

**A negatív szelekciót meghatározó *AIRE* gén  
allépolimorfizmusainak vizsgálata reumatoid  
arthritis kockázatára klinikai minták  
felhasználásával**

Doktori (Ph.D.) értekezés

Bérczi Bálint



Klinikai Orvostudományok Doktori Iskola  
Daganatok Molekuláris Epidemiológiája  
Doktori Iskola vezetője: Prof. Dr. Bogár Lajos,  
Programvezető: Prof. Dr. Kiss István  
Témavezető: Dr. Gyöngyi Zoltán  
Pécsi Tudományegyetem Általános Orvostudományi Kar  
Orvosi Népegészségtani Intézet

Pécs, 2024.

## Rövidítésjegyzék

aCCP	anti-ciklikus citrullinált peptid antitest
AIRE	autoimmun regulator
<i>AIRE</i>	AIRE transzkripció faktort kódoló DNS szekvencia
APECED	I-es típusú autoimmun poliendokrin szindróma
CARD	caspase-activation and recruitment domain, kaszpáz-aktivációs és recruitment domén
CRP	C-reaktív protein
DAS28	disease activity score 28, betegség aktivitási érték
ESR	erythrocyte sedimentation rate, eritrocita szedimentációs ráta
HSR	homogeneously staining region, homogéneen festődő régió
mTEC	medullar thymoepitelial cells, medulláris timoepiteliális sejtek
NLS	nuclear localisation signal, nukleáris lokalizációs jel domén
RA	reumatoid artritisz
Rf	reumatoid faktor
SAND	Sp100 AIRE NucP41/75 és DEAF domén
SNP	single nucleotide polymorphism, egy pontos nukleotid polimorfizmus
TRA	tissue restricted antigen, szöveti antigének

## **A negatív szelekciót meghatározó AIRE gén allélpolimorfizmusainak vizsgálata reumatoid artritisz kockázatára klinikai minták felhasználásával**

### **1. BEVEZETÉS**

A reumatoid artritisz (RA) egy krónikus, szinoviális gyulladással járó autoimmun betegség, melynek során a következményes extraartikuláris tünetekkel járó szimmetrikus poliartikuláris artritisz funkcióvesztéssel jár. Fejlett országokban a betegség prevalenciája felnőtt populációban 0,5-1%, éves incidenciája 5-50/100 000. Annak ellenére, hogy a betegség etiológiája még nem teljesen tisztázott, számos közlemény egyetért abban, hogy a folyamatot az adaptív immunrendszerből kijutó autoimmun T-sejtek indítják el.

A betegséggel szembeni érzékenységet számos környezeti és genetikai tényező határozza meg. A legutóbbi etnikumok közötti genomszintű asszociációs vizsgálatok 69 825 eset és 330 798 illesztett kontroll bevonásával meghatározta a betegség kialakulásában résztvevő összes olyan jelenleg detektálható gént, melyeknek egyponos nukleotid polimorfizmusai (single nucleotide polymorphism, SNP) hozzájárulhatnak a betegség etiológiájához. A gének közül egynek döntő szerepe van az autoimmunitás kontrollálásában, ez az autoimmun regulator (AIRE) transzkripciós faktort kódoló szekvencia (AIRE). A gén a 21-es kromoszóma hosszú karján, 21q22.3 sávnál helyezkedik el, ~12.5 kilobázis hosszú és 14 exoniális szekvenciával egy 545 aminosavból álló, 58 kDa nagyságú fehérjét kódol.

Az AIRE fehérje egy transzkripciós faktor, mely az éretlen T-sejtek (timociták) negatív szelekcióját irányítja. Ennek kulcsfolyamata a csecsemőmirigy medulláris sejtjeiben végbemenő promiszkuis génexpresszió, melynek során DNS-kötő fehérjékkel együttműködve olyan szöveti antigének (tissue restricted antigen, TRA) termelődését engedi meg, melyek leképezik a test teljes szöveti fehérjeállományát. Ezek az antigének kikerülnek a velősejtek felületi fő hisztokompatibilitási komplex MHC receptoraira, mely által a folyamat egyfajta tükröt tart az érésben lévő T-sejteknek.

Amennyiben az éretlen T-sejt receptorával felismeri a medulláris timoepiteliális sejtek (medullar thymopitelial cells, mTEC) felületére kihelyezett test-antigéneket, autoimmun természetűe bebizonyosodik és apoptózissal klonálisan törlődik. Multidomén szerkezete alapján az AIRE azon fehérjék csoportjába tartozik, amelyek képesek a kromatinhoz kötődni és szabályozni a géntanszkripciót. Az AIRE az N-terminustól a C-terminusig öt funkcionálisan fontos szerkezeti doménből áll: homogéne festődő régió (homogeneously staining region, HSR) vagy más néven kaszpáz-aktivációs és recruitment domén (caspase-activation and recruitment domain, CARD), nukleáris lokalizációs jel domén (nuclear localisation signal, NLS), Sp100 AIRE NucP41/75 és DEAF domén (SAND), valamint két PHD domén (PHD1, PHD2).

Az AIRE fehérje egyik legfontosabb szakasza, mely nélkül nem lenne képes a medullában funkcióját kifejteni, a SAND domén. 1998-ban ez volt az AIRE második szerkezeti szekvenciája, amelyet leírtak, a PHD domén után. Az AIRE-t kódoló aminosavlánc közepén elhelyezkedő SAND domén szerkezete egy globuláris  $\alpha/\beta$  elemet tartalmaz, melynek  $\alpha$ -hélix részében egy konzervált KDWK motívum található. Ez az elem biztosítja, hogy az AIRE fehérje a DNS foszfátcsoportjaihoz kötődjön, mely a SAND domén legfőbb kromatinfüggő transzkripciós szabályozó funkciója. Ennek segítségével kötődik az AIRE a medulláris mTEC sejtek DNS-éhez. A kötődést követően elkezdődik a klaszterbe rendezett TRA-gének

expressziója, mellyel a medulláris negatív szelekciót biztosítják az ide érkező érésben lévő T-sejtek számára. Ezt a domént az exon 5,6 és 7-es régió kódolja az *AIRE* kódoló régiójában.

Az *AIRE* fehérjét illetve a SAND régiót kódoló DNS-szekvenciában előforduló funkcióvesztő mutációk következménye a szelekció elmaradása és a perifériára kijutó érett, viszont autoreaktív T-sejtek egy viszonylag ritka autoimmun betegséget, az I-es típusú autoimmun poliendokrin szindrómát (APECED, APS-1) okozzák. Számos közlemény egybehangzóan állítja, hogy a kódoló génszakasz egy nukleotidban megmutatkozó allélpolimorfizmusai (single nucleotide polymorphism, SNP) befolyásolhatják az *AIRE* transzkripcióját, következésképp a fehérje funkcióját, amely így emelheti az egyes betegségek iránti fogékonyságot.

Az első vizsgálat, mely az allélpolimorfizmusokat összefüggésbe hozta a centrális toleranciával *Lovewell és mtsai.* publikálták. Ennek során az *AIRE* promóter régiójától 5' irányban egy 591 bp-nyi szakaszt vizsgáltak illetve lehetséges allélpolimorfizmusok *AIRE* transzkripcióra mért hatását. Humán limfocita DNS vizsgálatára a denaturáló nagy hatékonyságú folyadék kromatográfia (denaturing High-Performance Liquid Chromatography, dHPLC) módszerét alkalmazták. Ezekben a régióban két allélpolimorfizmust találtak *AIRE-655R* (rs117557896) és az *AIRE-230Y*-t (rs751032), mely régiókat pCR2.1 TA vektorba ligálták és TOP10 sejtekbe transzfektálták. Minden egyes promóter változatot a luciferáz riportergén elé, a pGL3-basic vektorba klónozták, hogy létrehozzanak egy *AIRE* riporterkonstrukciót. A transzkripció vizsgálat eredménye az lett, hogy az *AIRE-655G AIRE-230T* allélpolimorfizmusok alkotta genotípus esetében a promóter aktivitás nagyon alacsony mértékű, szinte alig kimutatható, míg *AIRE-655G AIRE-230C* genotípus a legmagasabb értéket produkálta. Az érdekessége a vizsgálatnak, hogy ezáltal kimondható, adott lokuszok polimorfizmusai alkotta genotípus képes a transzkripciót oly mértékben befolyásolni, hogy az expresszió alig kimutatható. Vagyis amennyiben ez lehetséges, akkor az is kimondható, hogy genotípusonként a centrális toleranciát meghatározó negatív szelekció újszülöttkorban egyénekenként eltérhet és adott allélpolimorfizmusok alapjaiban határozhatják meg annak hatékonyságát, alacsony *AIRE* expressziót produkáló genotípus emelheti az autoimmun megbetegedések iníciációját már a posztnatális időszaktól.

