



IKER
GAZTE
NAZIOARTEKO
IKERKETA EUSKARAZ

V. IKERGAZTE

NAZIOARTEKO IKERKETA EUSKARAZ

2023ko maiatzaren 17, 18 eta 19a
Donostia, Euskal Herria

ANTOLATZAILEA:
Udako Euskal Unibertsitatea (UEU)



Aitortu-PartekatuBerdin 3.0

ZIENTZIAK ETA NATURA ZIENTZIAK

**1 motako diabetesarekin
asoziatu-rikoko RNA ez kodetzaile
batak mikropeptido bat ekoiztu
dezake**

*Jon Mentxaka Salgado,
Koldo García Etxebarria,
Henar Rojas Marquez,
Ane Olazagoitia Garmendia,
Leire Bergara,
Luis Manuel Mendoza Gomez,
Ainara Castellanos Rubio
eta Izortze Santin Gomez*

239-246 or.

<https://dx.doi.org/10.26876/ikergazte.v.05.30>

ANTOLATZAILEA:



BABESLEAK:



LAGUNTZAILEAK:



1 motako diabetesarekin asoziatuako RNA ez kodetzaile batek mikropeptido bat ekoiztu dezake

Jon Mentxaka^{1,2}, Koldo García³, Henar Rojas^{1,2}, Ane Olazagoitia¹, Leire Bergara¹, Luis Manuel Mendoza², Ainara Castellanos^{1,4,5}, Izortze Santin^{1,2,5}

¹Euskal Herriko Unibertsitatea UPV/EHU, ²Instituto Biocruces Bizkaia, ³Biodonostia Health Research Institute, ⁴IKERBASQUE Basque Foundation for Science, ⁵CIBERDEM
jon.mentxaka@ehu.eus

Laburpena

1 motako diabetesa oso zabaldurik dagoen eritasun autoimmune multifaktorial bat da, zeinean infekzio enterobiralek patologiaren garapenean eragiten duten genetikoki suszeptibleak diren indibiduoetan. Testuinguru honetan, ohituraz ez kodetzaile gisa sailkatzen diren RNA geneetatik ekoiztu litezkeen mikropeptidoak gaixotasunaren garapenean garrantzitsuak izan litezkeela uste dugu. Hipotesi honekin, erribosomei atxikiriko RNAREN sekuentziazioa burutu dugu pankreako β zelulen lerro batean, infekzio birala imitatzen duen baldintzen pean. Tresna bioinformatikoen bidez mikropeptidoren bat kodetzeko aukera duten hautagaien zerrenda bat lortu ez ezik, metodoa balioztatu dugu hautagai batetik eratorritako peptido baten ekoizpena *in vitro* frogaturik.

1 motako diabetesa, RNA luze ez kodetzailea, mikropeptidoak, pankreako β zelulak

Abstract

Type 1 diabetes is a chronic, multifactorial and autoimmune disease in which enteroviral infections seem to be a crucial triggering factor in genetically susceptible individuals. Although traditionally disregarded, non-coding RNAs have emerged as key players in the disease, especially those which can produce small peptides from short open reading frames. In this study we have sequenced ribosome-bound RNA in a pancreatic β cell line under basal and viral-insult-mimicking conditions. By usage of several bioinformatical tools, we have generated a list of RNA gene candidates with micropeptide coding potential. Additionally, we have validated this approach in vitro by identifying a non-coding RNA gene which can produce a putative 109 aminoacid peptide.

Type 1 diabetes, long non-coding RNA, micropeptides, pancreatic β cells

1. Sarrera eta motibazioa

1 motako diabetes mellitusa (1DM) gaur egun gero eta zabalduago dagoen eritasun kroniko eta autoimmune bat da. Bere ezaugarri nagusia pankreako β zelulen suntsipen progresiboa da, immunitate-sistemako zelulek eragindako pankreako irlatxoaren kontrako eraso jarraituak direla eta (Eizirik *et al*, 2009). β zelulen galerak eragindako odoleko glukosa maila altua orekatu ahal izateko, 1DMren pazienteak bizitza osorako intsulina tratamendua jasotzera beharturik daude, izan ere oraindik ez da 1DMri aurre egiteko sendabiderik ezagutzen.

Beste gaixotasun askoren antzera, 1DM gaixotasun multifaktoriala da; hau pairatzeko arriskuak osagai genetikoki garrantzitsuak eduki arren (Todd *et al*, 2007), genetikak soilik ezin dezake gaixotasunaren herentzia osoa azaldu. Badirudi ingurugiroko zenbait faktorek zeresan handia dutela, batez ere gaixotasunaren hasierako etapetan. Zehazki, ikerketa anitzek enterobirus infekzioen eta 1DMren garapenean arteko lotura azpimarratu dute predisposizio genetikoki esangarria daukaten indibiduoetan (Op de Beeck eta Eizirik, 2016, Yeung *et al*, 2011). Pare bat adibide jartzearen, 1DMrekin diagnostikatu-berri izan diren pazienteen odolean birus hauen RNA detektatu izan da ikerketa batean eta, beste batean, birusen kapsidaren VP1 proteina 1DM daukaten pazienteetan, indibiduo osasuntsuetan baino proportzio askoz altuagoarekin agertzen dela egiaztatzen da (Richardson *et al*, 2009).

