEK293FT zelula lerroan eremu baten klon mutanteak lortu ziren, eta populazio mistoa HCT15 lerroan. Zehazki, *PROCR* eta *ROMO1* geneak barneratzen dituzten bi TADen artean dagoen DHS eta CTCF lotura-gunea den 377 bp-ko zatia delezionatu zen. Delezioak HEK293FT zelula mutanteetan *PROCR* eta *ROMO1* geneen adierazpenean aldaketak eragiten ditu, bai eta koadierazpen aldaketak HEK293FT zelula mutanteetan eta HCT15 populazio mistoan. Hala ere, bi lerro zelularretan ikusitako koadierazpen aldaketak kontrajarriak dira; koadierazpena apurtu egiten da HEK293FT zeluletan, eta HCT15 zeluletan aldiz, koadierazpen berria ezartzen da.

Emaitza horiek ikusita, bi lerro zelularrek genotipo desberdinak zituzten galdetu genion gure buruari. Izan ere, genotipo ezberdina izateak TADen egituraren eragina duten CTCF loturen aldakortasunak gertatzea azal dezake, bai eta bi lerroetan ikusitako koadierazpen patroia ezberdinak ere. Helburu horrekin, rs6060369, rs224371 eta rs2104417 SNPak genotipatu ziren HCT15 eta HEK293FT lerroetan. Hiru SNPok, *PROCR-ROMO1* eskualdean, delezionatutako CTCF lotura-eremutik kanpoan eta bi TADen artean dagoen SNPen haplotipoa ordezkatzeko dute, eta EZri asoziatuta dagoela ikusi da. Bi lerro zelularrek genotipo ezberdinak zituzten, baina ezin dugu aldakortasun hori bi lerroen arteko koadierazpen patroia ezberdintasunak azaltzeko nahikoa denik ondorioztatu.

Hala ere, rs6060369 SNParen genotipoa *PROCR* genearen adierazpen aldaketekin asoziatuta dagoela ikusi da (<https://www.gtexportal.org/home/>), eta eQTL (*expression quantitative trait locus*) dela esan daiteke. HCT15 eta HEK293FT lerro zelularretan, rs6060369 SNParen genotipoa desberdina dela ikusi dugu; beraz, lerro zelular horietan ikusitako koadierazpen patroien arteko ezberdintasunak eQTL horren ondorioak izan litezke. Gainera, kontuan hartzekoa

da delezioa homozigosian zeramaten HEK293FT zelulak lortu direla, baina soilik populazio mistoa HCT15en kasuan. Elementu guzti horiek HEK293FT eta HCT15 lerroetan ikusitako koadierazpen patroiei ezberdinetan eragina izan lezakete.

Laburbilduz, hautatutako eremuak garrantzitsuak izan litezke EZren patogenezian; izan ere, *PROCR* eta *ROMO1* geneen inguruko bi TADen arteko eskualdean (*PROCR-ROMO1* eremua) izandako mutazioek eragina dute gene biren arteko koadierazpenean, eta gene horiek erlazionatuta daude gaixotasunean asaldatuta dauden bidezidorrekin. Hala ere, egun ditugun muga teknikoak eta *in vitro* eredu perfektuaren falta kontuan izanda, beharrezkoa ikusten dugu metodologia berrien garapena, adierazpen genikoen, TADen eta GWA azterketen seinaleen artean egon daitezkeen harremani buruzko ebidentzia gehiago lortzeko (Pervjakova eta Prokopenko, 2017).

Metilazio patroien aldaketa akutuak EZn

DNAREN metilazioak, gene adierazpenaren modulazioaren bitartez, paper garrantzitsua jokatzen du zenbait prozesu biologikoetan; hala nola, garapenean eta gaixotasunetan (Wan et al., 2015). Sarritan ikusi dira DNAREN metilazio patroiei aberranteak zenbait gaixotasunetan, bereziki minbizian (Jones et al., 2007). Hala ere, IBD edo T1D bezalako immunitateko gaixotasunetan ere eragina duela jakina da (Agardh et al., 2015; McDermott et al., 2016). Azkenik, hainbat ingurumen faktorek, kutsadura (Madrigano et al., 2011), tenperatura eta hezetasuna (Bind et al., 2014), edo erretzea kasu (Zeilinger et al., 2013), DNAREN metilazioaren aldaketekin erlazionatu da.

EZri dagokionez, pazienteen duodenoko biopsietan burututako DNAREN metilazio azterketak NFκBekin erlazionatutako zenbait generen promotoreetako metilazio aldaketak antzeman zituen (Fernandez-Jimenez et al., 2014). EZ

aztergai duen tesi honetan, gliadinarekiko 4 orduko esposizioa metilazio aldaketak sortzeko nahikoa ote zen jakin nahi genuen. Helburu horrekin, gure taldeak aurretik burututako adierazpen (Castellanos-Rubio et al., 2008, 2010) eta metilazio (Scientific Reports, errebisiopean) ikerketen datuetatik aukeratutako sei eskualde hautagaietan, metilazio azterketak egin ziren GGD-EZ eta kontrolen biopsia zatietan, gliadinarekin kitzikatuta eta kitzikatu gabe.

Diferentzialki metilatutako zitosen aurkitzeko, PCR metilazio-espezifikoak erabili zen bisulfitoarekin tratatutako DNA amplifikatzeko, eta eremu hautagaien aplikoiak NGS bidez sekuentziatu ziren. Irakurketak erreferentziako genomarekin erkatu ziren, eta haien %71,56 egoki lerrokatu zen. Orokorrean, lerrokatzea maila %40 ingurukoa izaten da (Krueger eta Andrews, 2011; Tran et al., 2014), beraz, lan honetan lortutako lerrokatze proportzioa ohikoa baino dezente handiagoa da. Gainera, lan honetan, ez-CpG zitosen %98,12 timina modura detektatu ziren, lortutako bisulfito efizientzia altua agerian utzita (Holmes et al., 2014).

Gliadinaren kitzikapen akutuak DNAREN metilazioan nola eragiten duen ikusteko, gliadinarekin kitzikatutako eta kitzikatu gabeko GGD-EZ eta kontrolen biopsia zatiak konparatu ziren. GGD-EZ laginek informazio asko eskaintzen dute; izan ere, paziente horien alterazio genomikoek heredatutako edo haurtzaroan oso goiz hartutako (EZ garatu aurretik) predisposizio genetiko edo epigenetikoekin erlazionatutako ezaugarri konstitutiboak isla litzakete. Bestela, GGDk ehunak osa ditzakeen arren, gaixotasunarekin erlazionatutako zenbait metilazio aldaketa itzultzezinak izan daitezke.

*HDAC4*an izan ezik, ez zen metilazio akutuaren bestelako aldaketa esangarririk topatu. *HDAC4* prozesu inflamatorioekin erlazionatuta dagoen genea da (Yang et al., 2018). Eremu hori hipometilatuta zegoen gliadinarekin kitzikatutako kontrolen biopsia zatietan, baina ez ordea, GGD-EZ pazienteetan. Aurretiko marka genetiko edo epigenetikoak bertan badaude, zaila izaten da kitzikapen

akutuaren ondoriozko metilazio aldaketak aurkitzea. Horixe izan daiteke, hain zuzen ere, GGD-EZ pazienteetan gliadinarekin kitzikatuta eta kitzikatu gabeko biopsien artean metilazio aldaketarik ez detektatzeko arrazoia; izan ere, *HDAC4* hipometilatua zegoen GGD-EZ paziente horietan. Emaitza horien arabera, *HDAC4*an identifikatutako metilazio aldaketak konsitutiboak (heredatuak edo EZ pairatzeko joera dutenetan bizitzan oso goiz hartuak) izan daitezke, edo bestela, normaltasunera itzultzen ez diren gaixotasunarekin erlazionatutako aldaketak, pazienteek GGD jarraitu arren (marka epigenetiko itzulezinak). Aipatzekoa da, *HDAC4* tesi honetan identifikatutako IRF1 TF hautagaiaren itu genea dela, eta gainadierazita dagoela gaixotasun aktiboan. Esangarria ez bada ere, bere adierazpenak patroi berdina jarraitzen du GGD-EZ pazienteetan, kontrolekin konparatuta, eta hori *HDAC4* inguruan ikusitako metilazio aldaketekin koherentea da.

Gainera, gure emaitzek erakutsi dutenez, kitzikapen akutuaren ondorioz *HDAC4*an ikusten diren metilazio aldaketak kontroletan jazotzen dira, baina aldaketok ez dira iraunkorrak, eta indibidio ez-zeliakoetan marka epigenetikoa itzulgarria da. Hauxe da pertsona ez-zeliakoen hestean gliadinaren kitzikapen akutuak metilazio aldaketak sortzen dituela erakusten duen lehenengo lana. Aurretik egindako beste lanek ordea, gliadina ez-zeliakoetan mRNA aldaketak sortzeko gai dela erakutsi dute (Fernandez-Jimenez et al., 2014).

Gure taldearen arestiko lan batean, metilazio patroiak aztertu ziren EZ aktibo eta kontrolen hesteko epitelio eta immunitate zeluletan aberastutako frakzioetan (Scientific Reports, errebisiopean). Lan horretan, beste zenbait metilazio aldaketen artean, *HLA-B* eta *TAPI* geneen promotoreak EZn hipometilatuta zeudela ikusi genuen. Zehazki, *TAPI* bi zelula populazioetan zegoen hipometilatua, eta *HLA-B* epitelioko zeluletan. Gainera, RNA-seq emaitzek gene adierazpena zelula mota desberdinetan ikusitako metilazio aldaketekin bat zetorrela erakutsi zuten.

Tesi honetan, gliadinarekin kitzikatu gabeko *HLA-B* eta *TAP1* geneen promotoreen hipometilazioa konfirmatu dugu GGD-EZ pazienteetan, EZn bi urte baino gehiagoz GGD jarraitu arren, irauten duen alterazio konstitutiboa edo itzulezina dela iradokiz. Biak, *HLA-B* eta *TAP1*, EZrekin erlazionatu izan dira aurretik (Bratanic et al., 2010; Pietz et al., 2017). TAP1 HLA I motako gainazaleko peptido garraiatzailea da, eta azpiadierazita dago koloneko minbizian. Azpiadierazpen hori promotoretik gertuko eskualdeko posizioetan ikusitako metilazioarekin alderantziz lotuta dago (Ling et al., 2017). Metilazio patroietan ikusitako aldaketek immunitate-sistema egokitzailearen erantzun okerra erakutsi lezakete.