Az első *AIRE* allélpolimorfizmussal kapcsolatos molekuláris epidemiológiai vizsgálatot 2002-ben *Tazi-Ahnini és mtsai.* végezték el alopecia areata autoimmun betegségek vizsgálva. A kórlefolyás során T-sejt populáció szűródik a folliculusokba, melyek autoantitest és citokintermeléssel kopaszodást idéz elő. Ennek egy súlyos fajtájánál – alopecia universalis – azt találták, hogy a *G961C* allél frekvenciája a 202 betegnél szignifikánsan magasabb a 175 kaukázusi kontrollhoz képest. Szisztémás szklerózis (SSc) vagy más néven szkleroderma egy komplex autoimmun betegség, amelyet fibrózis, vaszkuláris rendellenességek és immunrendszeri aktiváció jellemez T-sejt infiltrációval és B-sejt aktivációval, melynek következménye autoantitest és citokin termelés. 2007-ben 41 SSc beteg és 100 kontroll vizsgálatánál *Ferrara és mtsai.* azt találták, hogy *G11107A* allélikus intron-polimorfizmus esélye szignifikánsan magasabb a betegpopulációban. Az összefüggés még erősebbnek mutatkozott, ha az SSc populáción belül autoimmun tiroditisz alcsoportban. Autoimmun bőrbetegség, a vitiligoval való összefüggést 2008-ban fedezték fel *Tazi-Ahnini és mtsai.* 86 beteg és 363 kontroll vizsgálatkor az *AIRE 7215C* allélja szignifikánsan magasabb eséllyel volt jelen a vitiligo betegpopulációjában a kapott esélyhányados (odds ratio, OR=3,12; 95%CI 1,87-5,46;  $p=1,36 \times 10^{-5}$ ) alapján. *AIRE* allélpolimorfizmus miatti nem megfelelő hatékonyságú centrális tolerancia, hatással lehet a szervezet azon képességére, hogy a tumorsejteket és azok antigénjeit felismerje és elpusztítsa. *Conteduca és mtsai.* arra voltak

kíváncsiak, hogy kapcsolatban állhatnak-e bizonyos allélpolimorfizmusok melanomával. 2010-es vizsgálatukban inverz asszociáció részeként a rs1055311 (C8385T), rs1800520 (C8723G) és a rs1800522 (T16366C) allélja szignifikánsan gyakoribbak voltak az egészségesekben, mint a melanomás betegeknél, függetlenül a nemtől, az életkortól és a melanoma stádiumától, sugallva az allél esetleges protektív hatását. Autoimmun alapú súlyos izomgyengeséget, myasthenia gravist 2017-ben hozták összefüggésbe *Zhang és mtsai.* egy kínai eset-kontroll kapcsán, melynél az rs3761389 G-allélja illetve az erre az allélra homozigóta genotípus szignifikánsan magasabb eséllyel volt jelen a betegeknél a kontrollhoz képest (OR=1,68; 95%CI 1,14–2,48; p=0,027). A mellékvesekéreg autoimmun pusztulásával, az Addison-kórral 2018-ban sikerült inverz asszociációt találni. Az rs2075875 C-, rs2075876 A-, rs62220374 G-, rs9983695 C-allélja szignifikánsan magasabb eséllyel volt jelen a kontrollpopulációban, ezek az allélok protektívnek bizonyultak a betegséggel kapcsolatban. Szisztémás lupusz eritematózus (SLE) esetében az immunrendszer tévesen megtámadja több szerv szöveteit az egész testben, különösen a bőrt, a veséket, az agyat és az ízületeket. Ezt a ritka autoimmun betegséget 2021-ben hozták összefüggésbe a timociták negatív szelekcióját irányító, AIRE fehérje expresszióját meghatározó intron 5-ös régióban található rs2075876 A alléljával, mely protektívnek bizonyult az esetcsoportban (OR=0,16, 95%CI=0,09–0,28, p<0,001). Végül az AIRE allélpolimorfizmusok terén a legutóbbi molekuláris epidemiológiai eredményt 2023-ban publikálták és találtak összefüggést trombocitopéniás purpura betegséggel, mely autoimmun folyamatok által okozott, autoreaktív trombocita pusztulás következtében kialakuló vérzékenység. Ebben a rendellenességben szintén az rs2075876 lókus intron 5-ös régiójában található A allélja és az arra homozigóta genotípus jelenti a kockázatot (OR=4,299; 95%CI 1,650–11,202; p=0.003). Ahogy látjuk, ezen autoimmun betegségek esetében tömegével nem készítették asszociációs vizsgálatokat. Szakirodalom szerint viszont egy autoimmun betegség összefüggését kiterjedten vizsgálták az AIRE kódoló régió allélpolimorfizmusaiival, ez pedig a reumatoid arthritisz.

Ez a betegség egy gyakori autoimmun manifesztáció, amely krónikus ízületi gyulladással jár. Az ebből eredő szimmetrikus, poliartikuláris arthritisz extraartikuláris szövődeményekkel kombinálva funkcionális károsodáshoz vezet. Az AIRE promóter régiójában lévő polimorfizmusokat 2011-ben genom szintű asszociációs vizsgálat keretében először *Terao és mtsai* vizsgálták. Ennek során 5415 RA és 6489 ázsiai kontroll körében a 277 420 SNP vizsgálatkor azt találták, hogy az intron régióban lévő rs2075876 A- és rs760426 G-allélja és ezen allélok alkotta homozigóta genotípusok szignifikánsan magasabb eséllyel van jelen a betegpopulációban (OR=1,18; 95%CI 1,11–1,24; p=3,6×10<sup>-9</sup>; OR=1,16; 95%CI 1,10–1,22; p=4,4×10<sup>-8</sup>). További eredményeikben publikálták, hogy dacára annak, hogy az rs2075876 splicing régióban foglal helyet, az AA genotípus jelenlétében az expresszió szignifikánsan alacsonyabb (p=6,8×10<sup>-5</sup>) szemben a GG genotípussal, mely eredmény felveti annak a lehetőségét, hogy a betegségre való kitétség már újszülöttkorban elkezdődik, a nem megfelelően működő negatív szelekció által. Hasonló eredményt értek el spanyol, japán, kínai, egyiptomi vizsgálatok során melyek döntő többsége eset-kontroll vizsgálat. Összességében elmondható, hogy maguk az allélpolimorfizmusok alkotta genotípusok nagymértékben hozzájárulhatnak az AIRE expresszió alakulásához, mint azt *Lovewell és mtsai.*-nak vizsgálatában láttuk, képesek alacsony expressziót eredményezni, mely az immunológiai tolerancia esetében az újszülöttkori negatív szelekciót alapjaiban határozhatja meg. Alacsony AIRE expresszió esetén, magasabb valószínűséggel kerülhetnek ki autoreaktív CD4<sup>+</sup> effektor sejtek a perifériára, melyek autoantigén jelenlétében elindíthatják a betegség iníciációs szakaszát már újszülöttkorban.

Az említett betegségek közül egyedül a reumatoid arthritiszt vizsgálta több molekuláris epidemiológiai megfigyelés, így számunkra optimálisnak bizonyult allélpolimorfizmusokkal

való pozitív vagy negatív összefüggések további elemzésére valamint a szakirodalom alapján a SAND régió egy kifejezetten fontos doménnek számít a negatív szelekció folyamatában, így vizsgálódásunk tárgyát azok a DNS szakaszok jelentik, melyek ezt a domént kódolják.

## **2. CÉLKITŰZÉSEK**

### **2.1. Az AIRE allépolimorfizmusok meta-analízise autoimmun kórképekben**

Az első AIRE allépolimorfizmussal kapcsolatos molekuláris epidemiológiai megfigyelést 2002-ben végezték, azóta számos eset-kontroll analízis született különböző autoimmun betegségeket vizsgálva. Mielőtt elkezdtük kutatásunkat, szisztematikus szakirodalmi áttekintéssel vizsgálódtunk, mely meta-analízisnek a célja az volt, hogy tisztázzuk, mely autoimmun betegség körében született a legtöbb molekuláris epidemiológiai eset-kontroll vizsgálat, a betegségen belül pedig, az egyes eset-kontroll vizsgálatok összefoglaló elemzésével új eredményeket és összefüggéseket találunk kandidáns allépolimorfizmusok és autoimmun betegségek között.

### **2.2. In silico vizsgálat AIRE allépolimorfizmusok meghatározására**

Annak érdekében, hogy rendelkezésünkre álljon egy adatbázis, mely lókuszkok allépolimorfizmusaival kívánunk a későbbiekben foglalkozni, célul tűztük ki egy olyan átfogó adatbázis létrehozását, melyben 5'-3' irányban szerepel az összes olyan allépolimorfizmus, melynek allélfrekvenciája legalább 1% vagy afölötti.

### **2.3. Pilot vizsgálat AIRE rs878081 és rs1003854 allépolimorfizmusok vizsgálatára reumatoid artritisz betegek mintáiban**

Mivel számunkra kezdetben ez egy ismeretlen terület volt, meg kellett figyelniünk, vajon a kiválasztott AIRE domén-kódoló régió lókuszaik adta allélfrekvenciák és vizsgálati elrendezésünk képesek-e értékelhető eredményt adni. Annak tisztázására, hogy az allépolimorfizmusok diszkriminációjára választott allélspecifikus próbák és a genotípezálás során kialakított protokoll megfelelőek-e, célul tűztünk ki egy olyan pilot vizsgálat megvalósítását, ahol ezekre a kérdésekre választ kaphatunk. Ennek érdekében kiválasztottuk a DNS-kötő domént kódoló SAND régió két lókusztát, rs878081-et, rs1003854-et és célul tűztük ki ezen lókuszkok adta allépolimorfizmusok és genotípusok összefüggésének vizsgálatát reumatoid artritisz kockázatával, eset-kontroll vizsgálat keretében, melyhez 6 különböző genetikai modellt használtunk. Azért ezt a régiót céloztuk meg, mert véleményünk szerint ezen exon szakasz által kódolt SAND domén a legfontosabb eleme az AIRE fehérjének az mTEC DNS-hez való kötődésnek, így a negatív szelekciónak. Ezen belül az rs878081-es lókuszon található allépolimorfizmus egy szinonim variáns, mely nagymértékben befolyásolhatja a termelődő fehérje mennyiségét, így a negatív szelekció hatékonyságát. Vizsgálatunkhoz célul tűztük ki egy olyan átfogó adatbázis létrehozását, mely magába foglalja a pilot study és az esetleges későbbi vizsgálat RA és kontrollpopulációinak teljes leletanyagát, rögzítve minden olyan paramétert, mely a leletanyagból kinyerhető, beleértve demográfiai adatokat, életmódi adatokat, laborvizsgálatok eredményeit és a múltban rögzített fertőzéseket is. További célul tűztük ki, amennyiben lehetséges az egyes genotípus-alcsoportok klinikai paramétereinek átlagértékeit statisztikailag összehasonlítsuk és amennyiben az szignifikánsnak bizonyul korrelációs vagy asszociációs összefüggéseket vizsgáljunk a klinikai paraméterek és a genotípusok között.