Hala eta guztiz ere, ez ditugu ezagutzen 1DMren garapena baldintzatzen duten aldagai guztiak. Azken urteotan, gero eta arruntagoak bihurtu dira genoma osoko asoziazio ikerketak (ingelesetik GWAS; *genome wide association studies*). Haietan, fenotipo edo gaixotasun zehatzen eta

genomako nukleotido bakarreko polimorfismoen (ingelesetik SNP; *single nucleotide polymorphism*) arteko lotura edo asoziazioa aztertzen da, gaixotasun horietan garrantzia nabarmena izan dezaketen SNPak identifikatu nahian. Bitxiki, SNP hauetako asko (%10a baino gehiago), proteina kodetzaileak diren geneetan egon beharrean, RNA luze ez-kodetzaile (ingelesetik lncRNA; *long non-coding RNA*) deritzen geneetan kokatzen dira. Ustez, gene hauek proteina funtzionalak sortzeko gai ez diren arren, gero eta gehiago dira zelula osasuntsuen funtzioan baita gaixotasunen garapenean duten garrantzia deskribatzen duten artikuluak (Széll *et al*, 2016). 1DM ez da salbuespena noski, eta gaixotasunaren garapenarekin erlazionatuta dauden lncRNA anitz deskribatu dira, beste hainbeste oraindik aurkitzeke daudelarik (Akerman *et al*, 2017, Motterle *et al*, 2015).

Gure ikerketan, 1DMren garapenarekin loturik egon daitezkeen lncRNAak aztertu ditugu, baina oraingoan pausu bat haratago joanda. Azken urteotan gene kodetzaile eta ez kodetzaileen paradigma tradizionala aldatuz joan da, sekuentziazio teknologia berrien agerpenarekin batera. Erribosomei loturiko RNAREN sekuentziazioak antzina nekez imajinatu genitzaieen ustekabeak aurkeztu ditu, izan ere, lncRNA mordoa antzeman dira erribosomei loturik (Ingolia *et al*, 2011). Beste era batera esanda, ustez proteinak kodetzeko gai ez diren transkrito horiek, zenbait baldintzaren pean behintzat, peptidoak edo proteinak ekoizteko ahalmena izan dezakete (Bazzini *et al*, 2014, Anderson *et al*, 2015). Hain zuzen ere, honelako lncRNA geneen bila aritu gara gure ikerketan, zehazki 100 aminoazido inguru edo gutxiagoko mikropeptidoak kodetu ditzaketenen lncRNA-en bila. Horretarako, ohiko metodo neketsua ordeztu dezakeen metodo sinple eta azkarrago bat diseinatu dugularik.

2. Arloko egoera eta ikerketaren helburuak

lncRNA geneetatik eratorritako mikropeptidoek funtzio garrantzitsuak bete ditzakete; literaturan zenbait kasu daude deskribaturik (Li *et al*, 2021, Senís *et al*, 2021). Guri dagokigunez, 1 motako diabetes mellitusarekin erlazionaturiko mikropeptidoak identifikatu nahiko genituzke. Geure hipotesia hurrengoa da: Azaldu dudanez, arrisku genetikoak duten indibiduoetan, infekzio enterobiralak 1DMren abiarazle izan daitezke. Mekanismoa ezezaguna bada ere, egoera horretan baliteke normalean erribosomearekin interakziorik ez daukaten lncRNA molekulek erribosomara lotzeko joera izatea; izan ere deskribaturik dago estres egoeretan erribosomen erregulazioa galtzen dela, egoera normalen itzultzen ez diren transkritoak itzuli ditzaketelarik (Thomaidou *et al*, 2021). Honen ondorioz, gaixotasunaren garapenean eragin dezaketen mikropeptidoak ekoiztu daitezke; adibidez, pankreako β zelulen neoautoantigeno bezala funtzionatu ditzaketanak.

Horrela, lan honetan bi helburu nagusi finkatu ditugu:

- 1) Potentzial kodetzailea duten lncRNA gene hautagaiak identifikatzeko metodo bat garatzea.
- 2) Garatutako metodoan oinarrituz, pankreako β zeluletan 1DM-aren garapenean eragina izan ditzaketenen mikropeptidoak identifikatzea.