AZKEN HAUSNARKETAK

Tesi honek informazio genomikoaren maila ezberdinak kontuan hartzearen garrantzia azpimarratzen du, eta TF, miRNA, TAD eta DNAREN metilazio moduko elementu erregulatuak EZren sorreran eta garapenean eragina izan dezaketela iradokitzen du, bai eta beste gaixotasun konplexu batzuen ere.

Gaur egun arte, EZn burutu diren transkriptoma azterketa gehienak, adierazpen aldaketetan fokatu dira (Castellanos-Rubio et al., 2008, 2010). Tesi honetan, genoma osoko koadierazpen aldaketak hartu dira kontuan transkriptoma aztertzerakoan. Horri esker, EZn gertatzen diren koadierazpen patroien asalduren ikuspegi orokorra lortu dugu.

Alde batetik, TFek azal litzakete koadierazpen aldaketa horiek. Lan honetan ikusi dugu, CREB1 eta IRF1 inplikaturik egon daitezkeela EZn identifikatutako hainbat koadierazpen aldaketetan; izan ere, bi TF horien zenbait iturri generen adierazpena aldatuta zegoen, eta beraien nukleorako translokazioaren emendioa ikusi genuen gliadina bidezko *in vitro* kitzikapenaren ondorioz.

Bestetik, kromatinaren egituraketa ere izan daiteke gaixotasunean ikusitako koadierazpen aldaketen beste erantzuleetako bat. Lan honetan, kontserbatutako TADen, eta EZrekin erlazionaturiko DKGen eta SNPen lerrokatzea aztertu dugu, EZn inplikaturik egon daitezkeen TADen berrantolaketak aurkitzeko. Hori da *PROCR-ROMO1* eskualdearen kasua, non TAD mugen delezio partzialak inguruko koadierazpen patroia aldatzen duen.

DNAREN metilazioa gene erregulazioan eragina duen mekanismo epigenetiko garrantzitsua da (Lou et al., 2014). Orain arte, oso gutxitan aztertu da metilazioak EZn duen eragina (Fernandez-Jimenez et al., 2014; Scientific Reports, under review). Tesi honetan, metilazioaren garrantzia konfirmatu da NGS teknologiaren bidez. Izan ere, gliadina zeliakoak ez diren kontrolen hesteko zeluletan, aldaketa akutuak sortzeko gai dela ikusi da. Aldaketa horiek, gainera,

ez dira itzulgarriak EZdun pazienteetan, bi urte baino gehiagoz GGD jarraitu arren.

Aipatzekoa da, lan osoan zehar prozedura bioinformatiko berritzaileak erabili ditugula aurretik existitzen ziren geruza genomiko ezberdinetako datuak beraztertze eta EZren patogenesisia azal dezaketen proposamen originalak egiteko. Datuen birziklatze-hurbilketa horiek arrakastatsuak izan daitezkeela frogatu dugu, eta EZn zein beste ezaugarri konplexu batzuetan erabilgarriak izan daitezkeela.

Zentzu honetan, azken hamarkadan datu genomiko asko lortu dira. Metodo biokimiko tradizionalak denbora asko eskatzen dute eta inefizienteak izan daitezke; omika teknologiek, aldiz, etekin-altuko analisiak egin eta datuak ekoitzi ditzakete eskala handitan. Gaixotasunekin asoziatutako lokusak identifikatu dituzten GWA azterketei jarraituz, epigenomika (Gupta eta Hawkins, 2015), transkriptomika (McHale et al., 2013), proteomika (Breker eta Schuldiner, 2014) edo metabolomika (Fearnley eta Inouye, 2016) moduko beste diziplina batzuk ere abiatu dira. Prozesu biologikoen ikuspegi ulerkor bat erdiesteko, maila ezberdinetan lortutako datuak integratu eta aztertu behar dira, tesi honetan egin den bezala. Hurbilketa multi-omikoek geruza biologiko ezberdinetatik lorturako datuak integratzen dituzte errendimendu altuko hurbilketa analitikoak erabiliz, horrela, beraien arteko erlazioak eta sistema konplexuagoek nola funtzionatzen duten ulertzeko (Kohl et al., 2014).

Beraz, omika maila ezberdinen erabilerak eta hauen analisi integratzaileak gaixotasunen prozesu patologikoen ulermenean laguntzen dute, eta erakutsi ditzakete diagnosi zehatzagoetara eta tratamendu egokiagoetara eramango gaituzteen parte-hartzaile, bidezidor eta mekanismo garrantzitsuak. Gainera, hurbilketa horrek analisi funtzionaletarako hautagaiak lehenesten ere lagundu dezake, eta EZ moduko gaixotasun konplexuen inguruan dugun ezagutza handi

dezake, modu horretan gaixotasun konplexu ezberdinen kontrako terapia eta farmakoen garapena sustatuz.

Ondorioz, aurkikuntza berri horiek komunitate zientifikoak ikuspuntu berri bat har dezan lagunduko dute. Horren arabera, mekanismo erregulazaile konplexu eta integratzaileek, maila desberdinetatik, gene erregulazioa aldatu, eta gaixotasun konplexuekin asoziatutako aldakortasun genetikoaren azalpen molekularra osa dezakete.

Ondorioak

1. Gliadinak genoma osoaren koadierazpenaren berrantolaketa eragiten du, bai esposizio akutuan zein epe-luzekoan. Hala ere, egoera osasuntsuarekin erlazionatutako koadierazpen moduluak aldakorragoak dira gaixotasunarekin erlazionatutakoak baino; gliadinaren estimulupean identifikatutako moduluak aldiz, egonkorragoak dira, eta oinarritzko koordinazio unitateak azaleratzen dituzte.

Aurkikuntza hauek koadierazpena EZren garapenean inplikaturik dagoen informazio genomikoaren beste maila bat dela azpimarratzen dute.

2. DKGetan egindako aberaste analisisiek 10 TF eta 46 miRNA identifikatu zituzten, EZn ikusitako koadierazpen aldaketan parte baten erregulatuzaile posible modura.
 - a. Balioztatze biologikorako hautatutako 5 TFtatik, IRF1, CREB1, NFKB1 eta ELK1 gainadierazita zeuden EZ aktibodun pazienteen biopsietan, eta haien itu gene gehienek ere adierazpen aldaketak zituzten. Bestalde, 6 pri-miRNA (hsa-mir-33a, hsa-mir-520b, hsa-mir-520e, hsa-let-7b, hsa-mir-655, hsa-mir-26b) azpiadierazita zeuden EZn, eta horien itu geneak EZ aktibodun pazienteen biopsien epitelio-frakzioan immunean baino gainadieraziago zeuden, eta beraz, mekanismoa zelula-espezifikoa dela ondoriozta daiteke.

Orokorrean, emaitza hauek in silico hurbilketaren baliogarritasuna indartzen dute, bai eta aukeratutako erregulatuzaile hautagaiek EZn parte-hartzen dutela ere.

- b. Hesteko epitelio zeluletan (C2BBel), CREB1 eta IRF1 TFak nukleora translokatzeko dira zelulak gliadinarekin kitzikatzean. Gainera, IRF1 DKGen multzo batean identifikatutako bi itu generen promotoreei lotzen

zaie, *HDAC4* eta *WDR43*en promotoreei, hain zuzen, gliadinak loturan aldaketarik sortzen ez duen arren.

Emaitzok, gliadinak TFetan dituen efektuen bidez, gene adierazpenean duen eragina azpimarratzen dute, baina gaur egun eskuragarri dauden eredu zelularren mugak ere agerian uzten dituzte.

3. Kontserbatutako TADen, eta EZrekin erlazionatutako DKGen eta SNPen arteko lerrokatzea EZn inplikatuak egon eta TADen fusio edo hausturak gerta litezkeen 25 eskualde identifikatzeko gai izan da. Horietako bi eskualdetan (*PROCR-ROMO1* eskualdea fusio modura, eta *HSCB-XBP1* eskualdea haustura modura) TADen mugetako CRISPR-Cas9 bidezko delezioek koadierazpen aldaketak sortzen dituzte zelula ereduetan, aztertutako bi zelula lerroek aurkako patroiak aurkeztu zituzten arren.

Emaitzok 3D genoma tokiko koadierazpenaren erregulatzailerik posibleetako bat dela, eta gaixotasunean asaldatuta dagoela iradokitzen dute. Gainera, eskualde hautagaietan dauden gaixotasunarekin asoziatutako SNPen genotipoek zelula lerro batetik bestera desberdinak diren aldaketak azal ditzakete.

4. Gliadinak *HDAC4* lokusaren hipometilazioa induzitzen du zeliakoak ez diren indibiduen biopsietan. Aldaketa hori ez da gliadinarekin kitzikatutako GGD-EZ pazienteetan ikusten; izan ere, lokus horretan hipometilazio kronikoa ikus daiteke kontrolekin erkatzean. Bestetik, aurretik EZ aktibodun pazienteen hesteko zelula populazioetan identifikatutako *HLA-B* eta *TAP1* promotoreetako DMRak erreplikatu egin ziren GGD-EZ pazienteak kontrolekin konparatzean.

Orokorrean, datu hauek konsitutiboa (eta beraz heredatua edo garapenean oso goiz ezarria) edo GGD bidez itzulezina den metilazio-signatura bat erakusten

dute EZen. Edozein kasutan, DNAREN metilazioa kontutan hartzekoa da EZren genomika aztertzeko orduan.

Bibliografia

Abadie, V., Sollid, L.M., Barreiro, L.B., and Jabri, B. (2011). Integration of Genetic and Immunological Insights into a Model of Celiac Disease Pathogenesis. *Annu. Rev. Immunol.* 29, 493–525.

Agardh, E., Lundstig, A., Perfilyev, A., Volkov, P., Freiburghaus, T., Lindholm, E., Rönn, T., Agardh, C.D., and Ling, C. (2015). Genome-wide analysis of DNA methylation in subjects with type 1 diabetes identifies epigenetic modifications associated with proliferative diabetic retinopathy. *BMC Med.* 13, 182.

Al-hassany, M. (1975). Coeliac disease in Iraqi children. *J. Trop. Pediatr.* 21, 178–179.

Al-Sadi, R., Guo, S., Ye, D., and Ma, T.Y. (2013). TNF- α modulation of intestinal epithelial tight junction barrier is regulated by ERK1/2 activation of Elk-1. *Am. J. Pathol.* 183, 1871–1884.

Al-Tawaty, A.I., and Elbargathy, S.M. (1998). Coeliac disease in north-eastern Libya. *Ann. Trop. Paediatr.* 18, 27–30.