## **2.4. Az AIRE exon 5 és 6 valamint intron 5 és 7-es régióinak vizsgálata reumatoid artritisz kockázatára**

Annak függvényében, a pilot vizsgálatunk milyen eredménnyel jár, a vizsgálat folytatásaként célul tűztük ki nagyobb elemszámú eset és kontrollpopuláció és több allélpolimorfizmus vizsgálatát annak tisztázására, vajon a DNS-kötő régió allélpolimorfizmusai kapcsolatban állnak-e a betegséggel. Ennek érdekében az *in silico* adatbázist és a pilot study eredményeit felhasználva, egy olyan panelt terveztünk, mely magába foglalja mindazon potenciális lókuszokat és azok adta allélpolimorfizmusokat, melyek SAND domént kódoló régiót (exon 5 és 6) valamint a köztük lévő splicing intronális régiót (intron 5 és 7) határozzák meg. Célunk ezen panelben szereplő allélpolimorfizmusok illetve az allélok adta genotípusok vizsgálata reumatoid artritisz kockázatára, 6 különböző genetikai modell felhasználásával.

## **2.5. Az AIRE exon 5 és 6 valamint intron 5 és 7-es régiók összefüggéseinek vizsgálata a betegséget meghatározó klinikai paraméterekkel**

A továbbiakban kíváncsiak voltunk, hogy az egyes allélpolimorfizmusok adta genotípusok összefüggésben állnak-e az RA-t meghatározó klinikai paraméterekkel és a betegség aktivitásával. Célul tűztük ki a panelben szereplő lókuszok alléljai által meghatározott genotípusok vizsgálatát, 5 különböző genetikai modellt felhasználva a következő klinikai paraméterek vizsgálatára: vérsüllyedést meghatározó eritrocita szedimentációs ráta (erythrocyte sedimentation rate, ESR), C-reaktív protein (CRP), reumatoid faktor (Rf), anti-ciklikus citrullinált peptid antitest (aCCP), betegség aktivitási (disease activity score 28, DAS28) érték.

## **2.6. Allélpolimorfizmusok regulatórikus kötőmotívumokra gyakorolt hatásainak *in silico* elemzése**

A vizsgálat zárásaként arra voltunk kíváncsiak, az egyes lókuszok alléljai milyen hatással lehetnek a DNS azon kötőhelyeire, melyekhez regulatórikus egységek, transzkripció faktorok (TF) kötődhetnek, így célul tűztük ki a panelon szereplő azon allélpolimorfizmusok bevonását, melyekkel kapcsolatban statisztikailag szignifikáns eredményt kaptunk az alléldiszkriminációs vizsgálat során RA-val illetve ha ott nem, de adott allélok adta genotípus statisztikailag szignifikáns összefüggésben volt klinikai paraméterekkel. A kötőmotívumok és az azokhoz kapcsolódó regulatórikus egységre mért hatást célul tűztük ki a HaploReg v4.2 *in silico* adatbázisban elemezni.

## **3. ANYAG ÉS MÓDSZER**

### **3.1. AIRE allélpolimorfizmusok meta-analízise autoimmun kórképekben**

A szakirodalmat adatbázisokban kulcsszavas kereséssel kerestük. A közleményeket beválasztási és kizárási kritériumok alapján válogattuk. Statisztikai vizsgálat során öt különböző genetikai modellt használtunk fel, valamint Hardy-Weinberg egyensúlyt, heterogenitást, szenzitivitás-analízist és publikációs torzítást számoltunk.

### **3.2. *In silico* vizsgálat AIRE allélpolimorfizmusok meghatározására**

Adatbázisunkba a variáns akkor került beválogatásra, amennyiben a ritka allél frekvenciája 0,01 (1%) és 0,5 (50%) érték közé esett illetve referencia, annotációs adatbázisban szerepel. A

variáns kizárásra került ha a ritka allél frekvenciája 0,01 alá esett. Ehhez két genom-szintű annotációs rendszert, UCSC Genome Browser (GRCh38 assembly) valamint az Ensembl automatikus humán annotációs rendszert használtuk.

### **3.3. Pilot vizsgálat AIRE rs878081 és rs1003854 allélpolimorfizmusok vizsgálatára reumatoid arthritisz betegek mintáiban**

Vizsgálatunk beavatkozással járó kutatás, így ennek engedélyeztetésére etikai engedélyre irányuló kérelmet nyújtottunk be az Országos Tisztifőorvosi Feladatokért Felelős Helyettes Államtitkársága, Egészségügyi Igazgatási Főosztályára (továbbiakban: országos tisztifőorvos) 2018-ban, négy éves intervallumra, a Harkányi Termál Rehabilitációs Centrum Közhasznú Nonprofit Kft. bevonásával. Ehhez az országos tisztifőorvos felkérte az Egészségügyi Tudományos Tanács és Kutatásetikai Bizottságát (ETT TUKEB) szakhatósági állásfoglalás kialakítására. Ennek során az ETT TUKEB a kutatás engedélyezése iránti kérelmet szakmai-etikai szempontból megfelelőnek találta, a kutatás országos tisztifőorvos általi engedélyezéséhez szakhatósági hozzájárulását adta. Ennek alapján az országos tisztifőorvos 2018. április 24-én kérelmünket 4 évre hagyta jóvá, 11871-7/2018/EÜIG iktatószámmal.

A vizsgálatban szereplő valamennyi személy a Harkányi Termál Rehabilitációs Központ ambuláns és bennfekvéses betegek közül kerültek ki. A vizsgálatra jelentkezés önkéntes alapon történt, melyhez betegtájékoztatót és beleegyező nyilatkozatot csatoltunk. A vérmintákat 4mL-es citrátsövekbe vettük személyenként. A teljes vért  $-75^{\circ}\text{C}$ -on tároltuk az analízisig. Az önkéntesen vállalkozó személy a betegcsoportba került, amennyiben (1) 18 évnél idősebb volt, (2) a beleegyező nyilatkozatot aláírta, valamint (3) az American College of Rheumatology diagnosztikus és besorolási kritériumai alapján reumatoid arthritisz diagnózisával rendelkezik. Azokat a betegeket, akiknek a kórtörténetében egyidejűleg más autoimmun kórképek, pl. Sjögren-szindróma, psoriasis, SSc, SLE szerepeltek, kizártuk a további elemzésből, annak érdekében, hogy csökkentjük az RA-tól eltérő autoimmun betegségek RA kockázatára mért zavaró hatását.

Mivel már a vizsgálat elején tudtuk, hogy gondosan meg kell terveznünk a megfelelő kontrollpopulációt, azon gondolkodtunk, hogyan korlátozhatnánk az esetleges fals negatív esetek előfordulási gyakoriságát az RA kimenetelére vonatkozóan. A betegség lefolyásáról az a tapasztalatunk, hogy az első tünetek és a kórházi kezelés átlagéletkora 65 év körüli. A szakirodalomból és a betegek orvosi jelentésének áttekintése során azonban azt láttuk, hogy az esetek kisebb részében az első kórházi kezelés 70 éves korban is jelentkezhet. Ezért úgy döntöttünk, hogy a kontrollok életkori felvételi kritériumát 75 év vagy annál idősebbre emeljük, mert ennek hiányában esetleg olyan alanyok kerültek volna a kontrollcsoportunkba, ahol az RA a preklinikai stádiumban van, és ahol a betegség patomechanizmusa az iníciációs vagy centrális stádiumban vannak, tünetek jelenléte nélkül. 75 éves korukra a betegek nagyobb valószínűséggel jelentkeznek tünetekkel, és egyszer már az ellátórendszer részévé válnak. Így csökkentettük a zavaró hatásokat a vizsgálatunkban. Kontrollszemélyekre vonatkozó felvételi kritériumaink tehát a következők voltak: (1) 75 éves vagy annál idősebbnek kellett lennie, (2) aláírta a beleegyező nyilatkozatot, (3) az American College of Rheumatology diagnosztikus és besorolási kritériumai alapján nem szenvedett RA-ban, (4) leletanyaga alapján nem szenvedett egyéb autoimmun betegségben (pl. SLE, SSc, psoriasis, Sjögren-szindróma), (5) CRP-értéke  $\leq 20$  mg/dl és Rf szintje  $\leq 25$  IU/ml (153). Végül pilot vizsgálatunk során 100 RA és 100 kontrollszemély vérmintáit sikerült összegyűjtenünk.



Az RA és kontroll vérmintákból elsődlegesen teljes vérképet, koleszterinszintet, vércukorszintet, ESR-értéket, CRP-értéket, nátrium-, kálium-, karbamid-, kreatinin-, glutamát-oxalacetát-transzamináz-szintet (GOT), glutamát-piruvát-transzamináz-szintet (GPT), a gamma-glutamiltranszferáz-szintet (gamma-GT), alkalikus foszfatáz, laktát-dehidrogenáz, kreatin-kináz, fehérjealbumin, magnézium, foszfor és a szérum bilirubin-értéket határoztunk meg. Az RA szeropozitivitásának és –negativitásának meghatározására Rf (IgG) és aCCP autoantitest szinteket határoztunk meg mind az RA mind pedig a kontrollcsoportban, az utóbbiban legfőképp azért, hogy az autoimmun álnegatívok ezúton is kizárásra kerüljenek.

A genomiális DNS extrakcióját és tisztítását követően multiplex qPCR vizsgálatot végeztünk a SAND régióban található rs878081 és rs1003854 allélpolimorfizmusok genotipizálására, melyet mind az RA betegek és kontroll személyek esetében TaqMan technológiával végeztünk. A pontos allélikus diszkrimináció érdekében Minor Groove Binder (MGB) TaqMan® Assay-eket (Thermo Fisher Scientific, Waltham, USA) használtunk, ahol az MGB az oligonukleotid próba 3' végében megnöveli az olvadási hőmérsékletet, így stabilizálva a próba-target hibrideket.