3. Ikerketaren muina

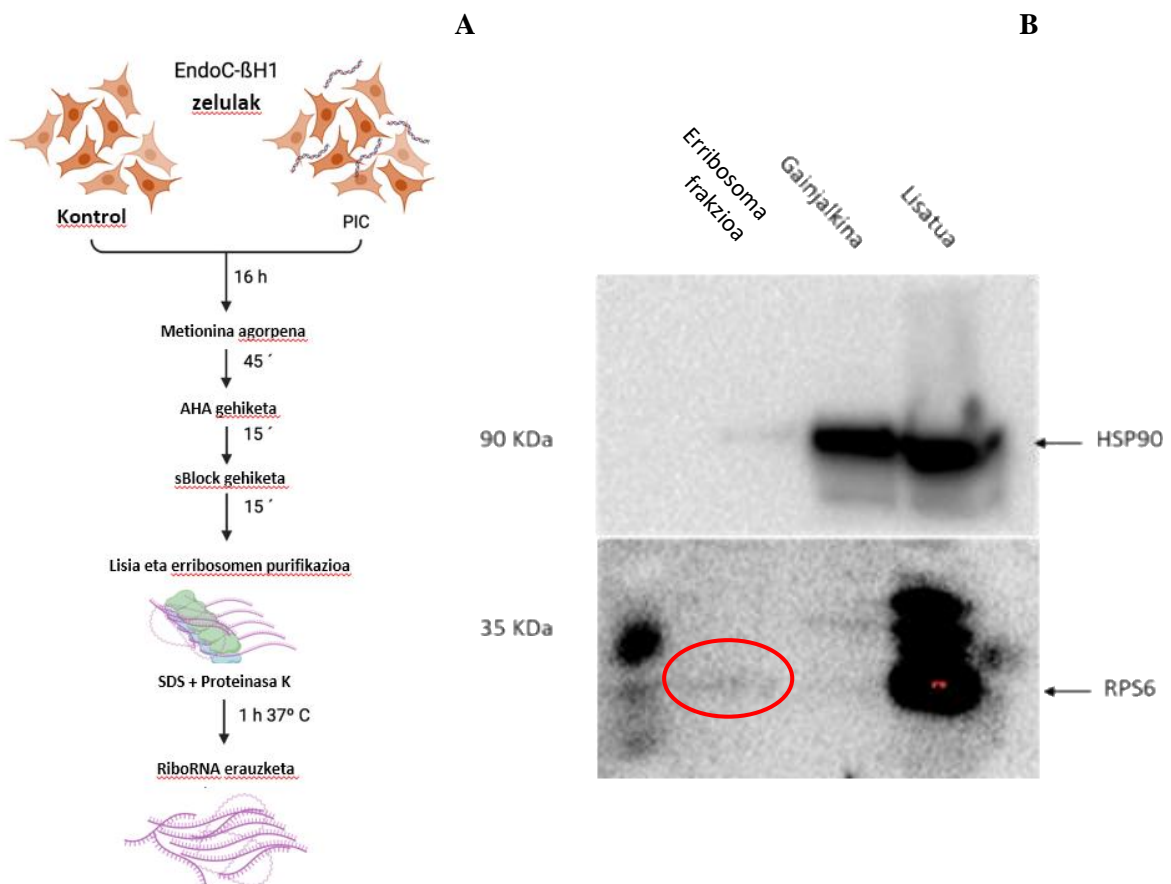
3.1 Erribosomei asoziatutako RNAREN erauzketa

Lehenik eta behin, gure ikerketa garatzeko, 1DM gaixotasunaren hasieran pankreako β zelulek jasaten duten egoera errealarri nolabait hurbiltzen zaion *in vitro* ingurunea prestatu dugu. Gaixotasunaren etapa goiztiarretan genetikoki suszeptibleak diren indibiduoetan infekzio biralak eman ohi dira. Egoera hori irudikatuzko, EndoC- β H1 deitzen den pankreako giza β zelula-lerroa erabili dugu. Infekzio biralaren ordeztu, azido poliinosiniko:polizitidiliko (PIC) deritzon harizpi bikoitzeko RNA sintetikoak transfektatu diegu zelulei, benetako birus infekzio baten aurrean ematen den erantzun zelularra abiarazteko.

Ondoren, erribosomari asoziatutako RNA erauzteari ekin genion. Erribosoman dagoen RNA aztertzeke ohiko metodoa “*Ribosome profiling*” deiturikoa da, zeinean, zelulak lisatu, RNAsarekin tratatu eta erribosomak ultrazentrifugazio bidez sedimentatzen diren, azkenik erribosomen barruan

babesturik dauden RNA zati motzak erauzi eta sekuentziatzeko. Nahiko prozesu neketsua eta konplexua denez, erribosomak “afinitatez” isolatzen dituen metodo komertzial bat erabili dugu: AHARIBO (Minati *et al*, 2021) (1A. Irudia). Horretarako, PICekin trataturiko eta tratatu gabeko zelulei hazkuntza medioa aldatu genien 16 ordutara, metionina gabeko medio bat jarri. Behin zelulen metionina erreserbak agortuta, metioninaren homologoa den molekula bat gehitu zitzaion medioari (AHA, L-azidohomoalanina), zelulek metioninaren ordean erabili zezaten itzultzeko prozesuan ari ziren proteina eta peptido jaioberrietan. Azkenik, erribosomen elongazioa eta itzulpena blokeatzen duen molekula txikia (sBlock, AHARIBOren molekula propioa) gehitu genion medioari, erribosoma-RNA-peptido konplexu horiek “izozteko”. EndoC-βH1 zelulak lisatu ziren, eta erribosomen purifikazioa AHArekin afinitate ligando batez estalitako bihi magnetikoak erabiliz lortu zen (1B. Irudia). Erribosomei asoziatutako RNA (riboRNA hemendik aurrera) hori erauzteko fenol:kloroformo:alkohol isoamiliko erauzketa tradizionala erabili genuen.

1. irudia. Erribosomei asoziatutako RNAREN prestaketaren diagrama (A) eta RPS6 erribosoma proteinaren western blot baten irudia, HSP90 karga kontrol gisa erabilia (B). Gorriz markatuta RPS6 proteina erribosomikoari dagokion banda, erribosoma frakzioan aberastua.