Allen, E.K., Randolph, A.G., Bhangale, T., Dogra, P., Ohlson, M., Oshansky, C.M., Zamora, A.E., Shannon, J.P., Finkelstein, D., Dressen, A., et al. (2017). SNP-mediated disruption of CTCF binding at the IFITM3 promoter is associated with risk of severe influenza in humans. *Nat. Med.* 23, 975–983.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

Anderson, R.P., Degano, P., Godkin, A.J., Jewell, D.P., and Hill, A.V.S. (2000). In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat. Med.* 6, 337–342.

Arentz-Hansen, H., Körner, R., Molberg, O., Quarsten, H., Vader, W., Kooy, Y.M., Lundin, K.E., Koning, F., Roepstorff, P., Sollid, L.M., et al. (2000). 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.* 191, 603–612.

Arnone, M.I., and Davidson, E.H. (1997). The hardwiring of development: organization and function of genomic regulatory systems. *Development* 124, 1851–1864.

Bae, Y.S., Oh, H., Rhee, S.G., and Yoo, Y.D. (2011). Regulation of reactive oxygen species generation in cell signaling. *Mol Cells* 32, 491–509.

Banka, S., de Goede, C., Yue, W.W., Morris, A.A.M., von Bremen, B., Chandler, K.E., Feichtinger, R.G., Hart, C., Khan, N., Lunzer, V., et al. (2014). Expanding the clinical and molecular spectrum of thiamine pyrophosphokinase deficiency:

A treatable neurological disorder caused by TPK1 mutations. *Mol. Genet. Metab.* *113*, 301–306.

Barutcu, A.R., Maass, P.G., Lewandowski, J.P., Weiner, C.L., and Rinn, J.L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. *Nat. Commun.* *9*, 1444.

Batlle-López, A., Cortiguera, M.G., Rosa-Garrido, M., Blanco, R., Del Cerro, E., Torrano, V., Wagner, S.D., and Delgado, M.D. (2015). Novel CTCF binding at a site in exon1A of BCL6 is associated with active histone marks and a transcriptionally active locus. *Oncogene* *34*, 246–256.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. *Science* *304*, 1321–1325.

van Belzen, M.J., Koeleman, B.P.C., Crusius, J.B.A., Meijer, J.W.R., Bardoel, A.F.J., Pearson, P.L., Sandkuijl, L.A., Houwen, R.H.J., and Wijmenga, C. (2004). Defining the contribution of the HLA region to cis DQ2-positive coeliac disease patients. *Genes Immun.* *5*, 215–220.

Van Belzen, M.J., Meijer, J.W.R., Sandkuijl, L.A., Bardoel, A.F.J., Mulder, C.J.J., Pearson, P.L., Houwen, R.H.J., and Wijmenga, C. (2003). A major non-HLA locus in celiac disease maps to chromosome 19. *Gastroenterology* *125*, 1032–1041.

Bessa, C., Soares, J., Raimundo, L., Loureiro, J.B., Gomes, C., Reis, F., Soares, M.L., Santos, D., Dureja, C., Chaudhuri, S.R., et al. (2018). Discovery of a small-molecule protein kinase C δ -selective activator with promising application in colon cancer therapy. *Cell Death Dis.* *9*, 23.

Bhagwat, A.S., and Vakoc, C.R. (2015). Targeting Transcription Factors in Cancer. *Trends in Cancer* *1*, 53–65.

Bickmore, W.A. (2013). The Spatial Organization of the Human Genome. *Annu. Rev. Genomics Hum. Genet.* *14*, 67–84.

Bind, M.A., Zanobetti, A., Gasparini, A., Peters, A., Coull, B., Baccarelli, A., Tarantini, L., Koutrakis, P., Vokonas, P., and Schwartz, J. (2014). Effects of temperature and relative humidity on DNA methylation. *Epidemiology* *25*, 561–569.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* *16*, 6–21.

Bodé, S., and Gudmand-Høyer, E. (1996). Incidence and prevalence of adult coeliac disease within a defined geographic area in Denmark. *Scand. J. Gastroenterol.* *31*, 694–699.

- Bondar, C., Plaza-Izurieta, L., Fernandez-Jimenez, N., Irastorza, I., Withoff, S., Wijmenga, C., Chirido, F., and Bilbao, J.R. (2014). THEMIS and PTPRK in celiac intestinal mucosa: coexpression in disease and after in vitro gliadin challenge. *Eur. J. Hum. Genet.* *22*, 358–362.
- Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* *13*, 1097–1101.
- Bouhleb, M.A., Lambert, M., and David-Cordonnier, M.-H. (2015). Targeting Transcription Factor Binding to DNA by Competing with DNA Binders as an Approach for Controlling Gene Expression. *Curr. Top. Med. Chem.* *15*, 1323–1358.
- Bracken, S., Byrne, G., Kelly, J., Jackson, J., and Feighery, C. (2008). Altered gene expression in highly purified enterocytes from patients with active coeliac disease. *BMC Genomics* *9*, 377.
- Bratanic, N., Smigoc Schweiger, D., Mendez, A., Bratina, N., Battelino, T., and Vidan-Jeras, B. (2010). An influence of HLA-A, B, DR, DQ, and MICA on the occurrence of Celiac disease in patients with type 1 diabetes. *Tissue Antigens* *76*, 208–215.
- Breker, M., and Schuldiner, M. (2014). The emergence of proteome-wide technologies: Systematic analysis of proteins comes of age. *Nat. Rev. Mol. Cell Biol.* *15*, 453–464.
- Buoli Comani, G., Panceri, R., Dinelli, M., Biondi, A., Mancuso, C., Meneveri, R., and Barisani, D. (2015). miRNA-regulated gene expression differs in celiac disease patients according to the age of presentation. *Genes Nutr.* *10*, 482.
- Capuano, M., Iaffaldano, L., Tinto, N., Montanaro, D., Capobianco, V., Izzo, V., Tucci, F., Troncone, G., Greco, L., and Sacchetti, L. (2011). MicroRNA-449a overexpression, reduced NOTCH1 signals and scarce goblet cells characterize the small intestine of celiac patients. *PLoS One* *6*, e29094.
- Caputo, I., Barone, M.V., Martucciello, S., Lepretti, M., and Esposito, C. (2009). Tissue transglutaminase in celiac disease: Role of autoantibodies. *Amino Acids* *36*, 693–699.
- Castellanos-Rubio, A., and Bilbao, J.R. (2017). Profiling Celiac Disease-Related Transcriptional Changes. *Int. Rev. Cell Mol. Biol.* *336*, 149–174.
- Castellanos-Rubio, A., Martin-Pagola, A., Santin, I., Hualde, I., Aransay, A.M., Castaño, L., Vitoria, J.C., and Bilbao, J.R. (2008). Combined Functional and Positional Gene Information for the Identification of Susceptibility Variants in Celiac Disease. *Gastroenterology* *134*, 738–746.
- Castellanos-Rubio, A., Santin, I., Irastorza, I., Castaño, L., Carlos Vitoria, J., and Bilbao, J.R. (2009). TH17 (and TH1) signatures of intestinal biopsies of CD

patients in response to gliadin. *Autoimmunity* 42, 69–73.

Castellanos-Rubio, A., Santin, I., Martin-Pagola, A., Irastorza, I., Castaño, L., Vitoria, J.C., and Bilbao, J.R. (2010). Long-term and acute effects of gliadin on small intestine of patients on potentially pathogenic networks in celiac disease. *Autoimmunity* 43, 131–139.

Castellanos-Rubio, A., Fernandez-Jimenez, N., Kratchmarov, R., Luo, X., Bhagat, G., Green, P.H.R., Schneider, R., Kiledjian, M., Bilbao, J.R., and Ghosh, S. (2016). A long noncoding RNA associated with susceptibility to celiac disease. *Science* (80-.). 352, 91–95.

Cataldo, F., and Montalto, G. (2007). Celiac disease in the developing countries: a new and challenging public health problem. *World J. Gastroenterol.* 13, 2153–2159.

Catassi, C., Gatti, S., and Fasano, A. (2014). The new epidemiology of celiac disease. *J. Pediatr. Gastroenterol. Nutr.* 59, 7–9.

Chen, J.Q., Papp, G., Szodoray, P., and Zeher, M. (2016). The role of microRNAs in the pathogenesis of autoimmune diseases. *Autoimmun. Rev.* 15, 1171–1180.

Chung, J.S., Lee, S., and Yoo, Y. Do (2014). Constitutive NF- κ B activation and tumor-growth promotion by Romo1-mediated reactive oxygen species production. *Biochem. Biophys. Res. Commun.* 450, 1656–1661.

Cicerone, C., Nenna, R., and Pontone, S. (2015). Th17, intestinal microbiota and the abnormal immune response in the pathogenesis of celiac disease. *Gastroenterol. Hepatol. from Bed to Bench* 8, 117–122.

Collin, P., Reunala, T., Rasmussen, M., Kyrönpalo, S., Pehkonen, E., Laippala, P., and Mäki, M. (1997). High incidence and prevalence of adult coeliac disease: Augmented diagnostic approach. *Scand. J. Gastroenterol.* 32, 1129–1133.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex disease traits with global gene expression. *Nat Rev Genet* 10, 184–194.

Dalmay, T. (2008). Identification of genes targeted by microRNAs. *Biochem. Soc. Trans.* 36, 1194–1196.

Dekker, J., and Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. *FEBS Lett.* 589, 2877–2884.

Delpu, Y., Cordelier, P., Cho, W.C., and Torrisani, J. (2013). DNA methylation and cancer diagnosis. *Int. J. Mol. Sci.* 14, 15029–15058.

Dicke, W.K. (1950). *Coeliake*. MD Thesis.