Statisztikai elemzéseinket az SPSS szoftver 28.0 verziójával végeztük (IBM Corp. Released 2021, IBM SPSS Statistics for Windows, Version 28.0; IBM Corp, Armonk, NY, USA). A várható genotípus-gyakoriságokat a Hardy-Weinberg-egyensúly (HWE) alapján számoltuk ki, és a megfigyelt gyakoriságokkal való különbséget statisztikailag összehasonlítottuk nem-paraméteres  $\chi^2$  próbával, hogy lássuk, mennyire illeszkednek a megfigyelt genotípus-gyakoriságok a HWE szerinti várható gyakoriságokhoz. A klinikai paramétereket jelentő folytonos változókat átlagérték  $\pm$  standard deviáció (SD) formában ábrázoltuk. A normalitás meghatározására Shapiro-Wilk-próbát, a Levene-tesztet pedig a varianciák egyezésének elemzésére használtuk. A csoportok összehasonlítására a folytonos változók esetében független mintás t-próbát és varianciaanalízist alkalmaztunk. A nem normáloszlású változók mediánjainak összehasonlítására Mann-Whitney U és Kruskal-Wallis H próbákat alkalmaztunk kétszélű szignifikanciával. A korrelációt kétváltozós korrelációval, Pearson korrelációs együtthatóval és kétszélű szignifikanciával vizsgáltuk. Bináris logisztikus regressziót használtunk OR-értékek kiszámítására 95% CI intervallumot alkalmazva. Statisztikai szignifikanciát akkor határoztuk meg, ha a p-értékek kisebbek voltak 0,05-nél. Több okból is a hagyományos 0,05-ös p-értéket választottuk. Először is, ez egy szabványosított küszöbérték a tudományos kutatásban. Ezenkívül, tanulmányunk biológiai alapú kutatás, amelynek célja a feltételezett expozíció és a kimenetel közötti új összefüggések kimutatása, ezért el akartuk kerülni ennek a standardizált határértéknek a korrigálását vagy módosítását. Ebben az esetben a 0,05 választása ésszerű, mert elfogadható egyensúlyt teremt az első- és másodfajú hibák között. Ez azt jelenti, hogy az érték ésszerű egyensúlyt hoz létre a túl megengedő – ami a hamis pozitív eredmények kockázatának emelkedését okozza – és a túlságosan szigorú értékelés között, ahol a hamis negatív eredmények kockázata emelkedett volna. Bár a szignifikanciát 0,05-re állítottuk be, a táblázatokban magasabb szignifikanciák (<0,01, <0,001) is megtalálhatók.

### **3.4. Az AIRE exon 5 és 6 valamint intron 5 és 7-es régióinak vizsgálata reumatoid artritisz kockázatára**

Annak érdekében, hogy vizsgálatunkat folytatni tudjuk szükségünk volt további minták gyűjtésére és időre, melyhez elengedhetetlen volt a pilot vizsgálathoz kapott etikai engedély meghosszabbítása. Ennek során mintagyűjtésünk helyszínét a meglévő Harkányi Termál

Rehabilitációs Centrum Közhasznú Nonprofit Kft. mellett a Hévízgyógyfürdő és Szent András Reumakórházzal együttműködésben kibővítettük és kérelmeztünk az etikai engedély meghosszabbítását további négy évre. Benyújtott módosítási kérelmünket a szakhatósági véleményezésre felkért ETT TUKEB szakmai és etikai szempontból megfelelőnek találta és a kutatási terv kérelemben foglalt módosításának engedélyezéséhez hozzájárult, melyek alapján a Nemzeti Népegészségügyi Központ a módosítási kérelmünket az 57142-5/2022/EÜIG iktatószámú határozatban, további négy évre 2022. november 8-án engedélyezte.

A mintagyűjtés inentől kezdve párhuzamosan két intézményben zajlott. A Harkányi Termál Rehabilitációs Centrumban történt RA és kontroll személyek vérmintáinak gyűjtése, míg a Hévízgyógyfürdő és Szent András Reumakórházzal együttműködve kizárólag RA betegek vérmintáinak gyűjtése zajlott. A vizsgálatra jelentkezés jelen esetben is önkéntes alapon történt, melyhez beleegyező nyilatkozatot és betegtájékoztatót csatoltunk. A vérmintákat személyenként 4mL-es citrátcsövekbe vettük, melyet intézetünkben  $-75^{\circ}\text{C}$ -on tároltuk az analízisig. A vizsgálatra önkéntesen vállalkozó személy RA csoportba került, amennyiben (1) 18 évnél idősebb volt, (2) aláírta a beleegyező nyilatkozatot, valamint (3) az American College of Rheumatology diagnosztikus és besorolási kritériumai alapján az RA-t diagnosztizáltak nála. Azon RA betegeket, akiknek a leletanyagában fellelhető volt egyidejűleg más autoimmun kórkép is, mint pl. Sjögren-szindróma, pszoriázis, SSc, SLE, kizártuk a további elemzésből, annak érdekében, hogy kizárjuk az RA-tól eltérő autoimmun betegségek a kiválasztott allélpolimorfizmusokkal való lehetséges összefüggését, valamint hogy csökkentjük az RA-tól eltérő autoimmun betegségek RA kockázatára mért zavaró hatását. Jelen vizsgálatunkban a pilot studyban alkalmazott kontroll csoport életkori kritériumát használtuk a fals RA negatív esetek elkerülésére, így a kontrollok életkori felvételi kritériumát 75 év vagy annál idősebbre emeltük. Ennek hiányában esetleg olyan alanyok kerültek volna a kontrollcsoportunkba, akik az RA preklinikai, tünetmentes stádiumában vannak. Az önkéntesen jelentkező kontrollokra vonatkozó felvételi kritériumaink ennek megfelelően a következők voltak: (1) 75 éves vagy annál idősebbnek kellett lennie, (2) aláírta a beleegyező nyilatkozatot, (3) az American College of Rheumatology diagnosztikus és besorolási kritériumai alapján nem szenvedett RA-ban, (4) leletanyaga alapján nem szenvedett egyéb autoimmun betegségben (pl. SLE, SSc, psoriasis, Sjögren-szindróma), (5) CRP-értéke  $\leq 20$  mg/dl és Rf szintje  $\leq 25$  IU/ml (153). A teljes mintagyűjtés végére, 2023. decemberére összesen 592 jelentkező vérmintáit sikerült összegyűjteni, melyből 270 RA beteg és 322 kontroll személy volt.

Az 592 személy vérmintáját elsődlegesen a következő klinikai paraméterek meghatározására használtuk: teljes vérkép, koleszterin, vércukor, ESR, CRP, nátrium, kálium, karbamid, kreatinin, GOT, GPT, gamma-GT, alkalikus foszfatáz, laktát-dehidrogenáz, kreatin-kináz, fehérjealbumin, magnézium, foszfor és a szérum bilirubin. Szérum Rf (IgG) és aCCP meghatározás annak érdekében történt, hogy erősítsük RA-csoportunkat, valamint a kontrollból az álnegatívakat kizárjuk.

Reakcióként 20 ng gDNS-t használtunk multiplex qPCR vizsgálatunkhoz. Ebben a kiterjedtebb második vizsgálatban az allélpolimorfizmusok genotipizálására szintén TaqMan technológiával alkalmaztunk. A Minor Groove Binder (MGB) TaqMan® Assay-eket (Thermo Fisher Scientific, Waltham, USA) alkalmazva stabilabb próba-target hibrideket sikerült létrehozni az allélikus diszkrimináció során. A pilot vizsgálatunk során sikeresen alkalmaztuk a genotipizáláshoz szükséges termékeket, így a jelen esetben is azt a vizsgálati elrendezést választottuk. TaqMan® Assay-ek tartalmaznak egy primerpárt, egy VIC festékkel jelölt oligonukleotid próbát (530 nm-es sárga csatorna), amely az 1-es allél szekvenciát

detektálja, és egy FAM festékkel jelölt próbát (470 nm-es zöld csatorna), amely a 2-es allél szekvenciáját detektálja. Homozigóta a genotípus, ha kizárólag a VIC vagy a FAM csatornán mutatkozik fluoreszcens jel. Heterozigóta a genotípus amennyiben mindkét csatornán mutatkozik jel.

Statisztikai elemzéseinket a vizsgálat során SPSS szoftver 28.0 verziójával végeztük (IBM Corp. Released 2021, IBM SPSS Statistics for Windows, Version 28.0; IBM Corp, Armonk, NY, USA). A HWE alapján várható és megfigyelt genotípus frekvenciákat nem-paraméteres  $\chi^2$  próbával vizsgáltuk, hogy lássuk van-e szignifikáns eltérés a kettő között. A kapcsoltságot a Lewontin-féle „D” módszerrel értékeltük, ahol az LD-struktúrát magába foglaló haplotípus blokkot a Haploview szoftver (4.2. verzió) segítségével generáltuk.

A klinikai folytonos változókat átlagérték  $\pm$  SD formában ábrázoltuk. A normalitás meghatározására Shapiro-Wilk-próbát, a Levene-tesztet pedig a varianciák egyezésének elemzésére használtuk. A csoportok összehasonlítására a folytonos változók esetében független mintás t-próbát és varianciaanalízist alkalmaztunk. Mann-Whitney U és Kruskal-Wallis H próbákat alkalmaztunk kétszélű szignifikanciával a nem normáloszlású változók mediánjainak összehasonlítására. A korrelációt kétváltozós korrelációval, Pearson korrelációs együtthatóval és kétszélű szignifikanciával vizsgáltuk. OR-értékek kiszámítására bináris logisztikus regressziót alkalmaztunk, melynek során 95% CI és p-értéket generáltunk. Statisztikai szignifikanciát a 0,05 értéknél kisebb p-érték jelentette.

### **3.5. AIRE allélpolimorfizmusok regulatórikus kötőmotívumokra gyakorolt hatásainak *in silico* elemzése**

Az egyes allélpolimorfizmusok jelenlétének hatását a transzkripciós faktorok kötődési affinitására HaploReg v4.2 adatbázis segítségével *in silico* elemeztük. Az adatbázis segítségével potenciális transzkripciós faktorok kötődési affinitási értékeiket számoltuk ki a ritka allél és a gyakori allél esetében.