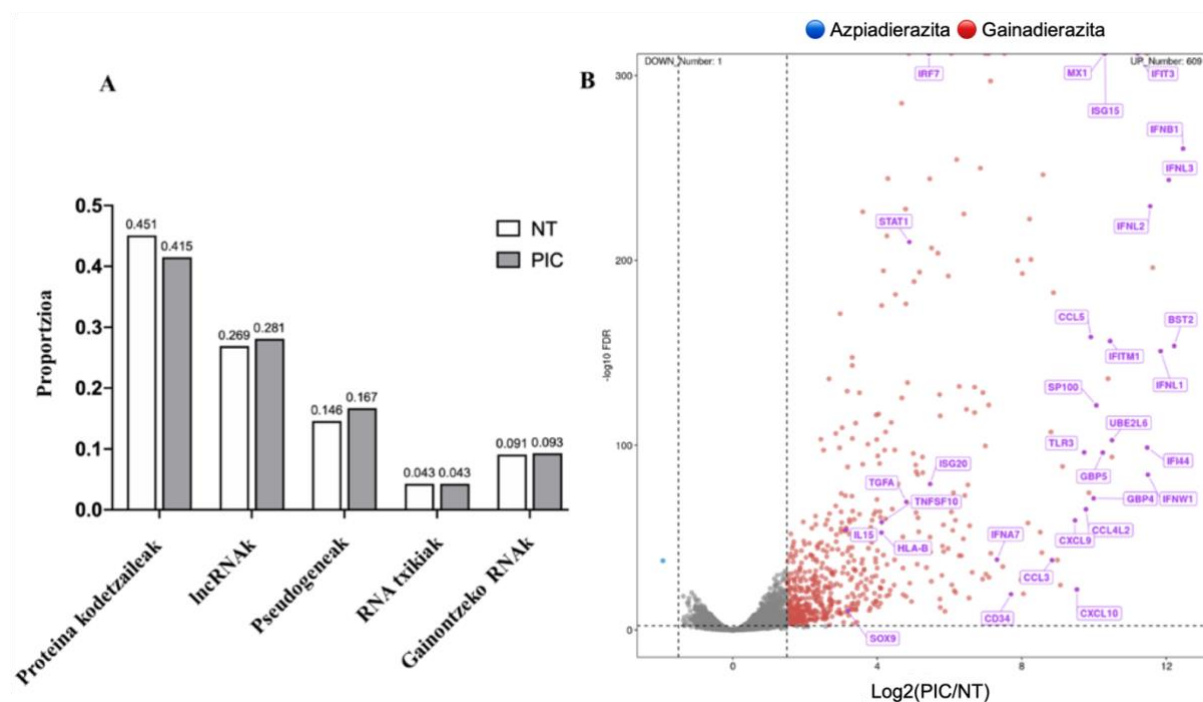
Azaldutako prozesu honekin 4 RNA lagin bikote prestatu genituen, PICekin trataturikoa eta kontrola bakoitzean. Gure riboRNAREN kalitatea frogatuta eta onargarriztat hartuta, sekuentziatzera bidali genituen, RNAseq arrunt batean. Batez beste, kalitate onargarria zuten 131 milioi irakurketa ekoiztu ziren lagin bakoitzeko.

3.2 RNAseq datuen analisia

Sekuentziazioaren analisiari ekinik, irakurketa gehienak gene proteina kodetzaileei zegozkien (2A. irudia). Hauetatik 614 zeuden esangarritasun estatistikoarekin areagotuta PIC laginetan (\log_2 Fold Change > 1.5 eta FDR<0.05) eta infekzio birikoa simulatzeko PICaren eraginkortasuna frogatuz,

erantzun antibiral eta immunearekin erlazioaturiko geneak esangarriki areagotutako geneen artean oso aberasturik zeudela behatu genuen (2B. irudia).

2. irudia. RNAseq bitartez detektaturiko transkrito guztien biotipoen proportzioak (A) eta proteina kodetzailen geneen *volcano plot* (B). Diferenzialki areagotutako gene askok erantzun immune eta antibiraleko seinaleztapen bidezidorretan dihardute



Beste alde batetik, lncRNA proportzio altua detektatu genuen, detektaturiko transkrito guztien %27 inguru zirelarik (2A. Irudia). Gainera PIC laginetan erribosomekin asoziatutako lncRNA kopurua kontrol laginetakoa baino handiago zen, estres egoeren pean erribosomaren interakzio ezohikoen teoria babestuz. Aurrekoen antzera, 454 lncRNA aurkitu genituen erribosomekiko asoziazioan esangarriki areagotuta PIC laginetan eta beste 470 lncRNA PIC laginetan soilik detektatu genituen.

lncRNA hauek pankreako β zelulen hanturan parte hartzen duten hautagaiak aurkitzeko abiapuntu ona izan daitezke, baina azkar aztertzeko gehiegi dira. Bilaketa fintzeko, gure hautagaien zerrenda, aurretiaz EndoC- β H1 zeluletan egindako *ribosome profiling* esperimentu batean (Thomaidou *et al*, 2021) identifikaturiko lncRNAekin eta GWAS ikerketetan 1 motako diabetesari loturiko SNPak dituzten lncRNAekin gurutzatu genuen. Horrela, interes bereziko 52 hautagairaino mugatu genuen zerrenda.