- Dieli-Crimi, R., Cénit, M.C., and Núñez, C. (2015). The genetics of celiac disease: A comprehensive review of clinical implications. *J. Autoimmun.* 64, 26–41.
- Le Dily, F.L., Baù, D., Pohl, A., Vicent, G.P., Serra, F., Soronellas, D., Castellano, G., Wright, R.H.G., Ballare, C., Filion, G., et al. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* 28, 2151–2162.
- Diop-Bove, N.K., Wu, J., Zhao, R., Locker, J., and Goldman, I.D. (2009). Hypermethylation of the human proton-coupled folate transporter (SLC46A1) minimal transcriptional regulatory region in an antifolate-resistant HeLa cell line. *Mol. Cancer Ther.* 8, 2424–2431.
- Diosdado, B., Wapenaar, M.C., Franke, L., Duran, K.J., Goerres, M.J., Hadithi, M., Crusius, J.B.A., Meijer, J.W.R., Duggan, D.J., Mulder, C.J.J., et al. (2004). A microarray screen for novel candidate genes in coeliac disease pathogenesis. *Gut* 53, 944–951.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. *Mol. Cell* 62, 668–680.
- Dubé, C., Rostom, A., Sy, R., Cranney, A., Saloojee, N., Garritty, C., Sampson, M., Zhang, L., Yazdi, F., Mamaladze, V., et al. (2005). The prevalence of celiac disease in average-risk and at-risk Western European populations: A systematic review. *Gastroenterology* 128, S57-67.
- Dubois, P.C., Trynka, G., Franke, L., Hunt, K.A., Romanos, J., Curtotti, A., Zhernakova, A., Heap, G.A.R., Ádány, R., Aromaa, A., et al. (2010). Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* 42, 295–302.
- Erson, A.E., and Petty, E.M. (2008). MicroRNAs in development and disease. *Clin. Genet.* 74, 296–306.
- Escudero-Hernández, C., Plaza-Izurietta, L., Garrote, J.A., Bilbao, J.R., and Arranz, E. (2017). Association of the IL-15 and IL-15R α genes with celiac disease. *Cytokine* 99, 73–79.
- Esteller, M. (2011). Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12, 861–874.
- Fang, Z., Lin, A., Chen, J., Zhang, X., Liu, H., Li, H., Hu, Y., Zhang, X., Zhang, J., Qiu, L., et al. (2016). CREB1 directly activates the transcription of ribonucleotide reductase small subunit M2 and promotes the aggressiveness of

human colorectal cancer. *Oncotarget* 7, 78055–78068.

Fasano, A., and Catassi, C. (2001). Current approaches to diagnosis and treatment of celiac disease: An evolving spectrum. *Gastroenterology* 120, 636–651.

Fearnley, L.G., and Inouye, M. (2016). Metabolomics in epidemiology: from metabolite concentrations to integrative reaction networks. *Int. J. Epidemiol.* 45, 1319–1328.

Feighery, C. (1999). Coeliac disease. *Bmj* 319, 236–239.

Feighery, C., Weir, D.G., Whelan, A., Willoughby, R., Youngprapakorn, S., Lynch, S., O’Moráin, C., McEneaney, P., and O’Farrelly, C. (1998). Diagnosis of gluten-sensitive enteropathy: is exclusive reliance on histology appropriate? *Eur. J. Gastroenterol. Hepatol.* 10, 919–925.

Felli, C., Baldassarre, A., and Masotti, A. (2017). Intestinal and circulating micrnas in coeliac disease. *Int. J. Mol. Sci.* 18.

Fernandez-Jimenez, N., Castellanos-Rubio, A., Plaza-Izurieta, L., Irastorza, I., Elcoroaristizabal, X., Jauregi-Miguel, A., Lopez-Euba, T., Tutau, C., De Pancorbo, M.M., Vitoria, J.C., et al. (2014). Coregulation and modulation of NFκB-related genes in celiac disease: Uncovered aspects of gut mucosal inflammation. *Hum. Mol. Genet.* 23, 1298–1310.

Fina, D., Sarra, M., Caruso, R., Del Vecchio Blanco, G., Pallone, F., MacDonald, T.T., and Monteleone, G. (2008). Interleukin 21 contributes to the mucosal T helper cell type 1 response in coeliac disease. *Gut* 57, 887–892.

Flavahan, W.A., Drier, Y., Liao, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* 529, 110–114.

Folk, J.E., and Chung, S.I. (1985). Transglutaminases. *Methods Enzymol.* 113, 358–375.

Folk, J.E., and Cole, P.W. (1966). Transglutaminase: Mechanistic features of the active site as determined by kinetic and inhibitor studies. *BBA - Enzymol. Biol. Oxid.* 122, 244–264.

Freeman, H.J. (2013). Non-dietary forms of treatment for adult celiac disease. *World J. Gastrointest. Pharmacol. Ther.* 4, 108–112.

Friendly, M. (2002). Corrgrams: Exploratory displays for correlatigon matrices. *Am. Stat.* 56, 316–324.

Frietze, S., and Farnham, P.J. (2011). Transcription factor effector domains.

Subcell. Biochem. 52, 261–277.

Fulton, D.L.L., Sundararajan, S., Badis, G., Hughes, T.R.R., Wasserman, W.W.W., Roach, J.C.C., and Sladek, R. (2009). TFCat: The curated catalog of mouse and human transcription factors. *Genome Biol.* 10, R29.

Galvao, L.C., Gomes, R.C., and Ramos, A.M. (1992). [Celiac disease: report of 20 cases in Rio Grande do Norte, Brazil]. *Arq Gastroenterol* 29, 28–33.

Garcia-Etxebarria, K., Jauregi-Miguel, A., Romero-Garmendia, I., Plaza-Izurieta, L., Legarda, M., Irastorza, I.I., and Bilbao, J.R.J.R. (2016). Ancestry-based stratified analysis of Immunochip data identifies novel associations with celiac disease. *Eur. J. Hum. Genet.* 24, 1831–1834.

Gargalovic, P.S., Gharavi, N.M., Clark, M.J., Pagnon, J., Yang, W.P., He, A., Truong, A., Baruch-Oren, T., Berliner, J.A., Kirchgessner, T.G., et al. (2006). The unfolded protein response is an important regulator of inflammatory genes in endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* 26, 2490–2496.

Garten, A., Petzold, S., Schuster, S., Körner, A., Kratzsch, J., and Kiess, W. (2011). Nampt and its potential role in inflammation and type 2 diabetes. *Handb. Exp. Pharmacol.* 203, 147–164.

Garvie, C.W., Stagno, J.R., Reid, S., Singh, A., Harrington, E., and Boss, J.M. (2007). Characterization of the RFX complex and the RFX5(L66A) mutant: Implications for the regulation of MHC class II gene expression. *Biochemistry* 46, 1597–1611.

Gee, S. (1888). On the celiac disease. *St Bart Hosp Rep.* 24, 17–20.

Geertz, M., Shore, D., and Maerkl, S.J. (2012). Massively parallel measurements of molecular interaction kinetics on a microfluidic platform. *Proc. Natl. Acad. Sci.* 109, 16540–16545.

Ghirlando, R., and Felsenfeld, G. (2016). CTCF: Making the right connections. *Genes Dev.* 30, 881–891.

Giorgio, E., Robyr, D., Spielmann, M., Ferrero, E., Di Gregorio, E., Imperiale, D., Vaula, G., Stamoulis, G., Santoni, F., Atzori, C., et al. (2014). A large genomic deletion leads to enhancer adoption by the lamin B1 gene: A second path to autosomal dominant adult-onset demyelinating leukodystrophy (ADLD). *Hum. Mol. Genet.* 24, 3143–3154.

Goel, G., King, T., Daveson, A.J., Andrews, J.M., Krishnarajah, J., Krause, R., Brown, G.J.E., Fogel, R., Barish, C.F., Epstein, R., et al. (2017). Epitope-specific immunotherapy targeting CD4-positive T cells in coeliac disease: two randomised, double-blind, placebo-controlled phase 1 studies. *Lancet Gastroenterol. Hepatol.* 2, 479–493.

Gong, J., Tian, J., Lou, J., Wang, X., Ke, J., Li, J., Yang, Y., Gong, Y., Zhu, Y., Zou, D., et al. (2018). A polymorphic MYC response element in KBTBD11 influences colorectal cancer risk, especially in interaction with an MYC-regulated SNP rs6983267. *Ann. Oncol.* *29*, 632–639.

Greco, L., Corazza, G., Babron, M.C., Clot, F., Fulchignoni-Lataud, M.C., Percopo, S., Zavattari, P., Bouguerra, F., Dib, C., Tosi, R., et al. (1998). Genome search in celiac disease. *Am. J. Hum. Genet.* *62*, 669–675.

Greco, L., Romino, R., Coto, I., Di Cosmo, N., Percopo, S., Maglio, M., Paparo, F., Gasperi, V., Limongelli, M.G., Cotichini, R., et al. (2002). The first large population based twin study of coeliac disease. *Gut* *50*, 624–628.

Green, P.H.R., and Cellier, C. (2007). Celiac disease. *N. Engl. J. Med.* *357*, 1731–1743.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* *34*, D140–D144.

Gross, D., and Garrard, W.T. (1988). Nuclease Hypersensitive Sites In Chromatin. *Annu. Rev. Biochem.* *57*, 159–197.

Guo, M., Mao, X., Ji, Q., Lang, M., Li, S., Peng, Y., Zhou, W., Xiong, B., and Zeng, Q. (2010). Inhibition of IFN regulatory factor-1 down-regulate Th1 cell function in patients with acute coronary syndrome. *J. Clin. Immunol.* *30*, 241–252.

Gupta, B., and Hawkins, R.D. (2015). Epigenomics of autoimmune diseases. *Immunol. Cell Biol.* *93*, 271–276.

Gutierrez-Achury, J., Zhernakova, A., Pulit, S.L., Trynka, G., Hunt, K.A., Romanos, J., Raychaudhuri, S., Van Heel, D.A., Wijmenga, C., and De Bakker, P.I.W. (2015). Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat. Genet.* *47*, 577–578.

Hang, L.-W., Hsia, T.-C., Chen, W.-C., Chen, H.-Y., and Tsai, F.-J. (2003). TAP1 gene AccI polymorphism is associated with atopic bronchial asthma. *J. Clin. Lab. Anal.* *17*, 57–60.

Harris, K.M., Fasano, A., and Mann, D.L. (2010). Monocytes differentiated with IL-15 support Th17 and Th1 responses to wheat gliadin: Implications for celiac disease. *Clin. Immunol.* *135*, 430–439.

He, D., Liu, Z.-P., Honda, M., Kaneko, S., and Chen, L. (2012). Coexpression network analysis in chronic hepatitis B and C hepatic lesions reveals distinct patterns of disease progression to hepatocellular carcinoma. *J. Mol. Cell Biol.* *4*, 140–152.

He, Z., Proudfoot, C., Whitelaw, C.B.A., and Lillico, S.G. (2016). Comparison of CRISPR/Cas9 and TALENs on editing an integrated EGFP gene in the genome of HEK293FT cells. *Springerplus* 5, 814.

van Heel, D.A., Franke, L., Hunt, K.A., Gwilliam, R., Zhernakova, A., Inouye, M., Wapenaar, M.C., Barnardo, M.C., Bethel, G., Holmes, G.K., et al. (2007). A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nat. Genet.* 39, 827–829.