## **4. EREDMÉNYEK**

### **4.1. Az AIRE allélpolimorfizmusok meta-analízise autoimmun kórképekben**

Keresés során négy irodalmi adatbázisból, a duplikátumok eltávolítását követően 11 releváns közleményt azonosítottunk, melyek közül öt foglalkozott reumatoid artritisszel míg a maradék publikációk 1-1 betegséggel. Így meta-analízisünket ennek vizsgálatára irányítottuk. Az öt beválasztott eset-kontroll analízisek vizsgálata során 11 SNP-t azonosítottunk, melyek közül az rs2075876 és rs760426 allélpolimorfizmusokat vizsgálta az eset-kontroll analízisek többsége, 7145 eset, 8579 kontroll valamint 6696 eset, 8164 kontroll fővel. Ezen SNP-k vizsgálatára kiválasztott öt genetikai modell 50 súlyozott esélyhányadosa 9 kivétellel statisztikailag szignifikánsan azt mutatta, hogy az rs2075876 A- és a rs760426 G-allélja ázsiaiak körében kockázati tényező a betegség kialakulására.

### **4.2. *In silico* vizsgálat AIRE allélpolimorfizmusok meghatározására**

UCSC Genome Browser és az Ensembl automatikus annotációs rendszerében a keresésünket az AIRE kódoló szekvenciájára (HGNC Symbol; Acc: HGNC:360) állítottuk be. Korlátozások nélkül 23 023 variációt azonosítottunk, melyek között egyaránt szerepelt pontmutáció, indel és SNP is. A ritka allél globális előfordulásai gyakorisága alapján 68 SNP-t detektáltunk, melyeket egyenként európai populációban is megvizsgálva végül 44 olyan SNP-t válogattunk

be adatbázisunkba, melyeknél a ritka allél frekvenciája 0,01 (1%) és 0,5 (50%) érték közé esett illetve referencia, annotációs adatbázisban szerepelnek. Ezek közül ezt követően további vizsgálatainkra a negatív szelekcióban betöltött szerepe alapján a SAND régiót kódoló DNS szakasz allélpolimorfizmusait választottuk.

#### **4.3. Pilot vizsgálat AIRE rs878081 és rs1003854 allélpolimorfizmusok vizsgálatára reumatoid artritisz betegek mintáiban**

Pilot vizsgálatunkhoz a Harkányi Termál Rehabilitációs Központ segítségével 100 esetet és 100 kontroll mintát gyűjtöttünk össze, melyeknél a betegek szignifikánsan magasabb reumatoid artritisz marker értéket mutattak.

Allélikus modellünk keretében a kockázatot jelentő allélt keressük a SAND régiót kódoló szakaszon lévő rs878081 és rs1003854 lókuszt esetében, vagyis először megvizsgáltuk, mely allélok frekvenciája magasabb a betegpopulációban, majd bináris logisztikus regressziós vizsgálattal esélyhányados értéket számoltunk, hogy megtudjuk, a betegségkockázat mely allél irányába mutat. Ezt követően a betegségkockázatot jelentő allélra alakítottuk ki domináns, recesszív, kodomináns heterozigóta és kodomináns homozigóta és over-domináns populációgenetikai modelljeinket mind az rs878081 és rs1003854 lókuszt esetében. A kontrollcsoportok mindegyike Hardy–Weinberg egyensúlyban volt.

A recesszív genetikai modell alapján statisztikailag szignifikáns összefüggést találtunk az RA-val, ahol a TT homozigóták gyakorisága szignifikánsan magasabb volt az RA-s betegekénél, mint a kontroll személyeknél, ami alátámasztja az RA kockázatát (OR=5,44, 95% CI=1,16-25,52,  $p=0,032$ ). A TT homozigóták genotípus-gyakorisága szintén magasabb volt az RA-s betegek körében a kodomináns homozigóta genetikai modell alapján, mint a kontroll személyek körében; az összefüggés azonban nem volt szignifikáns (OR=4,68, 95% CI=0,98-22,27,  $p=0,052$ ), de ugyanazon irányba mutatott, mint a recesszív modell (152). Az allélikus, domináns, kodomináns heterozigóta vagy over-domináns genetikai modellekben nem találtunk további szignifikáns összefüggést.

Az rs1003854 allélpolimorfizmusait allélikus, domináns, recesszív, kodomináns heterozigóta, kodomináns homozigóta és over-domináns genetikai modellek alapján vizsgáltuk. A genotípus gyakorisága az RA-s betegekénél statisztikailag szignifikánsan magasabb volt a TT homozigótáknál, mint a kontroll személyeknél, ami alátámasztja az RA kockázatát (OR=4,84, 95% CI=1,02-23,02,  $p=0,047$ ). Az allélikus, domináns, kodomináns heterozigóta vagy over-domináns genetikai modellekben nem találtunk további szignifikáns összefüggést.

RA betegekénél a CC genotípus-alcsoportban szignifikánsan magasabb volt az ESR szint a TT homozigótákhoz képest az rs1003854 kodomináns homozigóta modellje alapján ( $p=0,048$ ). A CC genotípus-alcsoport szignifikánsan magasabb CRP-szintet mutatott a TT homozigótákhoz képest a kodomináns homozigóta modellben valamint a TT+TC alcsoportéhoz képest a recesszív modellben ( $p=0,043$  és  $p=0,029$ ). Továbbá, a CC homozigóta RA-s betegek esetében szignifikáns korreláció mutatkozott a CRP-vel a kodomináns homozigóta és a recesszív modellben ( $p=0,006$  és  $p=0,002$ , illetve  $r=0,350$  és  $r=0,338$ ).

#### **4.4. Az AIRE exon 5 és 6 valamint intron 5 és 7-es régiókban található allélpolimorfizmusok vizsgálata reumatoid artritisz kockázatára**

A pilotban vizsgált két lókuszt mellé, választottunk három olyan lókuszt, melyek nagymértékben hozzájárulnak a SAND domén expressziójához. A vizsgálathoz kiválasztott lókusztok 5'-3' irányban így az rs878081 az exon 5-ös, rs1003853 és rs2075876 az intron 5-ös, rs1055311 az exon 6-os és rs1003854 az intron 7-es régióban.

Ebben a vizsgálatunkban két csoportot alakítottunk ki, melynek megfelelően az 1. csoport a teljes 592 fős populációt jelenti 270 esettel és 322 kontrollal, melyekhez választottuk a három további *AIRE* rs1003853, rs2075876 és rs1055311 lókuszt. A 2. csoport a teljes populációnak egy alcsoportja 170 RA és 222 kontrollal, melyekhez az *AIRE* rs878081 és rs1003854-es lókusztait választottuk.

Jelen esetben is először az allélikus genetikai modellt használtuk annak érdekében, hogy megtaláljuk mely allélok előfordulási gyakorisága magasabb a betegpopulációban, majd bináris logisztikus regressziós vizsgálattal OR értéket számoltunk, hogy megtudjuk, a betegségek kockázat melyik irányába mutat. Ezt követően a betegségek kockázatát jelentő allélra alakítottuk ki domináns, recesszív, kodomináns heterozigóta és kodomináns homozigóta és over-domináns populációgenetikai modelljeinket. Ezt a metodikát elvégeztük mind az öt lókuszt, az rs878081, rs1003853, rs2075876, rs1055311 és rs1003854 esetében, az 1. és 2. csoport alkotta teljes, 592 fős vizsgálati populáció esetében, melyhez 1184 allél diszkriminációs teszteredményeit használtuk fel.

Vizsgálatunk során az 1. csoportot jelentő teljes, 592 fős populáció 1184 allélját vizsgálva határoztunk meg szignifikáns összefüggést. Az *AIRE* intron 5-ös régiójában található rs1003853 splice variáns C-alléljának frekvenciája szignifikánsan magasabb volt RA betegek körében, mint a kontroll személyeknél, bináris logisztikus regresszió eredménye alapján elmondható, hogy az *AIRE* rs1003853 C-alléla és az RA betegség között statisztikailag szignifikáns összefüggés van (OR=1,33, 95% CI 1,01-1,74, p=0,037). Az 592 fős populációban a rs1003853 CC homozigóta genotípusának frekvenciája szintén statisztikailag szignifikánsan magasabb az RA betegek körében, mint kontrollokban, melynek eredményeképpen a recesszív genetikai modellben statisztikailag szignifikáns asszociációt tapasztaltunk a CC homozigóta genotípus és az RA között (OR=1,618, 95% CI 1,16-2,25, p=0,004).

Az öt lókuszt közül a 2. csoportban az *AIRE* SAND exon 5-ös SAND domént kódoló rs878081 C-alléla mutatott szignifikáns összefüggést az RA kockázatával, az allélikus modell szerint (OR=1,48, 95% CI 1,05-2,09, p=0,023). Itt a recesszív genetikai modell alapján a CC homozigóta RA-betegek genotípus-frekvenciája szignifikánsan magasabb volt a CT heterozigóták és a TT homozigóták genotípus-frekvenciáinak összegéhez képest. Recesszív genetikai modell alapján kijelenthető, hogy a CC genotípus és RA között az összefüggés statisztikailag szignifikáns (OR=1,64, 95% CI 1,09-2,48, p=0,01).

A 2. csoportban, a továbbiakban az *AIRE* intron 7-es rs1003854 lókusza volt a következő, ahol a T-allél előfordulásának az esélye szignifikánsan magasabb volt az RA allélok között a C-allélhoz képest (OR=1,52, 95% CI 1,08-2,12, p=0,014). A recesszív genetikai modell alapján a TT homozigóta genotípus esélye szignifikánsan magasabb volt RA betegek körében, a kontrollokhoz képest (OR=1,72, 95% CI 1,15-1,44, p=0,008).

Az RA betegek körében statisztikailag szignifikáns különbségeket találtunk a 2. csoport rs878081 recesszív modelljében, ahol a CC homozigóták szignifikánsan magasabb átlag ESR szinteket mutattak (p=0,027) a TT+CT genotípusú alpopulációhoz képest. Pearson-féle kétváltozós korrelációval statisztikailag szignifikáns összefüggést találtunk az ESR és a CC

homozigóta genotípus között ( $p=0,023$ ,  $r=0,190$ ), bináris logisztikus regresszióval szintén statisztikailag szignifikáns összefüggést találtunk ( $OR=1,01$ , 95% CI 1,002-1,032,  $p=0,026$ ) a két változó között. Az 1. csoport 592 alanyával 270 RA-betegnél szignifikánsan magasabb aCCP-szintet mutatott ki az rs1055311 lókuszt kodomináns heterozigóta modell CT és domináns modell CC+CT genotípus-csoportja ( $p=0,028$ ,  $r=0,250$  illetve  $p=0,044$ ,  $r=0,150$ ).

