

**SEMMELWEIS EGYETEM
DOKTORI ISKOLA**

Ph.D. értekezések

2708.

MIKÓ ÁGNES

**Krónikus betegségek gyermekkori prevenciója
című program**

Programvezető: Dr. Szabó Attila, egyetemi tanár
Témavezető: Dr. Tory Kálmán, egyetemi docens

Inkompletten penetráns mutációk és interallelikus
interakciók azonosítása populációgenetikai módszerrel
autoszomális recesszív kórképekben

Doktori értekezés

Dr. Mikó Ágnes

Semmelweis Egyetem
Klinikai orvostudományok Doktori Iskola



Témavezető: Dr. Tory Kálmán, Ph.D., egyetemi docens

Hivatalos bírálók: Dr. Nemoda Zsófia, Ph.D., egyetemi adjunktus

Dr. Egyed Balázs, Ph.D., egyetemi adjunktus

Szigorlati bizottság elnöke: Dr. Darvas Zsuzsanna, Ph.D., egyetemi docens

Szigorlati bizottsági tagok:

Dr. Haltrich Irén, Ph.D., tudományos főmunkatárs

Dr. Balogh István, Ph.D., egyetemi docens

Budapest

2022

Tartalomjegyzék

1. Rövidítések jegyzéke.....	3
2. Bevezetés.....	4
2.1. Penetrancia.....	4
2.1.1. Penetranciát befolyásoló tényezők AR betegségekben	5
2.2. Sztteroid-rezisztens nephrosis szindróma.....	9
2.3. Podocin	10
2.4. <i>NPHS2</i> gén	12
2.5. Domináns negatív hatás és interallelikus komplementáció	13
4. Célkitűzés	17
5. Módszerek.....	18
5.1. Variánsok vizsgálata populációgenetikai módszerrel.....	18
5.2. Autoszomális recesszív betegségek kiválasztása	18
5.3. Beteg adatok összegyűjtése	21
5.4. Mutációk nevezéktanának egyesítése	24
5.5. Cisz variánsok	24
5.6. A variánsok osztályozása.....	24
5.6.1. Genotípus szerinti beosztás: LOF/misszensz besorolás	25
5.6.2. Null-hipomorf beosztás	25
5.6.3. Fehérjedomén-lokalizáció	28
5.7. A beteg populáció európai eredetű betegekre való szűkítése	29
5.8. Vérrokonság hatásának kompenzálása	31
5.9. Patogenitás validálása.....	31
5.10. Intrauterin letalitással járó betegségek kizárása	33
5.11. Penetrancia számolás	33
5.12. Asszociációs tesztek	38
5.13. Populációgenetikai módszer validálása	39
5.14. Az <i>NPHS2</i> társulások vizsgálata FRET módszerrel.....	40
5.15. Statisztikai elemzések.....	42
5.16. Klinikai irányelv létrehozása az <i>NPHS2</i> R229Q asszociációk értelmezésére	43
6. Eredmények	44
6.1. Nem patogén variánsok azonosítása.....	44
6.2. Intrauterin letalitást okozó betegségek kizárása	49
6.3. Inkompletten penetráns variánsok azonosítása.....	53
6.4. Populációgenetikai módszer validálása	55
6.5. Interallelikus interakciók azonosítása.....	55
6.6. A podocin variánsok vizsgálatának eredményei	57
6.7. Klinikai irányelv az <i>NPHS2</i> R229Q asszociációk patogenitásának értelmezésére	60
7. Megbeszélés	66

8. Következtetések	73
9. Összefoglalás	74
10. Summary	75
11. Irodalomjegyzék	76
12. Saját publikációk jegyzéke.....	88
12.1. A disszertációhoz kapcsolódó közlemények jegyzéke	88
12.2. A disszertációtól független közlemények jegyzéke.....	89
13. Köszönetnyilvánítás	90

1. Rövidítések jegyzéke

AC	allele count, allélok száma
AF	allele frequency, allélfrekvencia
AR	autoszomális recesszív
BBS	Bardet-Biedl szindróma
CBAVD	congenital bilateral absence of the vas deferens, veleszületett kétoldali vas deferens hiány
CF	cisztás fibrózis
CTT	C-terminális helikális rész
ESRD	end stage renal disease, végstádiumú veseelégtelenség
ExAC	The Exome Aggregation Consortium
FRET	Förster-féle / fluoreszcens rezonancia-energia transzfer
FSGS	Focalis segmentalis glomerulosclerosis
gnomAD	Genome Aggregation Database
GFP	zöld fluoreszcens fehérje
HA	Influenza Haemagglutinin
HGMD	The Human Gene Mutation Database
HGVS	Human Genome Variation Society
LOF	Loss of Function, teljes funkcióvesztést okozó
NMD	nonsense-mediated mRNA decay, nonszensz mutáció által mediált mRNS lebontás
OMIM	Online Mendelian Inheritance in Man, a humán gének és genetikai rendellenességek online adatbázisa
SRNS	szteroid-rezisztens nephrosis szindróma

2. Bevezetés

2.1. Penetrancia

A penetrancia azt fejezi ki százalékosan, hogy az azonos genotípusú egyedek csoportjában az egyedek hány százaléka mutatja a genotípusra jellemző fenotípust. Ha a megegyező genotípusú egyedek mindegyikében megjelenik az adott fenotípus, akkor a gén teljes penetranciájú (100%), ha csak az egyedek egy részén látszik a fenotípus, akkor a gén penetranciája nem teljes, tehát inkomplett (1).

Bár gyanítható, hogy a módosító és a környezeti tényezők befolyásolják a monogénes rendellenességeket, csak a közelmúltban kezdték megérteni molekuláris alapjukat (1).

A közelmúltban három nagy tanulmányt is publikáltak, melyek az egészséges személyek genomjában jelenlévő számos patogén mutációra hívják a fel a figyelmet (2-4).

Az adatok alapján 185 egyén genom szekvenálása során 1285, vélhetően teljes funkcióvesztést okozó (LOF) mutációt (nonszensz, hasítási helyet érintő mutációk, mikro-inszerciók és deléciók) azonosítottak. Ezen eredmények szerint egy átlagos emberi genom 100 LOF mutációt tartalmaz és 20 gén mindkét kópiája inaktív. További vizsgálatok alapján 179 egészséges „1000 Genom projektben” résztvevő ember teljes genom-szekvenálása során minden egyén 281-515, nagy valószínűséggel kóroki misszensz mutációt hordoz, ebből 40-85 variánst homozigóta formában (2).