Hogen Esch, C.E., Rosen, A., Auricchio, R., Romanos, J., Chmielewska, A., Putter, H., Ivarsson, A., Szajewska, H., Koning, F., Wijmenga, C., et al. (2010). The PreventCD Study design: towards new strategies for the prevention of coeliac disease. *Eur. J. Gastroenterol. Hepatol.* 22, 1424–1430.

Holmes, E.E., Jung, M., Meller, S., Leisse, A., Sailer, V., Zech, J., Mengdehl, M., Garbe, L.A., Uhl, B., Kristiansen, G., et al. (2014). Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, aspirates, lavages, effusions, plasma, serum, and urine. *PLoS One* 9, e93933.

Holopainen, P., Naluai, A.T., Moodie, S., Percopo, S., Coto, I., Clot, F., Ascher, H., Sollid, L., Ciclitira, P., Greco, L., et al. (2004). Candidate gene region 2q33 in European families with coeliac disease. *Tissue Antigens* 63, 212–222.

Hong, S., and Kim, D. (2017). Computational characterization of chromatin domain boundary-associated genomic elements. *Nucleic Acids Res.* 45, 10403–10414.

Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., Khodiyar, V.K., Lush, M.J., Povey, S., Talbot, C.C., Wright, M.W., et al. (2004). Gene map of the extended human MHC. *Nat. Rev. Genet.* 5, 889–899.

Hüe, S., Mention, J.J., Monteiro, R.C., Zhang, S.L., Cellier, C., Schmitz, J., Verkarre, V., Fodil, N., Bahram, S., Cerf-Bensussan, N., et al. (2004). A direct role for NKG2D/MICA interaction in villous atrophy during celiac disease. *Immunity* 21, 367–377.

Hume, M.A., Barrera, L.A., Gisselbrecht, S.S., and Bulyk, M.L. (2015). UniPROBE, update 2015: New tools and content for the online database of protein-binding microarray data on protein-DNA interactions. *Nucleic Acids Res.* 43, D117–D122.

Hunt, K.A., Mistry, V., Bockett, N.A., Ahmad, T., Ban, M., Barker, J.N., Barrett, J.C., Blackburn, H., Brand, O., Burren, O., et al. (2013). Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. *Nature* 498, 232–235.

Husby, S., Koletzko, S., Korponay-Szabó, I., Mearin, M., Phillips, A., Shamir, R., Troncone, R., Giersiepen, K., Branksi, D., Catassi, C., et al. (2012).

European Society for Pediatric Gastroenterology, Hepatology, and Nutrition Guidelines for the Diagnosis of Coeliac Disease. *J. Pediatr. Gastroenterol. Nutr.* *54*, 136–160.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* *41*, 178–186.

Ivarsson, A., Persson, L., Nyström, L., Ascher, H., Cavell, B., Danielsson, L., Dannaeus, A., Lindberg, T., Lindquist, B., Stenhammar, L., et al. (2000). Epidemic of coeliac disease in Swedish children. *Acta Paediatr. Int. J. Paediatr.* *89*, 165–171.

Jabri, B., and Sollid, L.M. (2009). Tissue-mediated control of immunopathology in coeliac disease. *Nat. Rev. Immunol.* *9*, 858–870.

Jauregi-Miguel, A., Fernandez-Jimenez, N., Irastorza, I., Plaza-Izurietta, L., Vitoria, J.C., and Bilbao, J.R. (2014). Alteration of tight junction gene expression in celiac disease. *J. Pediatr. Gastroenterol. Nutr.* *58*, 762–767.

Jheng, J.R., Lin, C.Y., Horng, J.T., and Lau, K.S. (2012). Inhibition of enterovirus 71 entry by transcription factor XBP1. *Biochem. Biophys. Res. Commun.* *420*, 882–887.

Ji, X., Dadon, D.B., Powell, B.E., Misteli, T., Jaenisch, R., Young, R.A., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., et al. (2016). 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell Stem Cell* *18*, 262–275.

Jin, C., Jin, Z., Chen, N., Lu, M., Liu, C., Hu, W.-L., and Zheng, C. (2016). Activation of IRE1 α -XBP1 pathway induces cell proliferation and invasion in colorectal carcinoma. *Biochem. Biophys. Res. Commun.* *470*, 75–81.

Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. *Cell* *152*, 327–339.

Jolma, A., Yin, Y., Nitta, K.R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* *527*, 384–388.

Jones, P.A., Baylin, S.B., Erdjument-Bromage, H., Tempst, P., Bird, A., Reinberg, D., Sato, F., Meltzer, S.J., Sidransky, D., Badenhorst, P., et al. (2007). The Epigenomics of Cancer. *Cell* *128*, 683–692.

Juuti-Uusitalo, K., Mäki, M., Kaukinen, K., Collin, P., Visakorpi, T., Vihinen, M., and Kainulainen, H. (2004). cDNA microarray analysis of gene expression in

coeliac disease jejunal biopsy samples. *J. Autoimmun.* 22, 249–265.

Juuti-Uusitalo, K., Mäki, M., Kainulainen, H., Isola, J., and Kaukinen, K. (2007). Gluten affects epithelial differentiation-associated genes in small intestinal mucosa of coeliac patients. *Clin. Exp. Immunol.* 150, 294–305.

Karell, K., Louka, A.S., Moodie, S.J., Ascher, H., Clot, F., Greco, L., Ciclitira, P.J., Sollid, L.M., and Partanen, J. (2003). 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.* 64, 469–477.

Kaser, A., Lee, A.-H., Franke, A., Glickman, J.N., Zeissig, S., Tilg, H., Nieuwenhuis, E.E.S., Higgins, D.E., Schreiber, S., Glimcher, L.H., et al. (2008). XBP1 Links ER Stress to Intestinal Inflammation and Confers Genetic Risk for Human Inflammatory Bowel Disease. *Cell* 134, 743–756.

Katainen, R., Dave, K., Pitkänen, E., Palin, K., Kivioja, T., Välimäki, N., Gylfe, A.E., Ristolainen, H., Hänninen, U.A., Cajuso, T., et al. (2015). CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat. Genet.* 47, 818–821.

de Kauwe, A.L., Chen, Z., Anderson, R.P., Keech, C.L., Price, J.D., Wijburg, O., Jackson, D.C., Ladhams, J., Allison, J., and McCluskey, J. (2009). Resistance to Celiac Disease in Humanized HLA-DR3-DQ2-Transgenic Mice Expressing Specific Anti-Gliadin CD4+ T Cells. *J. Immunol.* 182, 7440–7450.

Khuffash, F.A., Barakat, M.H., Shaltout, A.A., Farwana, S.S., Adnani, M.S., and Tungekar, M.F. (1987). Coeliac disease among children in Kuwait: Difficulties in diagnosis and management. *Gut* 28, 1595–1599.

Kim, H.J., Jo, M.J., Kim, B.R., Kim, J.L., Jeong, Y.A., Na, Y.J., Park, S.H., Lee, S.Y., Lee, D.H., Lee, H.S., et al. (2017). Reactive oxygen species modulator-1 (Romo1) predicts unfavorable prognosis in colorectal cancer patients. *PLoS One* 12, e0176834.

Kim, M.S., Chang, X., Yamashita, K., Nagpal, J.K., Baek, J.H., Wu, G., Trink, B., Ratovitski, E.A., Mori, M., and Sidransky, D. (2008). Aberrant promoter methylation and tumor suppressive activity of the DFNA5 gene in colorectal carcinoma. *Oncogene* 27, 3624–3634.

Kishi, Y., Kondo, T., Xiao, S., Yosef, N., Gaublot, J., Wu, C., Wang, C., Chihara, N., Regev, A., Joller, N., et al. (2016). Protein C receptor (PROCR) is a negative regulator of Th17 pathogenicity. *J. Exp. Med.* 213, 2489–2501.

Kitamura, K., Seike, M., Okano, T., Matsuda, K., Miyanaga, A., Mizutani, H., Noro, R., Minegishi, Y., Kubota, K., and Gemma, A. (2014). MiR-134/487b/655 cluster regulates TGF- β -induced epithelial-mesenchymal transition and drug resistance to gefitinib by targeting MAGI2 in lung adenocarcinoma cells. *Mol. Cancer Ther.* 13, 444–453.

Kohl, M., Megger, D.A., Trippler, M., Meckel, H., Ahrens, M., Bracht, T., Weber, F., Hoffmann, A.C., Baba, H.A., Sitek, B., et al. (2014). A practical data processing workflow for multi-OMICS projects. *Biochim. Biophys. Acta* 1844, 52–62.

Kolodziej, L.E., Lodolce, J.P., Chang, J.E., Schneider, J.R., Grimm, W.A., Bartulis, S.J., Zhu, X., Messer, J.S., Murphy, S.F., Reddy, N., et al. (2011). TNFAIP3 maintains intestinal barrier function and supports epithelial cell tight junctions. *PLoS One* 6, e26352.

Krueger, F., and Andrews, S.R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572.

Krumm, A., and Duan, Z. (2018). Understanding the 3D genome: Emerging impacts on human disease. *Semin. Cell Dev. Biol.*

Kuo, P.-L., Liao, S.-H., Hung, J.-Y., Huang, M.-S., and Hsu, Y.-L. (2013). MicroRNA-33a functions as a bone metastasis suppressor in lung cancer by targeting parathyroid hormone related protein. *Biochim. Biophys. Acta* 1830, 3756–3766.

de la Fuente, A. (2010). From “differential expression” to “differential networking” - identification of dysfunctional regulatory networks in diseases. *Trends Genet.* 26, 326–333.

Lahdenperä, A., Ludvigsson, J., Fälth-Magnusson, K., Högberg, L., and Vaarala, O. (2011). The effect of gluten-free diet on Th1-Th2-Th3-associated intestinal immune responses in celiac disease. *Scand. J. Gastroenterol.* 46, 538–549.

Lai, L., Song, Y., Liu, Y., Chen, Q., Han, Q., Chen, W., Pan, T., Zhang, Y., Cao, X., and Wang, Q. (2013). MicroRNA-92a negatively regulates toll-like receptor (TLR)-triggered inflammatory response in macrophages by targeting MKK4 kinase. *J. Biol. Chem.* 288, 7956–7967.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Lebreton, C., Ménard, S., Abed, J., Moura, I.C., Coppo, R., Dugave, C., Monteiro, R.C., Fricot, A., Traore, M.G., Griffin, M., et al. (2012). Interactions among secretory immunoglobulin A, CD71, and transglutaminase-2 affect permeability of intestinal epithelial cells to gliadin peptides. *Gastroenterology* 143, 698–707.

Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in disease. *Cell* 152, 1237–1251.

- Lee, S., Park, Y.H., Chung, J.S., and Yoo, Y.D. (2015). Romo1 and the NF- κ B pathway are involved in oxidative stress-induced tumor cell invasion. *Int. J. Oncol.* *46*, 2021–2028.
- Leffler, D.A., Kelly, C.P., Green, P.H.R., Fedorak, R.N., Dimarino, A., Perrow, W., Rasmussen, H., Wang, C., Bercik, P., Bachir, N.M., et al. (2015). Larazotide acetate for persistent symptoms of celiac disease despite a gluten-free diet: A randomized controlled trial. *Gastroenterology* *148*, 1311–1319.
- Lenna, S., Assassi, S., Farina, G.A., Mantero, J.C., Scorza, R., Lafyatis, R., Farber, H.W., and Trojanowska, M. (2015). The HLA-B*35 allele modulates ER stress, inflammation and proliferation in PBMCs from Limited Cutaneous Systemic Sclerosis patients. *Arthritis Res. Ther.* *17*, 363.
- León, A.J., Garrote, J.A., Blanco-Quirós, A., Calvo, C., Fernández-Salazar, L., Del Villar, A., Barrera, A., and Arranz, E. (2006). Interleukin 18 maintains a long-standing inflammation in coeliac disease patients. *Clin. Exp. Immunol.* *146*, 479–485.
- Lerner, A. (2010). New therapeutic strategies for celiac disease. *Autoimmun. Rev.* *9*, 144–147.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* *25*, 2078–2079.
- Li, J., Zhou, D., Qiu, W., Shi, Y., Yang, J.J., Chen, S., Wang, Q., and Pan, H. (2018). Application of Weighted Gene Co-expression Network Analysis for Data from Paired Design. *Sci. Rep.* *8*, 622.
- Liang, H., Cheung, L.W.T., Li, J., Ju, Z., Yu, S., Stemke-Hale, K., Dogruluk, T., Lu, Y., Liu, X., Gu, C., et al. (2012). Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer. *Genome Res.* *22*, 2120–2129.
- Lindfors, K., Kaukinen, K., and Mäki, M. (2009). A role for anti-transglutaminase 2 autoantibodies in the pathogenesis of coeliac disease? *Amino Acids* *36*, 685–691.
- Ling, A., Löfgren-Burström, A., Larsson, P., Li, X., Wikberg, M.L., Öberg, Å., Stenling, R., Edin, S., and Palmqvist, R. (2017). TAP1 down-regulation elicits immune escape and poor prognosis in colorectal cancer. *Oncoimmunology* *6*, e1356143.
- Lister, R., Pelizzola, M., Downen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* *462*, 315–322.

Lou, S., Lee, H.M., Qin, H., Li, J.W., Gao, Z., Liu, X., Chan, L.L., Lam, V.K.L., So, W.Y., Wang, Y., et al. (2014). Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome Biol.* *15*, 408.

Louka, S., and Sollid, L.M. (2003). HLA in coeliac disease: unravelling the complex genetics of a complex disorder. *Tissue Antigens* *61*, 105–117.

Ludwig, H., Polymenidis, Z., Granditsch, G., and Wick, G. (1973). Association of HL-A1 and HL-A8 with childhood celiac disease. *Z. Immunitätsforsch. Exp. Klin. Immunol.* *146*, 158–167.

Lundin, K.E., Sollid, L.M., Qvigstad, E., Markussen, G., Gjertsen, H. a, Ek, J., and Thorsby, E. (1990). T lymphocyte recognition of a celiac disease-associated cis- or trans-encoded HLA-DQ alpha/beta-heterodimer. *J. Immunol.* *145*, 136–139.

Lundin, K.E., Scott, H., Hansen, T., Paulsen, G., Halstensen, T.S., Fausa, O., Thorsby, E., and Sollid, L.M. (1993). 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.* *178*, 187–196.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* *161*, 1012–1025.

Lupiáñez, D.G., Spielmann, M., and Mundlos, S. (2016). Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet.* *32*, 225–237.

Madrigano, J., Baccarelli, A., Mittleman, M.A., Wright, R.O., Sparrow, D., Vokonas, P.S., Tarantini, L., and Schwartz, J. (2011). Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. *Environ. Health Perspect.* *119*, 977–982.

Magni, S., Comani, G.B., Elli, L., Vanessi, S., Ballarini, E., Nicolini, G., Rusconi, M., Castoldi, M., Meneveri, R., Muckenthaler, M.U., et al. (2014). miRNAs Affect the Expression of Innate and Adaptive Immunity Proteins in Celiac Disease. *Am. J. Gastroenterol.* *109*, 1662–1674.

Maiuri, L., Troncone, R., Mayer, M., Coletta, S., Picarelli, A., De Vincenzi, M., Pavone, V., and Auricchio, S. (1996). 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.* *31*, 247–253.

Maiuri, L., Ciacci, C., Ricciardelli, I., Vacca, L., Raia, V., Auricchio, S., Picard,

- J., Osman, M., Quarantino, S., and Londei, M. (2003). Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *Lancet* 362, 30–37.
- Maiuri, L., Luciani, A., Vilella, V.R., Vasaturo, A., Giardino, I., Pettoello-Mantovani, M., Guido, S., Cexus, O.N., Peake, N., Londei, M., et al. (2010). Lysosomal accumulation of gliadin p31-43 peptide induces oxidative stress and tissue transglutaminase-mediated PPAR γ downregulation in intestinal epithelial cells and coeliac mucosa. *Gut* 59, 311–319.
- Mäki, M. (1995). The humoral immune system in coeliac disease. *Baillieres. Clin. Gastroenterol.* 9, 231–249.
- Mäki, M., and Collin, P. (1997). Coeliac disease. *Lancet* 349, 1755–1759.
- Marsh, M.N. (1992). Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* 102, 330–354.
- Martin-Pagola, A., Ortiz, L., De Nanclares, G.P., Vitoria, J.C., Castaño, L., and Bilbao, J.R. (2003). Analysis of the Expression of MICA in Small Intestinal Mucosa of Patients with Celiac Disease. *J. Clin. Immunol.* 23, 498–503.
- Martín-Pagola, A., Pérez-Nanclares, G., Ortiz, L., Vitoria, J.C., Hualde, I., Zaballa, R., Preciado, E., Castaño, L., and Bilbao, J.R. (2004). MICA response to gliadin in intestinal mucosa from celiac patients. *Immunogenetics* 56, 549–554.
- Martin, P., McGovern, A., Orozco, G., Duffus, K., Yarwood, A., Schoenfelder, S., Cooper, N.J., Barton, A., Wallace, C., Fraser, P., et al. (2015). Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat. Commun.* 6, 10069.
- Martinon, F., Chen, X., Lee, A.H., and Glimcher, L.H. (2010). TLR activation of the transcription factor XBP1 regulates innate immune responses in macrophages. *Nat. Immunol.* 11, 411–418.
- Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., et al. (2016). JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 44, D110–D115.
- Matys, V., Kel-Margoulis, O. V, Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., et al. (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* 34, D108-10.
- Mazzarella, G., Maglio, M., Paparo, F., Nardone, G., Stefanile, R., Greco, L., Van de Wal, Y., Kooy, Y., Koning, F., Auricchio, S., et al. (2003). 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* 52, 57–62.

McDermott, E., Ryan, E.J., Tosetto, M., Gibson, D., Burrage, J., Keegan, D., Byrne, K., Crowe, E., Sexton, G., Malone, K., et al. (2016). DNA methylation profiling in inflammatory bowel disease provides new insights into disease pathogenesis. *J. Crohn's Colitis* 10, 77–86.

McHale, C.M., Zhang, L., Thomas, R., and Smith, M.T. (2013). Analysis of the transcriptome in molecular epidemiology studies. *Environ. Mol. Mutagen.* 54, 500–517.

Meddens, C.A., Harakalova, M., van den Dungen, N.A.M., Foroughi Asl, H., Hijma, H.J., Cuppen, E.P.J.G., Björkegren, J.L.M., Asselbergs, F.W., Nieuwenhuis, E.E.S., and Mokry, M. (2016). Systematic analysis of chromatin interactions at disease associated loci links novel candidate genes to inflammatory bowel disease. *Genome Biol.* 17, 247.

Medina, I., Carbonell, J., Pulido, L., Madeira, S.C., Goetz, S., Conesa, A., Tárraga, J., Pascual-Montano, A., Nogales-Cadenas, R., Santoyo, J., et al. (2010). Babelomics: An integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucleic Acids Res.* 38, 210–213.

Miklos, G.L.G., and Rubin, G.M. (1996). The role of the genome project in determining gene function: Insights from model organisms. *Cell* 86, 521–529.

Molberg, Ø., Mcadam, S.N., Körner, R., Quarsten, H., Kristiansen, C., Madsen, L., Fugger, L., Scott, H., Norén, O., Roepstorff, P., et al. (1998). Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat. Med.* 4, 713–717.

Monteleone, G., Pender, S.L., Alstead, E., Hauer, A.C., Lionetti, P., McKenzie, C., and MacDonald, T.T. (2001). Role of interferon alpha in promoting T helper cell type 1 responses in the small intestine in coeliac disease. *Gut* 48, 425–429.

Monteleone, I., Sarra, M., Del Vecchio Blanco, G., Paoluzi, O.A., Franzè, E., Fina, D., Fabrizi, A., MacDonald, T.T., Pallone, F., and Monteleone, G. (2010). Characterization of IL-17A–Producing Cells in Celiac Disease Mucosa. *J. Immunol.* 184, 2211–2218.

Mumbach, M.R., Satpathy, A.T., Boyle, E.A., Dai, C., Gowen, B.G., Cho, S.W., Nguyen, M.L., Rubin, A.J., Granja, J.M., Kazane, K.R., et al. (2017). Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat. Genet.* 49, 1602–1612.

Murray, J.A., Van Dyke, C., Plevak, M.F., Dierkhising, R.A., Zinsmeister, A.R., and Melton, L.J. (2003). Trends in the identification and clinical features of

celiac disease in a North American community, 1950-2001. *Clin. Gastroenterol. Hepatol.* 1, 19–27.

Murray, J.A., Watson, T., Clearman, B., and Mitros, F. (2004). Effect of a gluten-free diet on gastrointestinal symptoms in celiac disease. *Am. J. Clin. Nutr.* 79, 669–673.

Nanayakkara, M., Lania, G., Maglio, M., Auricchio, R., De Musis, C., Discepolo, V., Miele, E., Jabri, B., Troncone, R., Auricchio, S., et al. (2018). P31–43, an undigested gliadin peptide, mimics and enhances the innate immune response to viruses and interferes with endocytic trafficking: a role in celiac disease. *Sci. Rep.* 8, 10821.