#### 4.5. Regulatorikus kötőmotívumok allélspecifikus affinitása transzkripció faktorokhoz

A SAND doménben az 5-ös exonon rs878081 és a 6-os exonon található rs1055311 allélpolimorfizmusok szinonim polimorfizmusok, melyek a fehérjében strukturális alapú funkcióvesztést nem, de allélspecifikus funkcionális különbségeket képesek okozni. A lehetséges transzkripció következmények tisztázása érdekében allélspecifikusan, a HaploReg v4.2 adatbázis segítségével *in silico* analízist végeztünk annak érdekében, hogy meglássuk, az egyes allélpolimorfizmusok helyet foglalnak-e bizonyos TF-k DNS kötőregiójának motívumaiban, és ha igen akkor hatással vannak-e az egyes allélok a TF kötődési affinitására (Regulatory motifs altered).

Azt a négy lókuszt vizsgáltuk, melyeknek allélpolimorfizmusai összefüggésbe hozhatóak RA kockázatával illetve RA csoporton belül klinikai paraméterekkel úgy, mint ESR vagy aCCP. A helysúlyozott mátrix predikciós modellje az rs878081 esetében kimutatta, hogy maga a lókuszt pozíciója NF- $\kappa$ B és a Yin-Yang 1 (YY1) kötőmotívumainak szekvenciájában helyezkedik el.

Az NF- $\kappa$ B NF-kappaB\_disc2 DNS kötőmotívumához számított LOD-érték +11,1 értékkel magasabb a T-allél esetében, a NF-kappaB\_known4 esetében pedig ez az érték +2, mely eredmények azt mutatják, hogy a referencia C-allél jelenlétében az NF- $\kappa$ B gyengébben kötődik a DNS-hez. Ugyanezen a lókuszon a YY1\_disc1 kötőmotívumhoz az alternatív T-allél esetében az affinitást jelentő LOD-érték +0,9 értékkel magasabb, ami szintén azt jelenti, hogy a referencia C-allél jelenlétében ez a TF gyengébben kötődik.

A továbbiakban a helysúlyozott mátrix predikciós modellje az rs1003853 kimutatta, hogy maga a lókuszt pozíciója a Krüppel-like zinc finger protein AP-2rep (AP-2rep) és a Rad21 transzkripció faktorok kötőmotívumainak szekvenciájában helyezkedik el. Az AP-2rep DNS kötőmotívumához számított LOD-érték +7,1 értékkel, a Rad21\_disc5 kötőmotívum esetében +0,5 értékkel magasabb a T-allél esetében, vagyis C referencia allél jelenlétében ez a két TF gyengébben kötődik.

A helysúlyozott mátrix predikciós modellje szerint az rs1055311 esetében a lókuszt pozíciója a myeloid cink-ujj 1-nek (myeloid zinc finger 1, MZF1) a MZF1::1-4\_1 motívumában, NF- $\kappa$ B transzkripció faktor NF-kappaB\_disc2 és p300 transzkripció koaktivátor p300\_known1 kötőmotívumának szekvenciájában foglal helyet. Prediktív LOD-érték következetesen az MZF1::1-4\_1 és NF-kappaB\_disc2 esetében +1,6 és +11,1, mely azt mutatja, hogy az alternatív T-allél jelenlétében erősebb és a referencia C-allél jelenlétében gyengébb affinitással kötődik a TF. A p300\_known1 esetében a LOD-érték megfordul a referencia C-allél javára + 1,3 értékkel, mely fordított helyzetnél az előzőektől eltérően az alternatív T-allél köti gyengébben a p300 transzkripció koaktivátort.

A predikciós modell alapján az rs1003854 lókuszt a GATA a cink-ujj fehérje család – melyek képesek a (T/A)GATA(A/G) konszenzusos DNS-szekvenciához kötődni – GATA\_disc3 kötőmotívumának szekvenciájában foglal helyet. A prediktív LOD-érték +1 értékkel

magasabb az alternatív C-allél esetében, vagyis a referencia T-allél ezt a transzkripció faktorot gyengébben köti. Ezen vizsgálatrés eredményeit az 1. táblázat szemlélteti.

Lókusz	Ref	Alt	Position Weight Matrix (helysúlyozott mátrix) ID	LOD value Ref	LOD value Alt
rs878081	C	T	NF-kappaB_disc2	1,5	12,6
			NF-kappaB_known4	12,5	14,5
			YY1_disc1	8,6	9,5
rs1003853	C	T	AP-2rep	3,0	10,1
			Rad21_disc5	10,2	10,7
rs1055311	C	T	MZF1::1-4_1	10,3	11,9
			NF-kappaB_disc2	1,7	12,8
			p300_known1	11,1	9,8
rs1003854	T	C	GATA_disc3	12,2	13,2

Ref: referencia allél; Alt: alternatív allél; LOD: log odds ratio

**1. táblázat – HaploReg4.2 adatbázis predikciós modellje alapján az AIRE allélpolimorfizmusok pozíciója transzkripció faktorok DNS kötőrégiójának motívumaiban, és allélspecifikus hatásuk a TF kötődési affinitására,**

## 5. MEGBESZÉLÉS

Kutatásunk egyik legfőbb célja az volt, hogy behatóan megismerjük az AIRE transzkripció faktor összefüggését lehetséges autoimmun betegségekkel. Mivel az AIRE funkciója az autoreaktív timociták kiszűrése és törlése, a tímusz zsíros elcsökevényesedését – mely pubertáskorban kezdődik – követően fiatal felnőttkorban az expresszió helyileg kimutathatatlan, igazolja azt a feltevést mely szerint az AIRE transzkripció faktor működése a csecsemőkorral van összefüggésben. Továbbá a tímusz életkorhoz köthető természetes funkcióvesztése eredményezi azt, hogy az immunrendszer felnőttkorra egy meglévő, a továbbiakban nagymértékben nem gyarapodó T-sejt repertoárral dolgozik, a naív T-sejtek termelése szignifikánsan alábbhagy. Amennyiben tehát a genetikai alapú betegségkockázat megléte bebizonyosodik azon túl, hogy ez összefüggést jelent az AIRE transzkripció faktorral, egyben összefüggést jelent a csecsemőkorral, ami a betegség patomechanizmusát új megvilágításba helyezné. Az AIRE a negatív szelekció alapjául szolgáló szöveti antigén termelésért felelős. Amennyiben nem vagy nem kellő mértékben termelődik, a centrális tolerancia ezen lépése hiányos lesz vagy elmaradhat, megemelve az autoreaktív, naív T-sejtek klonális túlélését. A fentiekben ismertettem Lovewell és mtsai.-nak munkáját, melyben kettő, az rs117557896, rs751032 lókuszon található AIRE-655G AIRE-230T allélpolimorfizmusok alkotta genotípus esetében a promóter aktivitása alig kimutatható és a negatív szelekció elmaradt. Amennyiben két allélpolimorfizmus megléte ilyen drasztikus expressziós különbséget képes eredményezni, elképzelhető, hogy az AIRE promóter régióján allélpolimorfizmusok sora és az azok általi haplotípusok egy személyre szabott AIRE expressziós profilt hoznak létre. A populáció egészét tekintve az esetek többségében ez elegendő expressziót biztosít, viszont előfordulhat olyan AIRE transzkripciót érintő haplotípus kombináció, mely nem biztosít kellő mennyiségű AIRE fehérjét a szöveti antigének expressziójához, a negatív szelekció elmaradhat, mely az autoreaktív, naív T-sejtek klonális expanziójához vezethet. Számos autoimmun betegség kiváltó oka az autoreaktív T-sejt populáció, így amennyiben az AIRE egy kulcsszereplő az immunológiai toleranciában elképzelhető, hogy allélpolimorfizmus által számos autoimmun betegség közös nevezője, oki