A molekuláris genetika fejlődésével az elmúlt néhány évben az inkomplett penetrancia mögöttes mechanizmusainak egy részére találtak magyarázatot. Ide tartozik a mutáció típusa, az epigenetikai tényezők, a gén-környezet kölcsönhatás, az életkor és a nem befolyása, az oligo- és a digenikus öröklődésű betegségek, a módosító gének, a kópiaszám variációk, valamint a további génvariánsok hatása (1).

A közelmúltig az inkomplett penetrancia kifejezést szinte kizárólag az autoszomális domináns betegségekben azonosították. Ilyen betegség a congenitalis cataracta (*GJA3*) (5), a retinitis pigmentosa (*PRPF8* & *PRPF31*) (6) és a hosszú QT szindróma (7).

Inkomplett penetrancia azonban az autoszomális recesszív rendellenességekben is előfordulhat, amikor ugyanazon mutáció nem minden esetben vezet a fenotípus kialakulásához, részben a másik allélon levő mutáció függvényében. Inkomplett penetranciára utal az is, hogy molekuláris analízissel igazoltak homozigóta kóroki

mutációt egészséges egyéneknél (1). Ezek alapján függetlenül az öröklődés módjától, a penetrancia az adott variáns függvénye (8).

2.1.1. Penetranciát befolyásoló tényezők AR betegségekben

Mutáció típusa

Az inkomplett penetrancia ritka esetei előfordulhatnak a variánsok hipomorf jellege miatt, ami lehetővé teszi a fenotípus mentesítését, ahogyan az az R117H *CFTR* allél okozta cisztás fibrózisban (CF) leírt. Ez a misszensz variáns homozigóta, illetve összetett heterozigóta formában is a CF enyhe fenotípusát eredményezi. Ez a hatás annyira enyhe, hogy az R117H variáns bizonyos variánsokkal társulva, az esetek többségében nem okoz betegséget (9).

A familiáris mediterrán lázért felelős *MEFV* génben is azonosítottak összetett heterozigóta variánst hordozó egészséges egyéneket (10).

Epigenetikai faktorok

Az epigenetikai szabályozás olyan környezeti hatások összessége, melyek a DNS nukleotidsorrendjét nem változtatják meg, viszont olyan változásokat idézhetnek elő (DNS metiláció, hiszton acetiláció és metiláció, mikro RNS hatások), melyek befolyásolhatják a génexpressziót és a penetranciát (8).

Gén-környezet kölcsönhatás

A környezeti hatások (nem, étrend, gyógyszerek, alkoholfogyasztás, metabolikus szindróma, fizikai aktivitás) befolyásolhatják a betegség súlyosságát. Jól ismert, hogy fenilketonuriában (PKU) a betegség súlyossága jelentősen befolyásolható fenilalanin szegény diéta szigorú betartásával (1, 11).

Digenikus és oligogénes öröklődés

Inkomplett penetrancia lehet a betegség oligogén jellegének következménye. Ez esetben több gén kóroki variánsa áll a betegség hátterében.

Oligogénes öröklődést igazoltak nephronophthises (NPHP) betegek esetében. A NPHP egy autoszomális recesszív módon öröklődő cisztás vesebetegség, mely veseelégtelenséghez vezet. Eddig több mint 20 NPHP-ért felelős gént azonosítottak (12). Egy vizsgálat során 44 családban a két mutációt két különböző *NPHP* génben azonosították (*NPHP1-4*). Továbbá, ugyancsak digenikus öröklődést mutattak ki egy betegnél, aki egy mutációt hordoz az *NPHP2*-ben, és egy második mutációt az *NPHP4* génben (13).

Digenikus öröklődésről akkor beszélünk, ha két kóroki mutáció két különböző génben vezet az adott fenotípus megnyilvánulásához. A Bardet-Biedl szindróma (BBS) egy autoszomális recesszív betegség, amelyet retina dystrophia, elhízás, posztaxiális polydactylia, tanulási zavarok, veseérintettség, valamint férfi hypogenitalizmus jellemez. Egy tanulmány során komplex digenikus, triallelikus öröklődésre utaló bizonyítékokat írtak le: 163 BBS beteg vizsgálata során négy családban két génben is találtak kóroki mutációt (*BBS2* génben két mutációt, *BBS6* génben egyet). Két családban két egészséges egyénnél találtak a *BBS2* génben két kóroki mutációt, *BBS6* mutáció nélkül (14).

Ezen eredmények arra utalnak, hogy a BBS nem minden esetben monogénes, a fenotípus kialakulásához egyes esetekben három allél mutációja vezet, két különböző génben (14).

Módosító gének

A monogénes betegségekben alapvetően egy gén felelős a betegség kialakulásáért. Bizonyos betegségekben kimutattak már egy vagy több gént, melyek befolyásolják a betegség fenotípusát (8).

A CF-ban a mutációk ismerete alapján nem lehet pontosan megjósolni a betegség fenotípusát. Eddig legalább hét módosító gént írtak le, melyek befolyásolják a betegség lefolyását (15). Ilyen módosító gén az *MBL2*, mely a mannóz kötő lektint kódolja. Egy vizsgálat során azt igazolták, hogy azoknak a cisztás fibrózisos betegeknek a tüdőfunkciója jelentősen csökkent volt, akiknél az *MBL2* gén mindkét allélja kóros volt. A heterozigóták esetében nem figyelték meg ezt a hatást. A pontos mechanizmus még nem világos, de a jelenlegi adatok alapján úgy tűnik, hogy a mannóz-kötő lektin, amely jelen van a bronchoalveoláris felszíni folyadékban gyulladásgátló hatást fejthet ki, illetve a kóros fehérje érzékenyebb lehet a mátrix metalloproteinázok lebontására, amelyek bőségesen jelen vannak a bronchoalveoláris folyadékban (16).

További, a CF fenotípusát befolyásoló módosító variánsok a *TGFBI* gén polimorfizmusai, az Alfa1-antitripszin kódoló *SERPINI* (17), a Beta-2 adrenerg receptort kódoló *ADRB2* gén polimorfizmusai, melyek mind heterozigóta, mind homozigóta formában súlyosbítják a CF lefolyását (18).

Interallelikus interakciók

Az inkomplett penetrancia hátterében interallelikus interakciók is állhatnak.

Az interallelikus interakciókról több évtizede ismert, hogy befolyásolják a fajok fenotípusát (19), valamint humán enzimopátiák kapcsán is leírták már.

Interallelikus interakcióra utaló bizonyítékot a fenilketonuriaért felelős *PAH* gén variánsainak vizsgálata esetében figyeltek meg. A tanulmányban *in vivo* és *in vitro* kísérleteket végeztek majom vesesejteken. Koexpresszió során a *PAH* bizonyos mutációinak enzimaktivitását vizsgálták összetett heterozigóta mutációk esetében. Az enzimaktivitást összehasonlították a várt enzimaktivitással és fenotípussal, melyek a PKU adatbázisában érhetőek el (BIOPKU: <http://www.biopku.org>).