Nilsen, E.M., Lundin, K.E., Krajci, P., Scott, H., Sollid, L.M., and Brandtzaeg, P. (1995). Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma. *Gut* 37, 766–776.

Nistico, L., Fagnani, C., Coto, I., Percopo, S., Cotichini, R., Limongelli, M.G., Paparo, F., D’Alfonso, S., Giordano, M., Sferlazzas, C., et al. (2006). Concordance, disease progression, and heritability of coeliac disease in Italian twins. *Gut* 55, 803–804.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385.

Ong, C.T., and Corces, V.G. (2014). CTCF: An architectural protein bridging genome topology and function. *Nat. Rev. Genet.* 15, 234–246.

Özcan, U., Yilmaz, E., Özcan, L., Furuhashi, M., Vaillancourt, E., Smith, R.O., Görgün, C.Z., and Hotamisligil, G.S. (2006). Chemical chaperones reduce ER stress and restore glucose homeostasis in a mouse model of type 2 diabetes. *Science* (80-.). 313, 1137–1140.

Paoletta, G., Lepretti, M., Barone, M.V., Nanayakkara, M., Di Zenzo, M., Sblattero, D., Auricchio, S., Esposito, C., and Caputo, I. (2017). Celiac anti-type 2 transglutaminase antibodies induce differential effects in fibroblasts from celiac disease patients and from healthy subjects. *Amino Acids* 49, 541–550.

Parmar, A., Greco, D., Venäläinen, J., Gentile, M., Dukes, E., and Saavalainen, P. (2013). Gene Expression Profiling of Gliadin Effects on Intestinal Epithelial Cells Suggests Novel Non-Enzymatic Functions of Pepsin and Trypsin. *PLoS One* 8, e66307.

Paterson, B.M., Lammers, K.M., Arrieta, M.C., Fasano, A., and Meddings, J.B. (2007). The safety, tolerance, pharmacokinetic and pharmacodynamic effects of

single doses of AT-1001 in coeliac disease subjects: a proof of concept study. *Aliment. Pharmacol. Ther.* *26*, 757–766.

Pervjakova, N., and Prokopenko, I. (2017). The TAD-pathway for GWAS signals. *Eur. J. Hum. Genet.* *25*, 1179–1180.

Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* *153*, 1281–1295.

Picarelli, A., Di Tola, M., Sabbatella, L., Anania, M.C., Di Cello, T., Greco, R., Silano, M., and De Vincenzi, M. (1999). 31-43 amino acid sequence of the alpha-gliadin induces anti-endomysial antibody production during in vitro challenge. *Scand. J. Gastroenterol.* *34*, 1099–1102.

Pietz, G., De, R., Hedberg, M., Sjöberg, V., Sandström, O., Hernell, O., Hammarström, S., and Hammarström, M.-L. (2017). Immunopathology of childhood celiac disease—Key role of intestinal epithelial cells. *PLoS One* *12*, e0185025.

Plaza-Izurieta, L., Fernandez-Jimenez, N., Irastorza, I., Jauregi-Miguel, A., Romero-Garmendia, I., Vitoria, J.C.J.C., and Bilbao, J.R.J.R. (2015). Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in celiac disease. *Eur. J. Hum. Genet.* *23*, 1–6.

Ploski, R., Ek, J., Thorsby, E., and Sollid, L.M. (1993). 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* *41*, 173–177.

Plot, L., and Amital, H. (2009). Infectious associations of Celiac disease. *Autoimmun. Rev.* *8*, 316–319.

Powrózek, T., Mlak, R., Dziedzic, M., Małecka-Massalska, T., and Sagan, D. (2018). Investigation of relationship between precursor of miRNA-944 and its mature form in lung squamous-cell carcinoma - the diagnostic value. *Pathol. Res. Pract.* *214*, 368–373.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* *26*, 841–842.

Quinn, E.M., Coleman, C., Molloy, B., Dominguez Castro, P., Cormican, P., Trimble, V., Mahmud, N., and McManus, R. (2015). Transcriptome Analysis of CD4+ T Cells in Coeliac Disease Reveals Imprint of BACH2 and IFN γ Regulation. *PLoS One* *10*, e0140049.

Rabassa, E.B., Sagaró, E., Fragoso, T., Castañeda, C., and Gra, B. (1981). Coeliac disease in Cuban children. *Arch. Dis. Child.* *56*, 128–131.

- Ráki, M., Schjetne, K.W., Stammaes, J., Molberg, Jahnsen, F.L., Issekutz, T.B., Bogen, B., and Sollid, L.M. (2007). Surface expression of transglutaminase 2 by dendritic cells and its potential role for uptake and presentation of gluten peptides to T cells. *Scand. J. Immunol.* *65*, 213–220.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* *8*, 2281–2308.
- Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* *159*, 1665–1680.
- Rauhavirta, T., Qiao, S.W., Jiang, Z., Myrsky, E., Loponen, J., Korponay-Szabó, I.R., Salovaara, H., Garcia-Horsman, J.A., Venäläinen, J., Männistö, P.T., et al. (2011). Epithelial transport and deamidation of gliadin peptides: A role for coeliac disease patient immunoglobulin A. *Clin. Exp. Immunol.* *164*, 127–136.
- Robertson, K.D. (2005). DNA methylation and human disease. *Nat. Rev. Genet.* *6*, 597–610.
- Rodón, L., González-Juncà, A., Del Mar Inda, M., Sala-Hojman, A., Martínez-Sáez, E., and Seoane, J. (2014). Active creb1 promotes a malignant TGFβ2 autocrine loop in glioblastoma. *Cancer Discov.* *4*, 1230–1241.
- Rodríguez-Fraticelli, A.E., Bagwell, J., Bosch-Fortea, M., Boncompain, G., Reglero-Real, N., García-León, M.J., Andrés, G., Toribio, M.L., Alonso, M.A., Millán, J., et al. (2015). Developmental regulation of apical endocytosis controls epithelial patterning in vertebrate tubular organs. *Nat. Cell Biol.* *17*, 241–250.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res.* *14*, 1902–1910.
- Rosenfeld, M.G., Lunnyak, V. V., and Glass, C.K. (2006). Sensors and signals: A coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev.* *20*, 1405–1428.
- Rubio-Tapia, A., and Murray, J.A. (2010). Classification and management of refractory coeliac disease. *Gut* *59*, 547–557.
- Di Sabatino, A., and Corazza, G.R. (2009). Coeliac disease. *Lancet* *373*, 1480–1493.
- Sagaró, E., and Jimenez, N. (1981). Family studies of coeliac disease in Cuba. *Arch. Dis. Child.* *56*, 132–133.
- Samasca, G., Sur, G., Lupan, I., and Deleanu, D. (2014). Gluten-free diet and

quality of life in celiac disease. *Gastroenterol. Hepatol. from Bed to Bench* 7, 139–143.

Sari, C., Bayram, N.A., Doğan, F.E.A., Baştuğ, S., Bolat, A.D., Sarı, S.Ö., Ersoy, O., and Bozkurt, E. (2012). The evaluation of endothelial functions in patients with celiac disease. *Echocardiography* 29, 471–477.

Sarra, M., Cupi, M.L., Monteleone, I., Franzè, E., Ronchetti, G., Di Sabatino, A., Gentileschi, P., Franceschilli, L., Sileri, P., Sica, G., et al. (2013). IL-15 positively regulates IL-21 production in celiac disease mucosa. *Mucosal Immunol.* 6, 244–255.

Savic, S., Ouboussad, L., Dickie, L.J., Geiler, J., Wong, C., Doody, G.M., Churchman, S.M., Ponchel, F., Emery, P., Cook, G.P., et al. (2014). TLR dependent XBP-1 activation induces an autocrine loop in rheumatoid arthritis synoviocytes. *J. Autoimmun.* 50, 59–66.

Sawle, A.D., Keschull, M., Demmer, R.T., and Papapanou, P.N. (2016). Identification of Master Regulator Genes in Human Periodontitis. *J. Dent. Res.* 95, 1010–1017.

Schmitges, F.W., Radovani, E., Najafabadi, H.S., Barazandeh, M., Campitelli, L.F., Yin, Y., Jolma, A., Zhong, G., Guo, H., Kanagalingam, T., et al. (2016). Multiparameter functional diversity of human C2H2 zinc finger proteins. *Genome Res.* 26, 1742–1752.

Schuppan, D., Junker, Y., and Barisani, D. (2009). Celiac Disease: From Pathogenesis to Novel Therapies. *Gastroenterology* 137, 1912–1933.

Sexton, T., Schober, H., Fraser, P., and Gasser, S.M. (2007). Gene regulation through nuclear organization. *Nat. Struct. Mol. Biol.* 14, 1049–1055.

Shalimar, Das, P., Sreenivas, V., Gupta, S.D., Panda, S.K., and Makharia, G.K. (2013). Mechanism of villous atrophy in celiac disease: Role of apoptosis and epithelial regeneration. *Arch. Pathol. Lab. Med.* 137, 1262–1269.

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G.M., Sollid, L.M., and Khosla, C. (2002). Structural basis for gluten intolerance in Celiac Sprue. *Science* (80-.). 297, 2275–2279.

Siggers, T., and Gordân, R. (2014). Protein-DNA binding: Complexities and multi-protein codes. *Nucleic Acids Res.* 42, 2099–2111.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., De Wit, E., Van Steensel, B., and De Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38, 1348–1354.