tényezője lehet, ennél fogva az AIRE számos autoimmun betegség kiindulópontja lehet. Ennek kutatásához csecsemőkori AIRE expressziós profil kialakítására, haplotípus analízisre lenne szükség tímusz szövetmintákon, valamint a csecsemők élethossz-kohorsz utánkötése, egy nagy elemszámú vizsgálatban kontroll ellenében. Továbbá mindemellett a centrális tolerancia teljes szereplőgárdájának vizsgálata adhatna egy részletesebb képet. Ilyen fokú komplex vizsgálatok kivitelezésére nem törekedtünk. Vizsgálatsorozatunk egyes lépései egymásból alakultak ki. Meta-analízisünkkel a vizsgálat legelején arra törekedtünk, hogy meghatározzuk a tudomány jelenlegi állása szerint melyik az az autoimmun betegség, mely leginkább összefüggésbe hozható az AIRE allélpolimorfizmusokkal. Publikációnk során miután meghatároztuk, hogy ez a reumatoid artritisz, azt tisztáztuk, mely allélpolimorfizmusok melyik rasszt tekintve adnak szignifikáns összefüggést. A legnagyobb populáció melyet bevonhattunk ázsiai volt, melyben meghatároztunk RA-ra nézve az rs2075876 lókuszt A-allélla és a rs760426 G-allélla, illetve ezeket az allélokat hordozó genotípusok – legyen homo- vagy heterozigóta – szignifikáns összefüggésben állnak a betegséggel, majd hogyan az összes genetikai modellt tekintve. Meta-analízisünk tehát megállapította, mely allélpolimorfizmusok jelentenek kockázatot ázsiai populációban RA-ra a közlemények állása alapján, 2018-ig bezárólag. Minekutána bebizonyosodott, hogy az allélpolimorfizmusok összefüggésben állnak a betegséggel, ugyanakkor szignifikáns összefüggést nem kaptunk európai populációra így kíváncsiak voltunk, létezik-e kockázatot jelentő allélpolimorfizmus kaukázusi személyek között. Ennek hatására megterveztük a további lépéseket, melynek során *in silico* vizsgálattal 44 SNP-t detektáltunk az AIRE régióban melyek potenciálisan hozzájárulhatnak autoimmun kórkép kialakulásához. Ebből a 44-ből választottunk ki kettőt a DNS-kötő régiót kódoló SAND domén vizsgálatára: rs878081 és rs1003854 lókuszt. Pilot vizsgálatunk elsődleges szerepe az összefüggések keresése és a vizsgálati módszerek optimalizálása volt. Eredményeink alapján rs878081 és rs1003854-es lókuszt szignifikáns összefüggést mutatott a betegséggel, a betegségkockázatot jelentő allél hordozása kockázatot jelentett vérsüllyedés és CRP emelkedésre. Közleményünk elsőként hozta összefüggésbe rs878081 és rs1003854 lókuszt adta genotípusokat RA-val, továbbá szintén elsőként sikerült összefüggést találni genotípusok és klinikai paraméterek (ESR és CRP) között a betegpopulációban. Azon RA betegek, melyek kettő kockázatot jelentő allélt hordoztak az rs1003854-es lókuszon, szignifikánsan magasabb CRP-értékkel rendelkeztek mind a recesszív ( $p=0,029$ ), mind pedig a kodomináns ( $p=0,043$ ) genetikai modellben, mely genetikai konstitúció egyben pozitív korrelációt is mutatott CRP-vel mind a két modellben ( $p=0,002$  és  $p=0,006$ ). Ennek magyarázatára *Bergström és mtsai.*-nak vizsgálata adhat magyarázatot. A betegség centrális szakaszában alakul ki az ízületi tokban a gyulladási környezet, melynek során fibroblasztszerű szinoviocita (fibroblast like synovocyte, FLS) sejtek kemokineket (CXCL10, CCL2, CCL15, CCL8) termelnek, a jelenlévő VCAM-1 sejtekkel segítik az itt keringő autoreaktív T-sejtek extravazációját. A kísérlet során azt tapasztalták, hogy ha az AIRE expressziót kis, nemkódoló siRNA-sal némítják FLS populációban, a sejtek 26x magasabb kemokin produkciót mutatnak szemben a nem stimuláltakhoz képest, ahol kemokintermelés kimutathatatlan volt, vagyis az AIRE-némított FLS sejtpopuláció gyulladási aktivitása magasabb volt. Allélpolimorfizmusok mint láttuk nagymértékben hozzájárulhatnak az AIRE expressziójához, vagyis elképzelhető hogy az AIRE rs1003854 lókusztán a kockázati allélra homozigóta genotípus nagymértékben csökkentette az expressziót, mely évtizedekkel később az FLS sejteken keresztül kialakított egy kemokinekben gazdag környezetet autoreaktív T-sejtek számára. Ezt az is alátámasztja, hogy a kemokinek (CXCL10, CCL8 és CCL5) erősen asszociálnak CRP szintekkel. Pilot studyt követően vizsgálatunkat kiterjesztettük további 3 lókuszt és 592 főre. A genetikai expozíció és a betegség közti összefüggés érdekében öt lókuszt allélpolimorfizmusait vizsgáltuk, melyek közül a SAND domén 5-ös exonjában található rs878081, az 5-ös intron



splice helyen található rs1003853 és a 7-es intron splice helyen található rs1003854 szignifikáns pozitív asszociációt mutatott a betegségkockázattal abban az esetben, ha a kockázatot jelentő allélra a genotípus homozigóta ( $p=0.017$ ,  $p=0.004$ ,  $p=0.008$ ). Két lókuszesetében (rs878081, rs1055311) pedig összefüggés mutatkozott RA betegek genetikai alcsoportjai és ESR valamint aCCP klinikai paraméterek között. Eredményeink alapján az rs1055311 lókuszesetében az allél- és genotípus-frekvenciája nem mutat szignifikáns összefüggést a betegséggel, de korrelál az aCCP-szintekkel. Ez többek között azért fordulhat elő, mert az allélpolimorfizmus episztatikus úton, köztes változóra hat, mely az autoantitest termelés, vagyis egy köztes fenotípust határoz meg, mely allélpolimorfizmus így közvetlenül nincs összefüggésben a betegségkockázattal. Befolyásolhatja az autoreaktív T-sejtek túlélését, amelyek belsőleg elősegítik az aCCP antitestek termelődését, jellemezve a betegség köztes fenotípusát.

Mindezek betegségkockázat eredmények alapján tehát úgy gondoljuk, hogy mind az exonok szinonim, mind pedig splice régiók introniális allélpolimorfizmusai illetve az azok által kialakított genotípus jelenlétében az AIRE expresszió kellően alacsony ahhoz, hogy a szöveti antigének expressziója megvalósuljon, így a kontroll alól autoreaktív timociták szabadulhatnak ki, melyek naív T-sejtté alakulva a másodlagos limfoid szövetekben klonálisan túlélnek és egészen a centrális szakasz extravazációjáig rejtve maradnak. Innentől több irányban triggerelik a patomechanizmust, egyrészt citokineket termelnek (IFN- $\gamma$ , IL-17), melyek később makrofágokat idevonzva circulus vitiosus fenntartva, másrészt autoreaktív CD4<sup>+</sup> T-sejtek az FLS sejtpopuláció expanzióját okozza, melynek eredménye a heges kötőszövet képződése. Kockázatot jelentő allélra homozigóta egyének esetében elképzelhető, hogy genetikai kockázatból adódó alacsony AIRE expresszió egy az átlagnál aktívabb gyulladáshoz teremt, azáltal, hogy kemokinek széles spektruma termelődik AIRE hiányában az FLS sejtekben, melynek hatására további citokinek, autoantitestek, felgyülemlett immunkomplexek képződnek.

Vizsgálatunk utolsó lépéseként ok-okozati összefüggések mélyebb feltárása érdekében *in silico* elemeztük a kockázati alléljaink molekuláris következményeit. Ennek során azt találtuk, hogy az 5 lókuszból 4, melyek vagy a betegségkockázattal, vagy pedig a klinikai paraméterekkel mutattak statisztikailag szignifikáns összefüggést, transzkripciós faktorok DNS kötőmotívumainak szekvenciájában foglalnak helyet és *in silico* adatok alapján hatással lehetnek azok kötődési affinitására. A kapott eredményeket az alapján választottuk ketté, hogy a referencia (Ref) allél jelenlétében erősebb-e vagy gyengébb a transzkripciós faktor kötődés a DNS kötőmotívum szekvenciájához, melyben az allélpolimorfizmus helyet foglal. A transzkripciós fátumot nem minden esetben lehetséges meghatározni, egy-két kivételtől eltekintve erősen kontextusfüggők. Azt találtuk, hogy az exon 5-ös rs878081 és exon 6-os rs1055311 Ref alléljának (C) jelenlétében gyengébben kötődik az NF- $\kappa$ B. *Haljasorg és mtsai*-nak vizsgálata alapján ennek a transzkripciós faktornak az erős affinitású kötődése az mTEC DNS-hez szükséges a negatív szelekciós AIRE expresszióhoz a tímuszban (35). Mindkét lókuszesetében a kockázatot jelentő allél a Ref C-allél volt, melynek homozigóta genotípusa asszociált a betegségkockázattal rs878081 esetében és aCCP-vel rs1055311 esetében. Esetkontroll és *in silico* eredményeink alapján elképzelhetőnek tartjuk azt a feltevést, hogy ezen allél jelenlétében nem volt képes megfelelő affinitással kötődni az NF- $\kappa$ B az mTEC DNS-éhez, melynek következménye az alacsony AIRE expresszió illetve az abból fakadó szöveti antigén expresszió, mely eredményezhette a negatív szelekció részleges vagy teljes elmaradását. Továbbiakban az rs878081 esetében Ref allélhoz gyengébben kötődött YY1, mely egy cink-ujj multifunkciós transzkripciós faktor, mely kontextus függően aktivátor vagy represszor. Az utóbbira példa, amennyiben hiszton deacetiláz (histone deacetylase, HDAC) jelenlétében kötődik a target promóter régiójához, transzkripciós repressziót, kromatinkondenzációt okoz. Nagyságrendekkel gyengébben kötődik az intron 5-ös

rs1003853 Ref (C) alléljához az AP-2 transzkripciós faktor, mely DNS-kötő doménjével promóter és enhancer régiókhoz kötődik, erős affinitása transzkripciót aktivál. Ezen lókuszt Ref allélja jelenlétében, szintén gyengébben kötődik a Rad21, melynek erős affinitású kapcsolódása szükséges a DNS-repairhez, kohezin-komplex része, mely segíti a kromoszómaszegregációt meiózis és mitózis alatt. Gyenge affinitású kötődése befolyással lehet ezeknek a funkcióknak a betöltésére (178). Eredményeink alapján elmondható, hogy az rs1055311 lókuszt Ref (C) alléljához MZF1 transzkripciós faktor gyengébben kötődik, melynek szerepe többek között a mieloid DC sejtek fejlődésének a biztosítása. Az intron 7-es rs1003854 Ref (T) allélja esetében a GATA transzkripciós faktor család gyengébben kötődik a DNS TF kötő motívumához. A GATA alapvető fontosságú a timociták érésében, hatással van a  $\beta$ -szelekcióra és meghatározzák a sejtvonalak elköteleződését a centrális toleranciában (180). Mivel ez az allélpolimorfizmus intronikus, elképzelhető, hogy alternatív splicing folyamatán lesz képes hatni az AIRE expresszióra (181). A TF-ok közül 1 kötődött erősebben a Ref allélhoz, ez pedig a p300 az rs1055311 C-alléljához. Hiszton acetil transzferáz (HAT) enzimeként acetilsoportot transzportál a hiszton molekulák lizin oldalláncára mellyel a kromatint elektromos töltések következményeként relaxálja és elérhetővé teszi a DNS-t RNS polimeráz II (RNA pol II) számára, vagyis transzkripciót aktivál. *In silico* vizsgálatunk eredményeként elmondható, hogy az 5 lókuszból 4 esetében a Ref allélt tartalmazó DNS kötőmotívumhoz számos TF gyengébb affinitással kapcsolódik, melyek közül hangsúlyosak voltak azon transzkripciós faktorok, melyek alapvetően a génexpressziót határozzák meg, előrevetítve ez úton is az AIRE allélpolimorfizmus-függő expresszióját ennek megfelelően TRA-vezérelt negatív szelekcióját.