A vizsgálat eredménye alapján arra következtettek, hogy az interallelikus interakció PKU esetében is lehet negatív és pozitív. Negatív, ahol a várthoz képest alacsonyabb enzimaktivitást tapasztaltak, pozitív, ahol magasabbat. A vizsgálat alapján felmerül, hogy az interallelikus interakció a variánsok hidrofób - hidrofil tulajdonságának megváltozásához kötött. Ezenkívül a két variáns lokalizációjával is összefügghet: amikor a regulációs doménban levő mutációk a katalitikus doménban levő mutációval társultak,

negatív interallelikus interakciót eredményezett, melyet a súlyos konformáció változásnak tulajdonítottak (20).

Továbbá két másik enzimopátiáért felelős gén variánsainak esetében igazoltak interallelikus interakciót. Ilyen a galaktozémia (*GALT*) (21) és az argininoszukcinát-liáz deficiencia (*ASL*) (22, 23).

2.2. Szteroid-rezisztens nephrosis szindróma

A nephrosis szindrómát (NS) a glomerulus szűrőfunkciójának károsodása következtében kialakuló súlyos proteinuria, hypalbuminaemia, ödéma és hiperlipidaemia jellemzi (24).

Az idiopátiás NS osztályozható szteroid terápiára adott válasz, szövettan, illetve genetikai mutációk alapján. Terápiára adott válasz alapján elkülönítünk szteroid-szenzitív és -rezisztens formát. A primer, idiopátiás forma leggyakrabban minimális szövettani elváltozással járó „minimal change” nephrosis szindróma (MCNS). Ez általában szteroid-szenzitív, és jó prognózissal jár (25).

A szteroid-rezisztens nephrosis szindróma (SRNS) egyik típusos szövettani formája a fokális szegmentális glomerulosclerosis (FSGS). Prognózisa rosszabb, mivel gyakran 5 éven belül végstádiumú veseelégtelenséghez (ESRD) vezet (25). Az SRNS a gyermekkori idiopátiás NS 10%-áért, felnőttkorban a 40%-áért felelős (26).

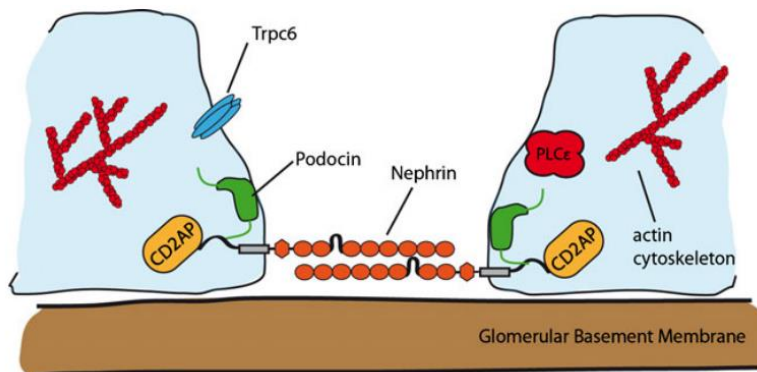
Kóreredit alapján elkülönítünk immunológiai és monogénes SRNS formát. Immunológiai eredet esetén a beteg immunszuppresszív terápia hatására remisszióba kerülhet, illetve a betegség vesetranszplantációt követően kiújulhat. Amennyiben monogénes ok áll az SRNS hátterében az immunszuppresszív terápia hatástalan, illetve a betegség nem újul ki vesetranszplantációt követően. A két forma elkülönítése fontos ezért a terápia adekvát alkalmazása és a prognózis megítélése szempontjából (27).

Az SRNS monogénes formáiban eddig több, mint 50 gén mutációit azonosították, amelyek többsége podocita-fehérjéket kódol. Ezen gének a glomeruláris résmembránt alkotó fehérjéket (*NPHS1*, *NPHS2*, *CD2AP*), a podociták szerkezeti elemeit (*MYO1E*, *ARHGDI1*, *TTC21B*, *WDR73*), illetve bazális membránhoz való adhéziójáért (*LAMB2*, *ITGA3*, *LAMB3*, *ITGB4*, *CD151*), differenciációjáért és megfelelő funkciójáért (*WT1*, *PLCE1*, *PTRO*) felelős fehérjéket kódolnak (28, 29).

A szignifikáns proteinuriával járó kórképek hátterében számos patomechanizmus állhat, melyek mind a glomeruláris filtrációs barrier (GFB) egy rétegének – a fendsztrált endothelnek, a glomeruláris bazális membránnak vagy a podocitáknak, - a károsodására vezethetőek vissza.

A GFB rétegei közül legfontosabbként a podocitákat tartják számon. A podociták olyan módosult epitheliális sejtek, melyek különleges lábnyúlványokat formáló sejtvázszerkezettel rendelkeznek és számos jelátviteli útvonalat tartalmaznak, ezek mind feltételei a megfelelő podocita működésnek. Az egészséges podociták fenntartásában, működésének szabályozásában jelentős szerepe van a glomeruláris résmembrán fehérjéinek (1. ábra), melyek egy jelátviteli központ részei (30).

A glomeruláris résmembrán normális körülmények között a víz, ionok és az albuminnál kisebb méretű fehérjék számára járható át. Nephrosis szindrómában a podociták szerkezete károsodik, így a podocita lábnyúlványok, és így maguk a podocita sejtek is leválnak a glomerulusról és nem képesek fenntartani a glomeruláris résmembránt, ami masszív proteinuriához vezet. Bár a podocitákban végbemenő celluláris és molekuláris változások kevéssé ismertek nephrosis szindrómában, a citoszkeletális fehérjék központi szerepe valószínű a kórlefoiyásban (31).