Azkenik, interes handia ezezik, mikropeptidoren bat kodetzeko ahalmena ere ebaluatu genuen gure hautagaien artean. Horretarako tresna eta baliabide bioinformatiko pare bat baliatu genituen. Alde batetik, nukleotido eta kodoien kontserbazio filogenetikoa kontuan hartuta, transkritoak puntuatzen dituen PhyloCSF programa eta, beste alde batetik, intereseko transkritoen irakurketa-ereduetatik (ORF, *open reading frame*) itzuli daitezkeen peptido edo proteinak analizatzen dituen CPC2 programa. Honela, gure irizpideen arabera, mikropeptido bat kodetzeko probabilitate handiena zuten hamar hautagai bildu genituen.

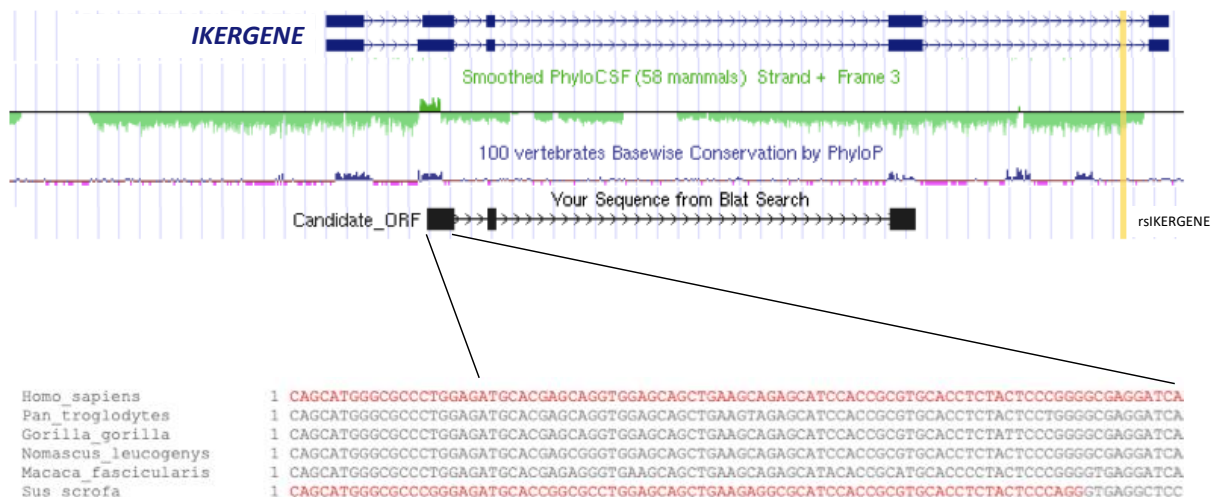
3.3 Metodoaren balioztapena

Hautagai horietatik lauk besterik ez zuten ATG trinukleotidoa euren mikropeptido potentzialaren hasiera kodoi gisa. Gauzak erraztearren, hasiera kanonikoa zeukatenekin jarraitu genuen. Horien artean *CRNDE* izeneko lncRNA genea agertu zen, gure ikerketaren arabera 84 aminoazidoko luzera duen

mikropeptido bat kodetzeko gai izan daitekeen hautagaia. Hain zuzen ere, gene hau jatorriz lncRNA bezala dago anotaturik, haatik, arestiko ikerlan pare batean HeLa zeluletan 84 aminoazidoko mikropeptido hori bera benetan kodetu dezakeela frogatu da (Szafron *et al*, 2015). Honek sendotasuna ematen dio orain arte jarraitutako estrategiari.

Horretaz gain, beste hautagai bat aukeratu genuen, hemen *IKERGENE* deitutako mikropeptidoa. *IKERGENE* 1DMrekin asoziatutako SNP bat dauka (rsIKERGENE), isoformen arabera exonikoa edo intronikoa dena. PhyloCSF algoritmoaren arabera, *IKERGENE* kodoi kontserbazio nabarmena daukan eskualde bat du, hirugarren exoian hasita (3. Irudia). Transkriptoaren sekuentzia CPC2 algoritmoan sartuta, kodetzailea izan daitekeen 330 nukleotidoko ORF bat identifikatu dezakegu. Intereseko ORF hori UCSC Genome Browserrean irudikatuta, agerikoa da ORFaren hasiera eta PhyloCSF zein kontserbazio gailurrak bat datozela (3. Irudia). Bestalde, hasierako ORFaren hasierako trinukleotidoa inguratzen duen sekuentzia, *Kozak* sekuentzia alegia, nahiko indartsua da TISPredictor programaren arabera.

3. Irudia. IKERGENE hautagaiaren kontestu genomikoa UCSC Genome Browserrean. Berdez PhyloCSF puntuazio tracka, puntuazio positiboko pikoarekin. Behean ORFaren hasieraren sekuentzia zenbait primate espezieetan (hasiera kodoia horiz).