Eredményeink mellett vizsgálatainknak számos limitációja van. Legfőbb limitáció, mely az összefüggés erősségét negatívan befolyásolhatta a korlátozott elemszám. Ebben elsődlegesen a COVID-19 világvárvány miatti betegforgalom visszaesés játszott szerepet. Másodlagosan pedig a két kórház nettó betegforgalma, mely 2019 és 2023 között ennyi RA beteg és kontroll vizsgálatba kapcsolására volt képes. Egy világvárvánnyal terhelt 4 éves mintagyűjtési időszakban ez volt az a maximum melyet a két kórház biztosítani tudott, mely teljesítményért nem győzünk hálásak lenni a Harkányi Termál Rehabilitációs Centrum és a Hévízgyógyfürdő és Szent András Reumakórház főorvosainak, orvosainak, nővéreinek akik nélkül ez a vizsgálat sorozat nem valósulhatott volna meg. A vizsgálatunkban résztvevők számának növelése jelentősen emelte volna mind az összefüggések számát mind pedig a statisztikai szignifikanciát. További korlátot jelent, hogy a vizsgálatba egy domén kódoló régiót vontunk be, melynek eredményei nem képesek átfogó betekintést nyújtani az AIRE funkciójába és egyes allélpolimorfizmusok AIRE-re mért expressziós következményeibe. Elképzelésünk szerint az összes domén kódoló régióban található összes allélpolimorfizmus és AIRE-partnerek genom szintű elemzése és eredményei nem adnának értékelhető eredményeket anélkül, hogy azokat izoláltan, egymástól elkülönítve ne vizsgálnánk meg, melynek folytatása jövőbeli terveink közé tartozik. További limitáció, hogy az AIRE mRNS-expresszióját nem állt módunkban vizsgálni, mely nagymértékben hozzájárult volna az allél- illetve a genotípusfüggő transzkripciós következmények jobb megértéséhez. Jövőbeli terveink között szerepel az RA és a centrális tolerancia közötti kapcsolat, valamint az AIRE partnerek szerepének átfogóbb megértése további genetikai asszociációs vizsgálatokon a fennmaradó 11 intronikus és 12 exonikus kódoló szekvenciára vonatkozóan.

## 6. ÚJ EREDMÉNYEK

Vizsgálatunk összességében 6 részből állt. Elsőként ázsiai populációkban az rs2075876 és rs760426 lókuszon lévő allélpolimorfizmusok meta-analízisét végeztük el. Itt 2018-ban elsőként határoztunk meg, hogy az rs2075876 kockázatot jelentő A-allélja valamint rs760426

kockázatot jelentő G-allélja illetve az általuk meghatározott genotípusok és genotípus csoportok statisztikailag szignifikáns összefüggést mutattak RA-val, kizárólag ázsiai populációban. Ezt az összefüggést kutatócsoportunk publikálta először.

Következő lépésként MAF allélfrekvenciák alapján meghatároztunk 44 olyan SNP-t, melyek potenciálisan vizsgálhatóak európai, kaukázusi populációkban melyekből kettőt, rs878081 és rs1003854-t kiválasztva elvégeztük pilot vizsgálatunkat, melynek végére elsőként határoztuk meg, hogy az rs878081 és rs1003854 kockázatot jelentő alléljai és genotípusai statisztikailag összefüggnek betegségkockázattal illetve ESR és CRP klinikai paraméterekkel. Európai populációban elsőként kutatócsoportunk pilot vizsgálata elemezte az allélpolimorfizmusok és RA, valamint klinikai paraméterek közötti összefüggést hat különböző (allélikus, domináns, recesszív, kodomináns hetero- és homozigóta) genetikai modell alapján.

Az addigi eredmények alapján következő vizsgálatunk elemszámát 529 főre és lókuszállományát 3 további pozícióra kiterjesztve, azt az eredményt kaptuk, hogy az öt kiválasztott lókuszból az rs878081 és rs1003853 kockázati C- valamint rs1003854 T-allélja illetve azok genotípusai statisztikailag szignifikáns összefüggésben állnak RA betegségkockázattal. Klinikai paraméterek közül ESR statisztikailag szignifikáns asszociációt és az aCCP pedig korrelációt mutatott a kockázatot jelentő C-allélra homozigóta genotípussal az rs878081 és rs1055311 esetében.

Vizsgálatsorozatunk utolsó lépéseként *in silico* elemeztük az egyes allélpolimorfizmusok adta transzkripciós következményeket, melynek során elsőként határoztuk meg 7 különböző transzkripciós faktor viselkedését kockázatot jelentő allélpolimorfizmus jelenlétében. Ezek közül egyezményesen kijelenthető, hogy a kockázatot jelentő SNP-k jelenlétében 7-ből 6 TF gyengébben kötődik a DNS kötőmotívumhoz, melyek közül 4-ből 2 SNP az NF- $\kappa$ B gyengébb kötődését mutatták, mely a neonatális negatív szelekció során felelős a TRA expressziót biztosító AIRE transzkripcióért. Eredményeink alapján az egyik lehetséges következmény a kockázatot jelentő allélok jelenlétében az NF- $\kappa$ B gyengébb kötődése, így a neonatális negatív szelekció során nem termelődik megfelelő mennyiségű AIRE illetve TRA, mely előíranyozza az autoreaktív timociták lehetséges klonális túlélését, megteremtve ez által autoimmun betegségek, köztük RA kialakulásának kezdeti immunológiai talaját.

## 7. TUDOMÁNYOS KÖZLEMÉNYEK

**Bérczi B**, Gerencsér G, Farkas N, Hegyi P, Veres G, Bajor J, Czopf L, Alizadeh H, Rakonczay Z, Vigh É, Erőss B, Szemes K, Gyöngyi Z. Association between AIRE gene polymorphism and rheumatoid arthritis: a systematic review and meta-analysis of case-control studies. *Sci Rep.* 2017 Oct 26;7(1):14096. doi: 10.1038/s41598-017-14375-z. PMID: 29074995; PMCID: PMC5658331 (**D1, IF: 4,6**)

**Berczi B**, Nusser N, Peter I, Nemeth B, Gyongyi Z. Association Between AIRE Polymorphisms rs870881(C>T), rs1003854(T>C) and Rheumatoid Arthritis Risk: A Hungarian Case-control Study. *In Vivo.* 2024 Mar-Apr;38(2):774-784. doi: 10.21873/invivo.13501. PMID: 38418155; PMCID: PMC10905445. (**Q2, IF: 2,4**)

**Bérczi B**, Nusser N, Péter I, Németh B, Kulisch Á, Kiss Z, Gyöngyi Z. Genetic Polymorphisms in Exon 5 and Intron 5 and 7 of AIRE Are Associated with Rheumatoid Arthritis Risk in a Hungarian Population. *Biology.* 2024; 13(6):439. <https://doi.org/10.3390/biology13060439> (**Q1, IF: 4,2**)

## 8. KÖSZÖNETNYILVÁNÍTÁS

Szeretnék köszönetet mondani az Orvosi Népegészségtani Prof. Dr. Kiss István intézetvezetőnek, hogy lehetővé tette, hogy munkámat az általa vezetett intézetben végezhessem.

Hálásan köszönöm Dr. Gyöngyi Zoltán egyetemi docensnek, témavezetőmnek, hogy szakmailag és emberileg is támogatott munkám során.

Köszönettel tartozom Dr. Péter Ivánnak, a Harkányi Termál Rehabilitációs Központ kórházigazgatójának, hogy lehetővé tette, hogy munkámat az általa vezetett kórházban végezhessem.

Köszönöm Dr. Németh Balázs, egyetemi adjunktusnak, hogy a kutatás kezdeti lépéseiben segítséget nyújtott.

Köszönettel tartozom Dr. Nusser Nórának, a Harkányi Termál Rehabilitációs Központ orvosigazgatójának, azért a mérhetetlen szakmai segítségért, szervezésért, továbbá jóindulatért és bizalomért, emberségért, melyet munkám során tőle kaptam, melyek nélkül ez a vizsgálat, ebben a formájában elképzelhetetlen lett volna.

Köszönettel tartozom Feledi Zsuzsannának, a Harkányi Termál Rehabilitációs Központ laboratórium vezetőjének, hogy a mintákon dolgozó csapatot összefogta, továbbá Tóthné Kamanczi Zita laboránsnak, hogy szisztematikusan és fáradhatatlanul gyűjtötte és katalogizálta mintáinkat. Köszönettel tartozom továbbá Pappné Netik Natáliának, Óbert Anita-nak és Szűcs Adriennek a harkányi mintagyűjtésben nyújtott segítségükért.

Köszönettel tartozom Dr. Holléné Dr. Mándó Zsuzsannának, a Hévízgyógyfürdő és Szent András Reumakórház kórházigazgatójának, hogy lehetővé tette, hogy munkámat az általa vezetett kórházzal együttműködésben végezhessem.

Köszönettel tartozom Dr. Kulisch Ágotának, a Hévízgyógyfürdő és Szent András Reumakórház osztályvezető főorvosának, azért a rengeteg segítségért mellyel lehetővé tette, hogy az együttműködés és mintagyűjtés létrejöhesse. Nélküle szintén elképzelhetetlen lett volna, hogy vizsgálatunk megvalósuljon.

Köszönettel tartozom Dr. Fülöp Jánosnének, a Hévízgyógyfürdő és Szent András Reumakórház főnővérének, a hévízi mintagyűjtés során nyújtott mérhetetlen segítségért.

Továbbá köszönet illeti az Orvosi Népegészségtani Intézet összes munkatársát munkám során nyújtott segítségükért.

Végül, de nem utolsó sorban köszönöm feleségemnek azt a támogatást és biztatást mellyel nagymértékben hozzájárult a dolgozat elkészüléséhez. Továbbá köszönöm egész családomnak szintén azt a támogatást és biztatást, mellyel nagyban hozzájárultak a dolgozat elkészüléséhez. A kutatások támogatásáért köszönet illeti a PTE ÁOK, Kari Kutatási Alap (ÁOK-KA) KA-304516 számú pályázatát.