1. ábra: glomeruláris résmembrán alkotóelemei (32)

2.3. Podocin

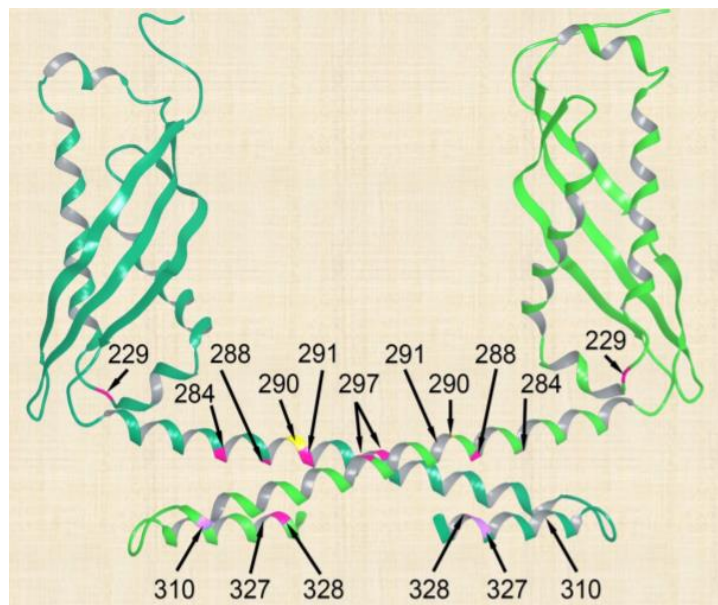
Az *NPHS2* gén által kódolt fehérje a podocin, mely kizárólag a podocitákban expresszálódik, innen is kapva a nevét (33). Egy 383 aminosavból álló, 42 kDa nagyságú fehérje, mely oligomer formában helyezkedik el a podocita résmembránjában, amely a glomeruláris szűrő legfontosabb alkotóeleme. A résmembrán további alkotóelemei a nephrin, a CD2AP, a Neph1 és az alfa-actinin-4 fehérjék (1. ábra) (34). A podocin egy integráns membránfehérje, mely a sztomatín fehérje családba tartozik. Hajtűhöz hasonló

intramembrán hurokkal és intracelluláris N- és C-terminális véggel rendelkezik, mely interakciós partnerek révén a citoskeletonnal teremt kapcsolatot, valamint a többi podocin molekulával homooligomereket képez (31, 33).

Az intramembrán szakasz a sejtmembrán lipidtutajaiban helyezkedik el, ahol kapcsolódik a podociták közötti résmembránt alkotó nephrin fehérjéhez (31).

A podocinnal interakcióba lépő további fehérjék a CD2AP, NOX2, CIN85 és a TRPC6. A podocin funkciója a nephrin kihorgonyzása, de részleteiben nem teljesen ismert. Korábbi vizsgálatok alapján azt feltételezték, hogy az oligomerizáció létrejöttében az N- és C terminális részek is szerepet játszanak (35).

A podocin fehérjeszerkezeti modellekben, amit a nagyfokú homológiát mutató *Pyrococcus hirokoshii* sztomatin kristályszerkezete alapján modelleztek (36), a vad homodimerek C-terminális részét vizsgálva (aminosavak 161-383) azt találták, hogy egy globuláris feji részből (PHB domén) (aminosavak 161-282) és egy helikális C-terminális régióból (CTT) áll, amit két helikális rész alkot (H1: 282-313, H2: 317-330) (2. ábra). További vizsgálatok során bebizonyosodott, hogy a C-terminális régió végső szakasza egy további hélixet tartalmaz (H3: 344-350), a végén egy rendezetlen szerkezetű résszel (351-383) (37).



2. ábra: Podocin homodimer szerkezete (36)

2.4. *NPHS2* gén

A leggyakoribb gyermekkori SRNS-t és felnőttkori FSGS-t okozó gén a podocint kódoló *NPHS2*. Az *NPHS2* gén mutációi 15-39%-ban állnak az SRNS congenitalis és infantilis formájának hátterében, illetve 13-14%-ban a gyermekkori forma mögött (26, 38).

Az *NPHS2* asszociált nephrosis szindróma autoszomális recesszív módon öröklődik. Ez idáig 126 variáns ismert (39). Az érintett egyéneknél általában hatéves koruk előtt alakul ki az SRNS, mely az élet első évtizedében végstádiumú veseelégtelenséghez vezet. A betegség lefolyásának súlyossága a kóroki mutációktól függ. Bizonyos mutációk, például a kóroki allélok 32-44%-át adó R138Q, a podocin endoplazmás retikulumban való retencióját okozva korai kezdetű nephrosishoz vezetnek (33).

2002-ben azonosították az *NPHS2* R229Q polimorfizmusát, melyre az európai emberek 3,6%-a heterozigóta, és adolescens vagy felnőttkorban kezdődő FSGS-hez vezethet. *In vitro* végzett vizsgálatokkal kimutatták, hogy az R229Q mutáns fehérje kevésbé kötődik a nephrin fehérjéhez, enyhe funkcionális károsodást okozva ezáltal (40).

Később, 2013-ban olyan személyekben azonosították az R229Q variánst homozigóta formában, akik nem voltak érintettek: igazolva azt, hogy az R229Q homozigóta formában nem vezet nephrosis szindróma kialakulásához (41), ennek ellenére a R229Q magas allélgyakorisága a késői SRNS -ben (5,7-7,5%) utal a variáns patogenitására (42).

2014-ben kutatócsoportunk francia munkatársakkal való kollaborációban igazolta, hogy az R229Q polimorfizmus csak meghatározott 3' mutációkkal társulva okoz betegségeket (36).

2.5. Domináns negatív hatás és interallélikus komplementáció

Kutatásunk alapjául kutatócsoportunk korábbi munkája szolgált, mely során igazolódott a podocint kódoló *NPHS2* példáján, hogy az AR módon öröklődő betegségekben egy allél patogénitása lehet a transz-asszociált allélon levő mutáció függvénye. Így az egyik allélról átíródó fehérje kifejthet negatív hatást a másik allélról átíródó fehérjére.

Ezt az interallelikus interakciót kutatócsoportunk az R229Q podocin variánsa kapcsán igazolta (36).

A kutatás alapját az adta, hogy a National Heart, Lung and Blood Institute (NHLBI) exome variáns adatbázisában azt figyelték meg, hogy az R229Q polimorfizmus allélfrekvenciája (2,7%) 15-ször gyakoribb, mint az összes ismert *NPHS2* mutáció (mut) kumulatív allélfrekvenciája (0,18%) (17-19, 24).

Ezen eredmények alapján azt várnánk, hogy azoknak a személyeknek köszönhetően, akik összetett heterozigóták az R229Q variánsra és egy más *NPHS2* mutációra ([R229Q];[mut]), a késői kezdetű FSGS 30x gyakrabban forduljon elő, (18, 19, 24) mint a nem R229Q variánst hordozó ([mut];[mut]) genotípus következtében megjelenő SRNS. Ezzel szemben egy francia és más nagy SRNS adatbázisok alapján az [R229Q];[mut] következtében kialakuló FSGS 3,5 - szer ritkábban fordult elő, mint a korai kezdetű SRNS [mut];[mut] (n=247) (25).

Ez a jelentős (100 – szoros) különbség a várt és megfigyelt arányok között azt sugallta, hogy az [R229Q];[mut] nem minden esetben okoz betegséget.