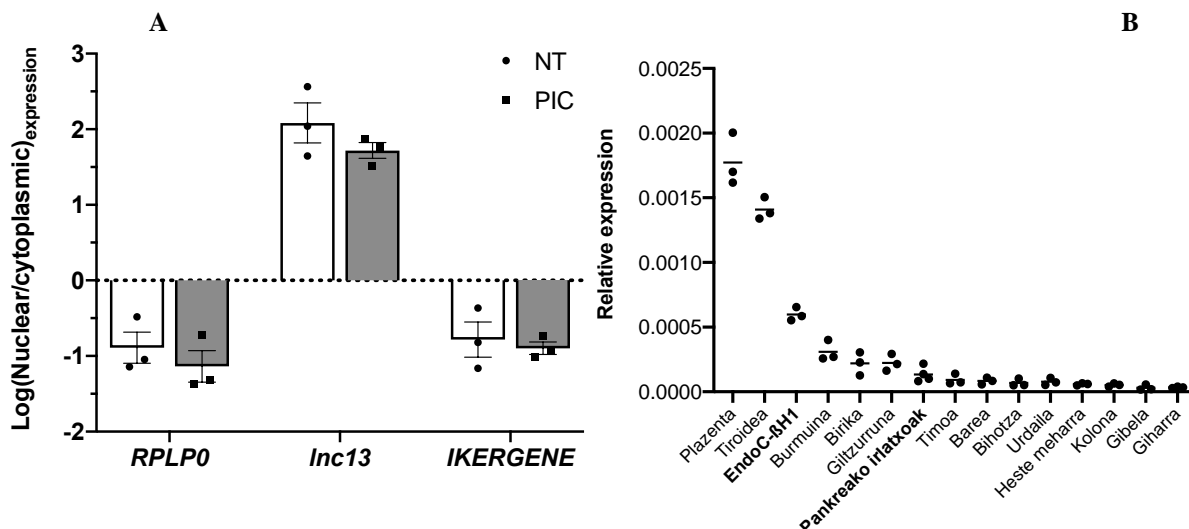


Hautagai honen aldeko ebidentziak ikusita, lncRNAren adierazpena aztertu genuen jarraian qPCRz primer espezifikoak erabilia, zenbait giza-ehunetan eta EndoC-βH1 zelulen frakzio azpizelularretan ere (4B. Irudia). Timo eta plazentan detektatu genuen adierazpen maila altuena, aztertutako beste ehun gehienetako adierazpen maila nahiko baxua zelarik. EndoC-βH1 zeluletan baita indibiduo osasuntsuen irlatxo pankreatikoetatik erauzitako RNA laginetan, *IKERGENE* transkriptoaren adierazpen maila ehun gehienena baino nabarmen altuago zen.

Adierazpen azpizelularrari dagokionez, nukleoan aberatsagoak diren lncRNA transkriptoak gene adierazpenaren erregulazioarekin lotu ohi dira, zitoplasman aberatsak direnak beste funtzio zelularrekin lotzen diren bitartean. Zentzuzkoa denez, mikropeptidoren bat kodetzeko gai diren lncRNAk zitoplasman daude aberastuta, RNA mezularien itzulpena zitoplasman dauden erribosometan gertatzen baita. Gure kasuan, *IKERGENE* bereziki β zelulen zitoplasman adierazten dela baieztatu genuen, bai egoera basalean baita PIC-ekin transfektatutako zeluletan (4A. Irudia).

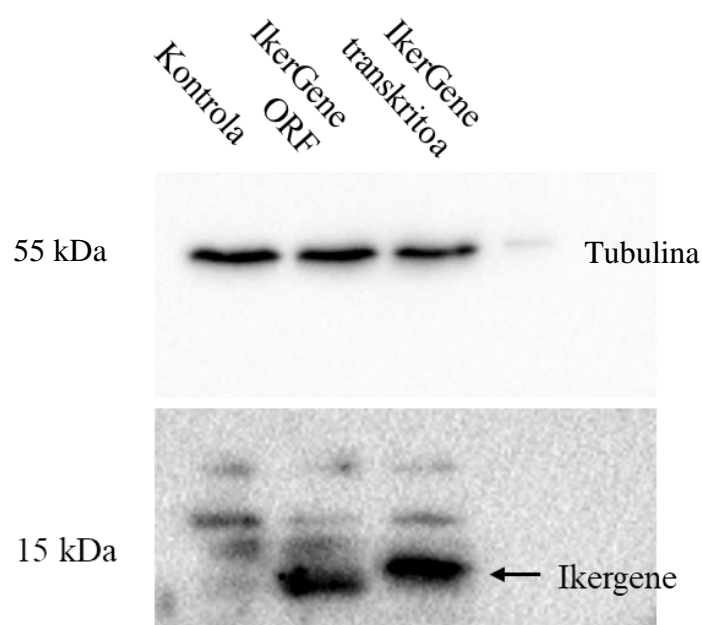
4. irudia. IKERGENEren adierazpena EndoC-βH1 zelulen RNA frakzio azpizelularretan (A). Aberasketa log(adierazpena nukleo frakzioan/adierazpena zitoplasma frakzioan) bezala kalkulatu zen; kontrol gisa RPLP0 eta lnc13 geneak erabili ziren, zeinak zitoplasman eta nukleoan aberasturik dauden hurrenez

hurren. (B) *IKERGENE*ren adierazpena zenbait giza ehun desberdinetan, erreferentzia gisa *RPLP0* erabilia



IKERGENE kodetu lezakeen ORFa gainadierazpen bektore batean klonatu genuen, 3' muturrean FLAG-tag etiketa bat gehituta. Mikropeptido potentzialaren tamaina 12 kDa ingurukoa da *Expsy Translate Tool* programaren arabera. Mikropeptidoaren presentzia *Western blot* bidez frogatu genuen gainadierazpenaren ondoren antiFLAG antigorputz poliklonal bat erabilia (5. Irudia). FLAG seinalea esperotako tokian agertu zen, 15 kDa inguruan, zentzuzkoa FLAG eta bereizlearen ~2 kDa gehiagarriak kontuan hartuta. Are gehiago, mikropeptidoaren ORF soila beharrean, RNA transkrito osoaren sekuentzia klonatzean, antzeko emaitzak lortu genituen (5. Irudia). Emaitza hauek mikropeptidoa β zeluletan benetan ekoiztu daitekeela eta egonkorra dela frogatzen dute.