Sjöström, H., Lundin, K.E.A., Molberg, Körner, R., Mcadam, S.N., Anthonsen,

- D., Quarsten, H., Norén, O., Roepstorff, P., Thorsby, E., et al. (1998). Identification of a gliadin T-cell epitope in coeliac disease: General importance of gliadin deamidation for intestinal T-cell recognition. *Scand. J. Immunol.* *48*, 111–115.
- Sollid, L.M. (1989). Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J. Exp. Med.* *169*, 345–350.
- Sollid, L.M. (2002). Coeliac disease: Dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.* *2*, 647–655.
- Sollid, L.M., and Thorsby, E. (1993). HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology* *105*, 910–922.
- Spurkland, A., Sollid, L.M., Polanco, I., Vartdal, F., and Thorsby, E. (1992). HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. *Hum. Immunol.* *35*, 188–192.
- Steinman, L. (2007). A brief history of TH17, the first major revision in the TH1/TH2 hypothesis of T cell-mediated tissue damage. *Nat. Med.* *13*, 139–145.
- Stenman, S.M., Lindfors, K., Korponay-Szabo, I.R., Lohi, O., Saavalainen, P., Partanen, J., Haimila, K., Wieser, H., Mäki, M., and Kaukinen, K. (2008). Secretion of celiac disease autoantibodies after in vitro gliadin challenge is dependent on small-bowel mucosal transglutaminase 2-specific IgA deposits. *BMC Immunol.* *9*, 6.
- Sturgess, R., Day, P., Ellis, H.J., Kontakou, M., Ciclitira, P.J., Lundin, K.E.A., and Gjertsen, H.A. (1994). Wheat peptide challenge in coeliac disease. *Lancet* *343*, 758–761.
- Suliman, G.I. (1978). Coeliac disease in Sudanese children. *Gut* *19*, 121–125.
- Sun, B., Hu, L., Luo, Z.-Y., Chen, X.-P., Zhou, H.-H., and Zhang, W. (2016). DNA methylation perspectives in the pathogenesis of autoimmune diseases. *Clin. Immunol.* *164*, 21–27.
- Sun, W., Julie Li, Y.S., Huang, H. Da, Shyy, J.Y., and Chien, S. (2010). microRNA: A Master Regulator of Cellular Processes for Bioengineering Systems. *Annu. Rev. Biomed. Eng.* *12*, 1–27.
- Takeuchi, A., Badr, M.E.S.G., Miyauchi, K., Ishihara, C., Onishi, R., Guo, Z., Sasaki, Y., Ike, H., Takumi, A., Tsuji, N.M., et al. (2016). CRTAM determines the CD4⁺ cytotoxic T lymphocyte lineage. *J. Exp. Med.*
- Teng, G. gen, Wang, W. hong, Dai, Y., Wang, S. jun, Chu, Y. xiang, and Li, J. (2013). Let-7b Is Involved in the Inflammation and Immune Responses Associated with Helicobacter pylori Infection by Targeting Toll-Like Receptor 4. *PLoS One* *8*, e56709.

Toosi, S., Orlow, S.J., and Manga, P. (2012). Vitiligo-inducing phenols activate the unfolded protein response in melanocytes resulting in upregulation of IL6 and IL8. *J. Invest. Dermatol.* *132*, 2601–2609.

Tovar, H., García-Herrera, R., Espinal-Enríquez, J., and Hernández-Lemus, E. (2015). Transcriptional master regulator analysis in breast cancer genetic networks. *Comput. Biol. Chem.* *59*, 67–77.

Tran, H., Porter, J., Sun, M.A., Xie, H., and Zhang, L. (2014). Objective and comprehensive evaluation of bisulfite short read mapping tools. *Adv. Bioinformatics* *2014*, 472045.

Trenkmann, M., Brock, M., Gay, R.E., Michel, B.A., Gay, S., and Huber, L.C. (2013). Tumor necrosis factor α -induced microRNA-18a activates rheumatoid arthritis synovial fibroblasts through a feedback loop in NF- κ B signaling. *Arthritis Rheum.* *65*, 916–927.

Troncone, R., Gianfrani, C., Mazzarella, G., Greco, L., Guardiola, J., Auricchio, S., and De Berardinis, P. (1998). Majority of gliadin-specific T-cell clones from celiac small intestinal mucosa produce interferon- γ and interleukin-4. *Dig. Dis. Sci.* *43*, 156–161.

Trujillo, R.D., Yue, S.B., Tang, Y., O’Gorman, W.E., and Chen, C.Z. (2010). The potential functions of primary microRNAs in target recognition and repression. *EMBO J.* *29*, 3272–3285.

Trynka, G., Hunt, K.A., Bockett, N.A., Romanos, J., Mistry, V., Szperl, A., Bakker, S.F., Bardella, M.T., Bhaw-Rosun, L., Castillejo, G., et al. (2011). Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat. Genet.* *43*, 1193–1201.

Vaira, V., Roncoroni, L., Barisani, D., Gaudio, G., Bosari, S., Bulfamante, G., Doneda, L., Conte, D., Tomba, C., Bardella, M.T., et al. (2014). microRNA profiles in coeliac patients distinguish different clinical phenotypes and are modulated by gliadin peptides in primary duodenal fibroblasts. *Clin. Sci. (Lond.)* *126*, 417–423.

Vallejo-Diez, S., Bernardo, D., De Moreno, M.L., Muñoz-Suano, A., Fernández-Salazar, L., Calvo, C., Sousa, C., Garrote, J.A., Cebolla, Á., and Arranz, E. (2013). Detection of specific IgA antibodies against a novel deamidated 8-mer gliadin peptide in blood plasma samples from celiac patients. *PLoS One* *8*, e80982.

Valton, A.L., and Dekker, J. (2016). TAD disruption as oncogenic driver. *Curr. Opin. Genet. Dev.* *36*, 34–40.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., and Luscombe, N.M. (2009). A census of human transcription factors: Function, expression and

evolution. *Nat. Rev. Genet.* *10*, 252–263.

Vriezinga, S.L., Auricchio, R., Bravi, E., Castillejo, G., Chmielewska, A., Crespo Escobar, P., Kolaček, S., Koletzko, S., Korponay-Szabo, I.R., Mummert, E., et al. (2014). Randomized Feeding Intervention in Infants at High Risk for Celiac Disease. *N. Engl. J. Med.* *371*, 1304–1315.

Wal, Y. Van De, Kooy, Y., Veelen, P. Van, Peña, S., Mearin, L., Papadopoulos, G., and Koning, F. (1998). Selective Deamidation by Tissue Transglutaminase Strongly Enhances Gliadin-Specific T Cell Reactivity. *J. Immunol.* *161*, 1585–1588.

Wan, J., Oliver, V.F., Wang, G., Zhu, H., Zack, D.J., Merbs, S.L., and Qian, J. (2015). Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation. *BMC Genomics* *16*, 49.

Wang, J., Zhuang, J., Iyer, S., Lin, X.Y., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* *22*, 1798–1812.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* *10*, 57–63.

Wapenaar, M.C., Monsuur, A.J., Bodegraven, A.A. Van, Weersma, R.K., Bevova, M.R., Linskens, R.K., Howdle, P., Holmes, G., Mulder, C.J., Dijkstra, G., et al. (2008). Associations with tight junction genes PARD3 and MAGI2 in Dutch patients point to a common barrier defect for coeliac disease and ulcerative colitis Associations with tight junction genes PARD3 and MAGI2 in Dutch patients point to a common barrier defect. *57*, 463–467.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and Inference of Eukaryotic Transcription Factor Sequence Specificity. *Cell* *158*, 1431–1443.

Weischenfeldt, J., Dubash, T., Drainas, A.P., Mardin, B.R., Chen, Y., Stütz, A.M., Waszak, S.M., Bosco, G., Halvorsen, A.R., Raeder, B., et al. (2017). Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* *49*, 65–74.

Wen, A.Y., Sakamoto, K.M., and Miller, L.S. (2010). The Role of the Transcription Factor CREB in Immune Function. *J. Immunol.* *185*, 6413–6419.

Widschwendter, M., Jiang, G., Woods, C., Müller, H.M., Fiegl, H., Goebel, G., Marth, C., Müller-Holzner, E., Zeimet, A.G., Laird, P.W., et al. (2004). DNA hypomethylation and ovarian cancer biology. *Cancer Res.* *64*, 4472–4480.

Wingender, E. (2000). TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res.* 28, 316–319.

Wingender, E., Schoeps, T., Haubrock, M., and Dönitz, J. (2015). TFClass: A classification of human transcription factors and their rodent orthologs. *Nucleic Acids Res.* 43, D97–D102.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: MicroRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* 11, 228–234.

Yadav, D., Ngolab, J., Lim, R.S.-H., Krishnamurthy, S., and Bui, J.D. (2008). Cutting Edge: Down-Regulation of MHC Class I-Related Chain A on Tumor Cells by IFN- γ -Induced MicroRNA. *J. Immunol.* 182, 39–43.

Yang, D., Xiao, C., Long, F., Su, Z., Jia, W., Qin, M., Huang, M., Wu, W., Suguro, R., Liu, X., et al. (2018). HDAC4 regulates vascular inflammation via activation of autophagy. *Cardiovasc. Res.* 114, 1016–1028.

Yang, J., Yu, H., Liu, B.H., Zhao, Z., Liu, L., Ma, L.X., Li, Y.X., and Li, Y.Y. (2013). DCGL v2.0: an R package for unveiling differential regulation from differential co-expression. *PLoS One* 8, e79729.

Yao, K., Lee, S.Y., Peng, C., Lim, D.Y., Yamamoto, H., Ryu, J., Lim, T.G., Chen, H., Jin, G., Zhao, Z., et al. (2018). RSK2 is required for TRAF6 phosphorylation-mediated colon inflammation. *Oncogene* 37, 3501–3513.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* (80-.). 356.

Yue, S.B., Trujillo, R.D., Tang, Y., O’Gorman, W.E., and Chen, C.Z. (2011). Loop nucleotides control primary and mature miRNA function in target recognition and repression. *RNA Biol.* 8, 1115–1123.

Zeilinger, S., Kühnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., Weidinger, S., Lattka, E., Adamski, J., Peters, A., et al. (2013). Tobacco Smoking Leads to Extensive Genome-Wide Changes in DNA Methylation. *PLoS One* 8, e63812.

Zhang, S., Shan, C., Kong, G., Du, Y., Ye, L., and Zhang, X. (2012). MicroRNA-520e suppresses growth of hepatoma cells by targeting the NF- κ B-inducing kinase (NIK). *Oncogene* 31, 3607–3620.

Zhang, Y., Schottker, B., Ordonez-Mena, J., Holleczeck, B., Yang, R.X., Burwinkel, B., Butterbach, K., and Brenner, H. (2015). F2RL3 methylation, lung cancer incidence and mortality. *Int. J. Cancer* 137, 1739–1748.

Zhao, N., Wang, R., Zhou, L., Zhu, Y., Gong, J., and Zhuang, S.M. (2014). MicroRNA-26b suppresses the NF- κ B signaling and enhances the chemosensitivity of hepatocellular carcinoma cells by targeting TAK1 and TAB3. *Mol. Cancer* 13, 35.

Zhou, R., Gong, A.Y., Chen, D., Miller, R.E., Eischeid, A.N., and Chen, X.M. (2013). Histone Deacetylases and NF- κ B Signaling Coordinate Expression of CX3CL1 in Epithelial Cells in Response to Microbial Challenge by Suppressing miR-424 and miR-503. *PLoS One* 8, e65153.