A hipotézis teszteléséhez *NPHS2* mutációk következtében érintett gyermekek egészséges szüleit szűrték az R229Q variánsra. Ők egy *NPHS2* mutációra obligát heterozigóták voltak. A 129 egészséges hordozó szülő közül hat (4,7%), akik 42 és 62 év között voltak, összetett heterozigóták voltak az R229Q mutációra és a következő mutációk egyikére: R138E, R138* (n = 2), R168H, c.534+1G>T, R238S. Ezek az eredmények egyértelműen igazolták az SRNS inkomplett penetranciáját az [R229Q]; [mut] egyénekben (36).

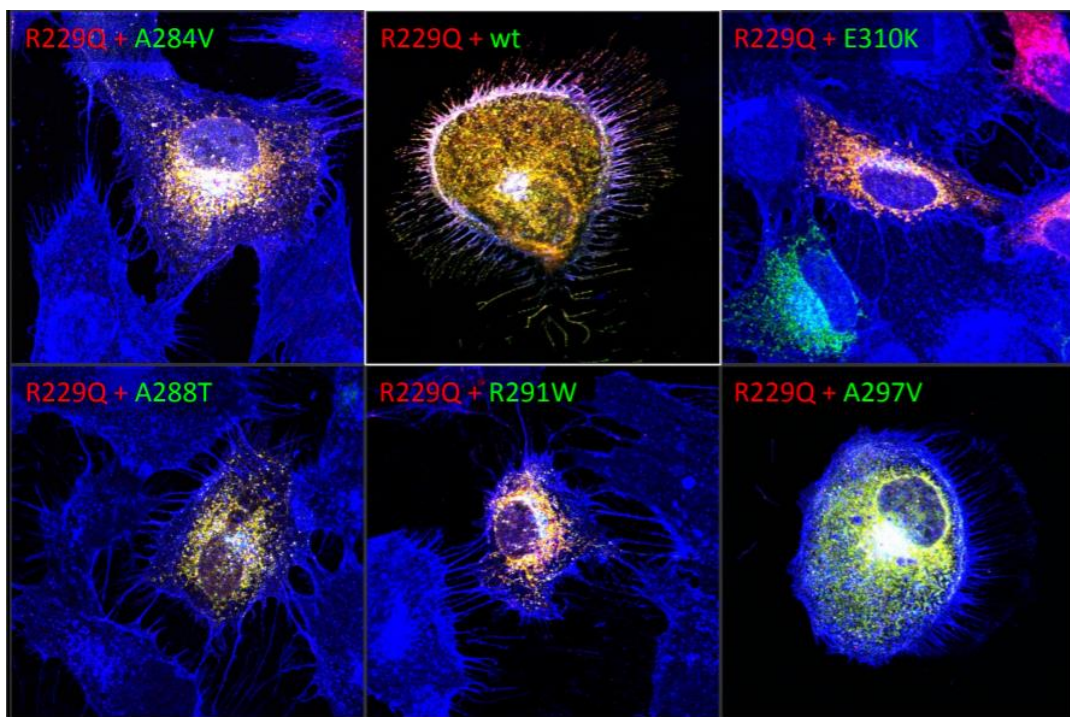
Tovább vizsgálva a társulásokat, azt figyelték meg, hogy az SRNS-ben szenvedő betegekben az R229Q szignifikánsan többször ($p = 1,2 \times 10^{-35}$) csak bizonyos, 7-8. exonokat érintő mutációkhoz társult. A leggyakrabban érintett aminosavak az A284, A288, R291, A297, E310, L327 és a Q328 voltak. Az ezen régiót érintő leggyakoribb

misszensz variáns, a V290M nem társult betegekben az R229Q variánshoz, valószínűtlenné téve a társulásuk patogenitását. Megtalálták ezzel összhangban az [R229Q];[V290M] társulást egy család két egészséges 40 év feletti családtagjában (36). Mivel az összes társuló mutáció a podocin C-terminális részén helyezkedett el, feltételezték ezek R229Q podocinnal szemben kifejtett domináns negatív hatását.

Sejtkultúrában végzett fehérjelokalizációs kísérletek és *in silico* számított fehérjeszerkezeti modellek alapján igazolták, hogy a C-terminális szakasz patogén mutációi domináns negatív hatást gyakorolnak az R229Q podocinra.

A különböző podocin konstrukciókat együtt expresszálva humán podocita sejtvonalakon igazolták, hogy a patogén podocin variánsokkal nem jut ki az R229Q podocin a sejtmembránba, így a heterooligomerek intracellulárisan rekednek.

Klinikai adatokkal tökéletesen összhangban, az R229Q podocin megfelelően lokalizálódott a plazmamembránban, amikor vad típusú podocinnal (3. ábra), R238S podocinnal vagy V290M podocinnal együtt expresszálták, de visszamaradt a citoplazmában, amikor A284V (3. ábra), A288T, R291W, A297V (3. ábra) vagy E310L podocinnal együtt expresszálták.



3. ábra: Az hemagglutinin (HA)-jelöléssel ellátott R229Q podocin piros, az együtt expresszált GFP-vel jelölt podocin fehérjék zöld színnel jelennek meg a fenti ábrán. A plazmamembrán búzacsíra agglutininnel (WGA) van jelölve, és kék színű.

A vad (wt) és az R229Q podocin koexpresszáva a plazmamembránban lokalizálódik, fehér szint mutatva. Az R229Q és az A284V, E310K, A288T, R291W, A297V társulások esetén a podocin nem jut ki a sejtmembránba (36).

A vad homodimerek C-terminális (aminosavak 161-383) részét vizsgálva fehérjeszerkezeti modellekben azt találták, hogy a két feji régiót helikális, coiled-coil doméneket alkotó szakaszok (aminosavak 271-372) kötik össze. Az R229Q podocin, a fehérje globuláris feji részében helyezkedik el, és három másik aminosavval (Q233, Q237, D244) képez hidrogénkötést. Az R229Q aminosavcsere következtében ebből két (Q233, Q237) kötés megszakad és két új jön létre (R286, K289), amelyek rigidebbé teszik a globuláris régió és a C-terminális helikális rész kapcsolatát.

Míg a patogén társulások (R229Q-A28V/A297V heterodimerek, A284 homodimerek) szerkezete eltért a nem patogén társulásokétól (pl. vad-R229Q, vad-A284V heterodimerek), a nem patogén társulások szerkezetükben nagymértékben egyeztek a vad-vad podocin homodimerével (36).

A patogén társulások megváltozott heterooligomer szerkezetével magyarázták az R229Q kóros sejten belüli lokalizációját. Tehát a kóros dimerizáció egy formája lehet az R229Q podocinra kifejtett domináns negatív hatásnak.