5. irudia. EndoC- β H1 zeluletan hautagaiaren ORF eta luzera osoko konstruktoak gainadierazi ziren. Mikropeptidoaren ekoizpena *Western blot* bidez frogatu zen antiFLAG antigorputza erabilia



4. Ondorioak

Hemen azaldutako ikerketan, 1DMrako espezifikoki, baina gaixotasun gehiagorako garrantzitsuak izan daitezkeen mikropeptido potentzialak identifikatzeko metodo eskuragarri bat garatu dugu. Gure ustetan, *in silico* egindako predikzioen eta *in vitro* egindako esperimentu biologikoen konbinazioan oinarritzeak indarra ematen diote estrategiari. Hain zuzen, jadanik mikropeptido kodetzaile bezala sailkatu den *CRNDE* genean identifikatu ezezik, orain arte deskribatu gabeko mikropeptido bat identifikatu dugula baieztatu dezakegu.

IKERGENE hautagaiak mikropeptido bat kodetzeko ezaugarri asko batzen ditu aztertutako tresna eta baliabideen arabera: kontserbazio filogenetiko puntuala, PhyloCSF piko positiboa, ATG *Kozak* sekuentzia indartsua kontserbazio pikoaren eskualdean eta baita CPC2 algoritmoak aurrerandako potentzial kodetzailea. Gainera, kontuan hartu beharrekoa da lncRNA honen iturria erribosomei loturiko RNA sekuentziazioa dela. Honi adierazpen eta *Western blot* esperimentuetan lortutako emaitzak gehituz, ez dirudi burugabea *IKERGENE* β zeluletan mikropeptido bat kodetzeko ahalmena duela baieztatzea. Peptido honek izan ditzakeen funtzio edo eraginak aurrerago aztertu beharko dira.

Erabilitako metodo honek izan dezakeen eragozpenetako bat purifikaturiko erribosometatik erauzitako RNAREN natura da. Testuan aipaturiko *ribosome profiling* teknikan erribosomaren barruan eta, hortaz, babesturik dauden RNA puskak erauzi eta sekuentziatzen dira. Gure AHARIBON, ordea, erribosoma konplexuari atxikituriko RNA oro. Horrek bere abantailak eta berezitasunak ditu; erribosoma konplexuetan agertu daitezkeen lncRNA transkrito guztiek ez dute zertan itzultzen egon behar, baliteke itzulpen prozesuaren erregulazioan aritzea, edota beste edozein funtzio betetzea.

5. Etorkizunerako planteatzen den norabidea

Aurreko atalean iradoki bezala, *IKERGENE*ren mikropeptidoaren azterketa funtzionala burutzea litzateke hurrengo pausoa. Orain arte burututako *in vitro* esperimentuetan, mikropeptidoaren presentzia ORFa eta lncRNA transkrito osoa gainadieraziz frogatu da. Hurrengo erronka EndoC- β H1 zeluletan dagoen mikropeptido endogenoa detektatzea izango litzateke, bai egoera basalean baita PIC tratamenduaren ostean, horretarako mikropeptidoaren aurkako antigorputz pertsonalizatu bat baliatuz *Western blot* eta immunofluoreszentzia esperimentuetan. Azken etapan, mikropeptidoak parte hartzen duen bidezidorra aurkitzen saiatuko ginateke, zelulak peptidoarekin berarekin tratatuz, *IKERGENE* lncRNA isilaraziz edota CRISPR bidez genea erabat ezabatuz adibidez, aldaketa hauek zelulen fisiologian eta gene adierazpenean daukaten eraginean aztertuz. Honela hobeto ulertuko genuke *IKERGENE*k 1DM gaixotasunean zenbaterainoko garrantzia izan lezakeen.

6. Erreferentziak

- Akerman, I., Tu, Z., Beucher, A., Rolando, D. M. Y., Sauty-Colace, C., Benazra, M., Nakic, N., Yang, J., Wang, H., Pasquali, L., Moran, I., Garcia-Hurtado, J., Castro, N., Gonzalez-Franco, R., Stewart, A. F., Bonner, C., Piemonti, L., Berney, T., Groop, L., Kerr-Conte, J., ... Ferrer, J. (2017). Human Pancreatic β Cell lncRNAs Control Cell-Specific Regulatory Networks. *Cell metabolism*, 25(2), 400–411
- Anderson, D. M., Anderson, K. M., Chang, C. L., Makarewich, C. A., Nelson, B. R., McAnally, J. R., Kasaragod, P., Shelton, J. M., Liou, J., Bassel-Duby, R., & Olson, E. N. (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*, 160(4), 595–606
- Bazzini, A. A., Johnstone, T. G., Christiano, R., Mackowiak, S. D., Obermayer, B., Fleming, E. S., Vejnar, C. E., Lee, M. T., Rajewsky, N., Walther, T. C., & Giraldez, A. J. (2014). Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *The EMBO journal*, 33(9), 981–993
- Eizirik, D. L., Colli, M. L., & Ortis, F. (2009). The role of inflammation in insulinitis and beta-cell loss in type 1 diabetes. *Nature reviews. Endocrinology*, 5(4), 219–226