Korábbi munkánk során, a szerkezeti modellezés alapján feltételezve, hogy a C-terminális hélixek elsődleges szereppel bírnak az oligomerizációban, ezen *in silico* számítást kísérletes módszerrel igazoltuk. Három természetben előforduló (R286Tfs*17, A317Lfs*31, F344Lfs*4) és egy mesterséges (F344*) C-terminálisukon változó mértékben csonkolt podocin variáns kapcsolatát vizsgáltuk az intakt C-terminális régióval rendelkező vad és R229Q mutáns podocinokkal, Förster-féle rezonáns energiaátadás (FRET) hatékonyság mérésével (37).

A munka során azt tapasztaltuk, hogy az a podocin variáns, mely C-terminális régióval egyáltalán nem rendelkezik, azaz mindhárom C-terminális helikális régiója hiányzik (R286Tfs*17), nem ad FRET jelet sem a vad, sem az R229Q podocinnal.

Ez az eredmény összhangban volt az expressziós vizsgálatokkal: az R286Tfs*17 podocin a koexpresszált podocin variánsokkal nem kolokalizált, és azok lokalizációját nem befolyásolta. Az eredmények alapján arra következtethetünk, hogy az oligomerizáció kizárólag a CTT révén valósul meg, az irodalmi adatokkal szemben az N-terminális régió nem vesz részt benne (35). Az R286Tfs*17 podocin ugyanakkor egy membrán-lokalizált, hypomorph variáns, igazolva, hogy az oligomerizáció nem feltétele a membránlokalizációnak, és CTT hiányában is, a membránba kijutó csonka podocin képes funkciójának egy részét ellátni. További három trunkáns mutáció (A317Lfs*31, F344Lfs*4, F344*) vizsgálata alapján igazoltuk, hogy a H1 régió megléte (A317Lfs*31) már önmagában elegendő az oligomerizációhoz és a H2, H2-H3 összekötő régió, illetve a teljes CTT régió megléte jelentősen befolyásolta a mért FRET hatékonyságot, melyből a megváltozott oligomerizációra következtethetünk (37).

Így tehát mind a családvizsgálatok és a populációgenetikai számítások határozottan alátámasztják bizonyos podocin mutánsok, R229Q podocinra gyakorolt domináns negatív hatását, illetve az *in silico* és *in vitro* vizsgálatok segítenek megérteni az ennek hátterében álló mechanizmust.

Ennek a jelenségnek közvetlen klinikai következményei vannak. A korábbi tanácsadási gyakorlatokkal ellentétben, amennyiben egy házaspár egy tagja az 1–6 exonban található *NPHS2* mutációt hordozza, a másik tag pedig az R229Q mutációt, az esetükben nem áll fenn az a veszély, hogy beteg gyermekük születik (36).

Ezen eredmények alapján felmerült annak lehetősége, hogy más autoszomális recesszív betegségek is öröklődhetnek hasonló mutációfüggő módon, különösen akkor, ha a kódolt fehérje oligomereket képez.

4. Célkitűzés

- A bevezetésben említett R229Q példájához hasonlóan inkomplett penetráns variánsok azonosítása gyakori autoszomális recesszív betegségekben, populációgenetikai módszerrel. Az inkomplett penetráns variánsok között az interallelikus interakció-függők azonosítása.
- Cél volt, hogy a számításainkat végző, általunk alkotott algoritmust egy digitalizált, R környezetben írt program segítse. A megírt program lehetőséget nyújt a növekvő szakirodalmi adatok gyors elemzésére.
- Patogén podocin oligomerek megváltozott szerkezetének igazolása experimentálisan.
- Az *NPHS2* R229Q ismeretlen patogenitású társulásainak megítélését segítő klinikai irányelv létrehozása.

5. Módszerek

5.1. Variánsok vizsgálata populációgenetikai módszerrel

Egy variáns patogénitásának, penetranciájának vizsgálatára alkalmas módszerek között szerepelnek a családvizsgálatok, biológiai vizsgálatok, illetve a populációgenetikai módszerek. Célunk volt szisztematikusan vizsgálni a variánsokat, amire a populációgenetikai módszer a legalkalmasabb. A módszer lehetőséget ad az allélgyakoriságok, illetve a geno- és fenotípus-gyakoriságok számszerű meghatározására a populációban. Mivel az irodalomban elérhetőek olyan adatbázisok, melyek kontrollcsoportként, „egészséges” (monogénes betegségben nem szenvedő) populációnak megfelelnek, a módszer alkalmazásához beteg populációt kellett létrehozni, amit a növekvő publikált esetszámok, adatbázisok lehetővé tettek.

5.2. Autoszomális recesszív betegségek kiválasztása

Az AR betegségek listáját az Online Mendelian Inheritance in Man adatbázisából (OMIM) töltöttük le. Összesen 1981 AR betegség szerepelt ebben a listában, 2016 áprilisában.

A populációgenetikai módszer csak gyakori betegségek, illetve gyakori variánsok esetén informatív. Azért, hogy minél több variáns penetranciáját meg tudjuk ítélni, meghatároztunk több frekvencia kritériumot, melyek alapján szűkítettük a vizsgálandó betegségeket:

- 1) A kóroki génben a teljes funkcióvesztést okozó (LOF) mutációk kumulatív allélfrekvenciája 0,02%-nál magasabb a gnomAD nevű adatbázis (Genome Aggregation Database) (43), európai, nem finn populációjában. Ez azért volt fontos, mert később a LOF mutációkhoz viszonyítjuk az összes többi variáns penetranciáját. A gyakoriságuk garancia arra, hogy kellően reprezentáltak legyenek minden európai alpopulációban.
- 2) Génenként legalább 100 különböző mutáció szerepeljen a HGMD adatbázisban (The Human Gene Mutation Database) (43). Az AR betegségekért felelős mutációk jellemzően alapító mutációk meghatározott földrajzi elterjedéssel. A

legalább 100 különböző mutáció (alapító hatás) biztosította, hogy a betegség közel egyenletesen legyen elterjedt az európai alpopulációkban.

- 3) A betegség szakirodalomban leírt prevalenciája legalább 1/200 000 legyen. A gyakori betegségekért felelős variánsok korábban jöttek létre a humán evolúcióban, így jobban elterjedtek a felelős variánsok, kisebb a különbség az alpopulációk között.
- 4) A publikált európai betegszám a szakirodalomban több, mint 100.

A fenti kritériumok alapján 15/1981 gént választottunk ki (*ATP7B* [MIM# 606882], *CAPN3* [MIM# 114240], *CFTR* [MIM# 602421], *CTNS* [MIM# 606272], *DHCR7* [MIM# 602858], *GAA* [MIM# 606800], *GALNS*, [MIM# 612222], *IDUA* [MIM# 252800], *MUT* [MIM# 211000], *NPHS1* [MIM# 602716], *NPHS2* [MIM# 602088], *PAH* [MIM# 6012349], *PKHD1* [MIM# 606702], *PMM2* [MIM# 601785], *SLC26A4* [MIM# 605646]). További két gént (*ASL* [MIM# 608310], *GALT* [MIM# 606999]) a szakirodalomban publikált *in vitro* kísérletek alapján leírt interallelikus interakcióra tekintettel választottuk be (21, 23). Így összesen 17 gén variánsait vizsgáltuk (1. táblázat).

1. táblázat: A vizsgált 17 autoszomális recesszív betegség, a kóroki gének és a hozzá tartozó MIM szám az OMIM adatbázisnak megfelelően, illetve referenciaszám a Human Reference Assembly GRCh37.p13 alapján (44).

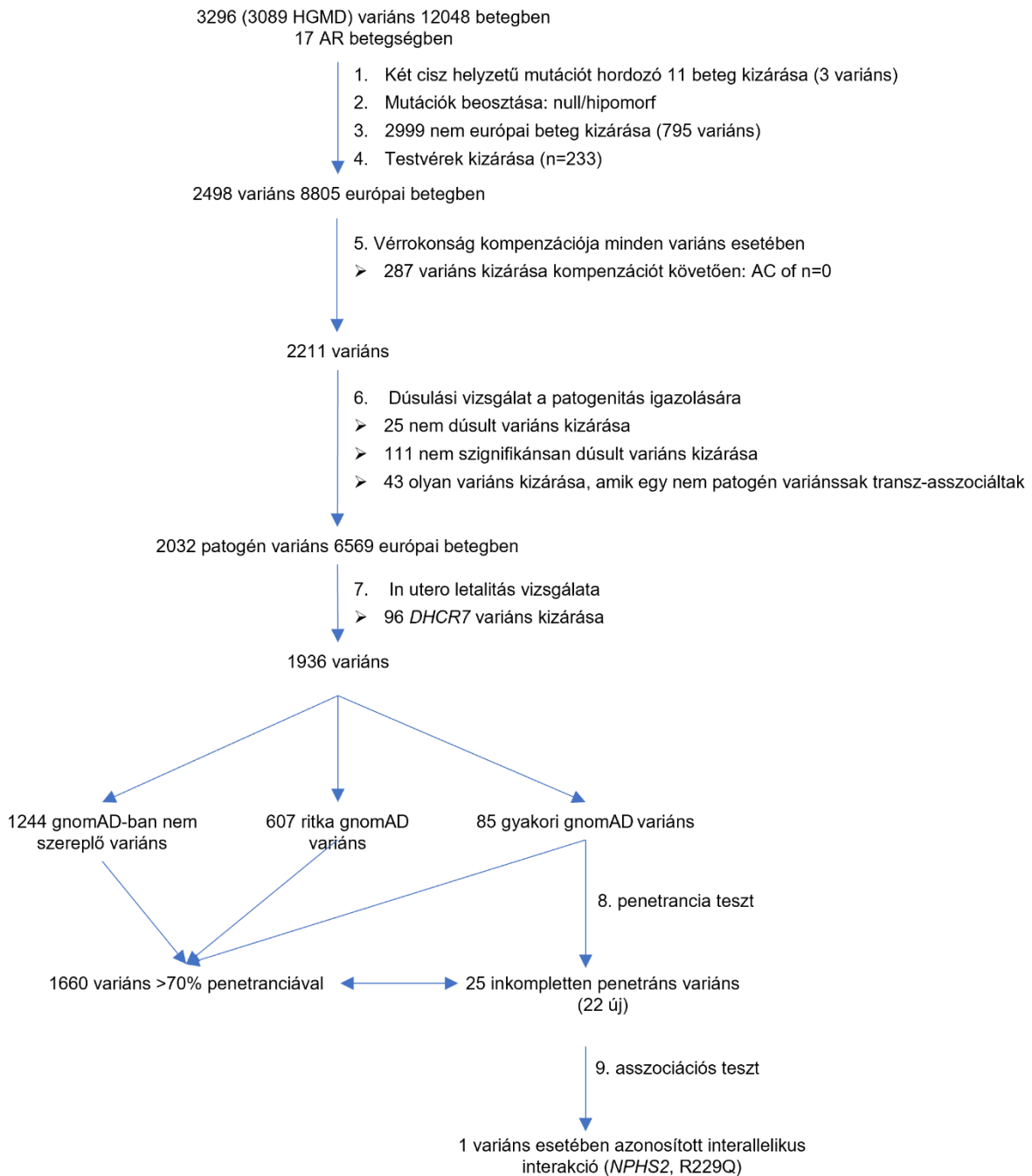
Gén	Betegség	MIM szám	Transzkriptum
<i>ASL</i>	Arginin aciduria	# 207900	ENST00000304874.9
<i>ATP7B</i>	Wilson kór	# 277900	ENST00000242839.4
<i>CAPN3</i>	Juvenilis calpainopathia (LGMD 2A)	# 253600	ENST00000397163.3
<i>CFTR</i>	Cisztás fibrózis	# 219700	ENST00000003084.6
<i>CTNS</i>	Cisztinózis	# 219800	ENST00000381870.3
<i>DHCR7</i>	Smith-Lemli-Opitz szindróma	# 270400	ENST00000355527.3
<i>GAA</i>	Pompe kór (II-es típusú Glycogenosis)	# 232300	ENST00000302262.3
<i>GALNS</i>	Morquio-szindróma (IV-es típusú Mucopolysaccharidosis)	# 253000	ENST00000268695.5
<i>GALT</i>	Galaktozémia	# 230400	ENST00000378842.3
<i>IDUA</i>	Hurler- Sheie szindróma (I-es típusú Mucopolysaccharidosis)	# 607014	ENST00000247933.4
<i>MUT</i>	Methylmalonsav acidémia	# 251000	ENST00000274813.3
<i>NPHS1</i>	I-es típusú nephrosis szindróma	# 256300	ENST00000378910.5
<i>NPHS2</i>	II-es típusú nephrosis szindróma	# 600995	ENST00000367615.4
<i>PAH</i>	Fenilketonuria	# 261600	ENST00000553106.1
<i>PKHD1</i>	Autoszomális recesszív policisztás vesebetegség	# 263200	ENST00000371117.3
<i>PMM2</i>	I-es típusú veleszületett glikolizációs zavar	# 212067	ENST00000268261.4
<i>SLC26A4</i>	Pendred szindróma (veleszületett siketség és golyva)	# 600791	ENST00000265715.3