- Ingolia, N. T., Lareau, L. F., & Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, *147*(4), 789–802
- Li, M., Shao, F., Qian, Q., Yu, W., Zhang, Z., Chen, B., Su, D., Guo, Y., Phan, A. V., Song, L. S., Stephens, S. B., Sebag, J., Imai, Y., Yang, L., & Cao, H. (2021). A putative long noncoding RNA-encoded micropeptide maintains cellular homeostasis in pancreatic β cells. *Molecular therapy. Nucleic acids*, *26*, 307–320
- Minati, L., Firrito, C., Del Piano, A., Peretti, A., Sidoli, S., Peroni, D., Belli, R., Gandolfi, F., Romanel, A., Bernabo, P., Zasso, J., Quattrone, A., Guella, G., Lauria, F., Viero, G., & Clamer, M. (2021). One-shot analysis of translated mammalian lncRNAs with AHARIBO. *eLife*, *10*, e59303
- Motterle, A., Gattesco, S., Caille, D., Meda, P., & Regazzi, R. (2015). Involvement of long non-coding RNAs in beta cell failure at the onset of type 1 diabetes in NOD mice. *Diabetologia*, *58*(8), 1827–1835
- Op de Beeck, A., & Eizirik, D. L. (2016). Viral infections in type 1 diabetes mellitus--why the β cells?. *Nature reviews. Endocrinology*, *12*(5), 263–273.
- Richardson, S. J., Willcox, A., Bone, A. J., Foulis, A. K. & Morgan, N. G. (2009). The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. *Diabetologia*, *52*, 1143–1151
- Senís, E., Esgleas, M., Najas, S., Jiménez-Sábado, V., Bertani, C., Giménez-Alejandre, M., Escriche, A., Ruiz-Orera, J., Hergueta-Redondo, M., Jiménez, M., Giralt, A., Nuciforo, P., Albà, M. M., Peinado, H., Del Toro, D., Hove-Madsen, L., Götz, M., & Abad, M. (2021). TUNAR lncRNA Encodes a Microprotein that Regulates Neural Differentiation and Neurite Formation by Modulating Calcium Dynamics. *Frontiers in cell and developmental biology*, *9*, 747667
- Szafron, L. M., Balcerak, A., Grzybowska, E. A., Pienkowska-Grela, B., Felisiak-Golabek, A., Podgorska, A., Kulesza, M., Nowak, N., Pomorski, P., Wysocki, J., Rubel, T., Dansonka-Mieszkowska, A., Konopka, B., Lukasik, M., & Kupryjanczyk, J. (2015). The Novel Gene CRNDE Encodes a Nuclear Peptide (CRNDEP) Which Is Overexpressed in Highly Proliferating Tissues. *PLoS one*, *10*(5), e0127475
- Széll, M., Danis, J., Bata-Csörgő, Z., & Kemény, L. (2016). PRINS, a primate-specific long non-coding RNA, plays a role in the keratinocyte stress response and psoriasis pathogenesis. *Pflugers Archiv : European journal of physiology*, *468*(6), 935–943
- Thomaidou, S., Sliker, R. C., van der Slik, A. R., Boom, J., Mulder, F., Munoz-Garcia, A., 't Hart, L. M., Koeleman, B., Carlotti, F., Hoeben, R. C., Roep, B. O., Mei, H., & Zaldumbide, A. (2021). Long RNA Sequencing and Ribosome Profiling of Inflamed β -Cells Reveal an Extensive Translatome Landscape. *Diabetes*, *70*(10), 2299–2312
- Todd, J. A., Walker, N. M., Cooper, J. D., Smyth, D. J., Downes, K., Plagnol, V., Bailey, R., Nejentsev, S., Field, S. F., Payne, F., Lowe, C. E., Szeszko, J. S., Hafler, J. P., Zeitels, L., Yang, J. H., Vella, A., Nutland, S., Stevens, H. E., Schuilenburg, H., Coleman, G., ... Clayton, D. G. (2007). Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nature genetics*, *39*(7), 857–864
- Yeung, W. C., Rawlinson, W. D., & Craig, M. E. (2011). Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. *BMJ (Clinical research ed.)*, *342*, d35

7. Eskerrak eta oharrak

- Lan hau Izortze Santin doktoreak jasotako Eusko Jaurlaritzaren Osasun Sailaren (erreferentzia: 2021111001) eta Espainiako Zientzia eta Berrikuntza Ministerioaren proiektuen (erreferentzia: PID2019-104475GA-I00) bidez finantziatu da.
- Halaber, ni neu Jon Mentxaka jauna, Eusko Jaurlaritzaren Hezkuntza Sailak eskainitako doktore ez diren ikertzaileak prestatzeko Doktoratu Aurreko Programako laguntzaren onuraduna naiz.