5.3. Beteg adatok összegyűjtése

A betegek adatainak kigyűjtését manuálisan végeztük a HGMD adatbázisban idézett szakirodalomból, a Google-Scholar, illetve a PubMed keresők segítségével. Azokat a kohorsz tanulmányokat vettük figyelembe, amelyekben a gén teljes kódoló szakaszát szekvenálták (n=341), és azon betegek adatait gyűjtöttük, akik mindkét allélon hordoztak kórokinak leírt variánsokat. Fokozott figyelmet fordítottunk arra, hogy kizárjuk annak esélyét, hogy ugyanaz a beteg kétszer szerepeljen a táblázatunkban. Kigyűjtöttük a betegek azonosítószámát, nemét, életkorát, az etnikai hovatartozást, a genotípust, valamint a fenotípus súlyosságát is (2. táblázat). Ha két beteg azonos központból, azonos mutációval, azonos etnikummal rendelkezett csak egyszer vettük figyelembe. A betegek etnikai hovatartozását a gnomAD adatbázisának megfelelően csoportosítottuk, így például az európai csoportot, európai nem finn és finn csoportra tagoltuk. Így 438 publikációból 12048 beteg adatait gyűjtöttük ki, akik összesen 3296 variánst hordoztak (4. ábra).

2. táblázat: Az *NPHS1* gén mutációit hordozó betegek összefoglaló táblázatának részlete: a táblázatban szerepel a beteg azonosítója (patients), az adott tanulmány első szerzője, illetve publikáció dátuma (Citation), a beteg neme (gender), gnomAD szerinti beosztás (ethnicity. Exac), az ország ahonnan a beteg származik (ethnicity), egyik, illetve másik allélon levő mutációk nukleotid (nucls1/nucls2) és aminosav szekvenciája (Amino_acid_1st/2nd), a betegsége kezdete (Age of onset), a veseelégtelenség kialakulásáig eltelt idő (Time to ESRD), a veseelégtelenség ideje (Age at ESRD).

no	dupl	patients	Citation	Gender	ethnicity.exac	ethnicity	nucls1	exon	Amino acid_1st	nucls2	exon	Amino acid_2nd	of onset(y or	Time to ESRD (y or	Age at ESRD)
57		YK	Berdelli et al, 2007		European (Non-Finnish	Turkey	c.59C>T		P20L	c.59C>T		P20L			
58		YSY	Berdelli et al, 2007		European (Non-Finnish	Turkey	c.344T>C		M115T	c.344T>C		M115T		No	
59	1		bertelli2003	M	European (Non-Finnish	Italy	c.413G>A		R138Q	c.413G>A		R138Q	11 m		84 m
60	2		bertelli2003	M	European (Non-Finnish	Italy	c.413G>A		R138Q	c.413G>A		R138Q	23 m		48 m
61	3		bertelli2003	M	European (Non-Finnish	Italy	c.413G>A	3	Arg138Gln	c.855_856delAA			12 m		136 m
62	4		bertelli2003	F	European (Non-Finnish	Italy	c.413G>A		R138Q	c.412C>T		R138X	13 m		57 m
63	5		bertelli2003	M	European (Non-Finnish	Italy	c.419delG		Gly140Aspfs*4:c.419delG		Gly140Aspfs*41	1 m		156 m	
64	6		bertelli2003	M	European (Non-Finnish	Italy	c.419delG		Gly140Aspfs*4:c.419delG		Gly140Aspfs*41	18 m		146 m	
65	7		bertelli2003	M	European (Non-Finnish	Italy	c.419delG		Gly140Aspfs*4:c.419delG		Gly140Aspfs*41	19 m		108 m	
66	8		bertelli2003	F	European (Non-Finnish	Italy	c.419delG		Gly140Aspfs*4:c.419delG		Gly140Aspfs*41	5 m		65 m	
67	9		bertelli2003	M	European (Non-Finnish	Italy	c.538G>A		V180M	c.467dupT		Leu156Phefs*11	96 m		192 m
68	E13		Buscher et al,	M	European (Non-Finnish	Germany	c.412C>T	3	Arg138*	c.413G>A	3	Arg138Gln	0	199 m	199 m
69	E17		Buscher et al,	F	European (Non-Finnish	Germany	c.503G>A	4	Arg168His	c.503G>A	4	Arg168His	2.8 y		
70	E22		Buscher et al,	M	European (Non-Finnish	Germany	c.413G>A	3	Arg138Gln	c.419delG	3	Gly140Aspfs*41	3.5 y		
71	E29		Buscher et al,	M	European (Non-Finnish	Germany	c.467dupT	4	Leu156Phe*11	c.467dupT	4	Leu156Phe*11	0.9 y	104 m	115 m
72	E41		Buscher et al,	M	European (Non-Finnish	Germany	c.413G>A	3	Arg138Gln	c.413G>A	3	Arg138Gln	0.7 y	103 m	112 m
73	E43		Buscher et al,	F	European (Non-Finnish	Germany	c.413G>A	3	Arg138Gln	c.413G>A	3	Arg138Gln	0.8 y	65 m	75 m
74	E9		Buscher et al,	M	European (Non-Finnish	Germany	c.413G>A	3	Arg138Gln	c.413G>A	3	Arg138Gln	0	82 m	82 m
75	MS16		Buscher et al,	F	European (Non-Finnish	Germany	c.413G>A	3	Arg138Gln	c.868G>A	7	Val290Met	11 y		
76	MS19		Buscher et al,	M	European (Non-Finnish	Germany	c.413G>A	3	Arg138Gln	c.413G>A	3	Arg138Gln	0.6 y		



4. ábra Az inkompletten penetráns variánsok azonosítására szolgáló algoritmus (44).

5.4. Mutációk nevezéktanának egyesítése

A kigyűjtött adatok feldolgozását egy R programozási nyelven általunk megírt algoritmus segítette, ezért különösen fontos volt a mutációk egyesítése, javítása.

A mutációkat a megfelelő referenciaszám (Human Reference Assembly GRCh37.p13) alapján (1. táblázat), az Ensembl (45) segítségével ellenőriztük. A nomenklatúrát a Human Genome Variation Society (HGVS) hivatalos irányelvei alapján jelöltük (46) illetve a Mutalyzer Name Checker (47) program segítségével javítottuk.

5.5. Cisz variánsok

Azokat a betegeket, akik három variánst hordoztak (kettőt ciszben) manuálisan vizsgáltuk ($n = 47/12048$ [0,39%]). A legtöbb esetben ($n=36/47$) a kóroki mutáció az allélfrekvencia, illetve a fehérje evolúciós konzerváltsága alapján eldönthető volt.

Ez esetben a másik cisz helyzetben levő, nem patogén variáns a vizsgálandó variánsok közül kizárásra került. Azokat a betegeket, akiknél a ciszben levő mindkét variáns patogén volt, vagy patogenitásuk nem volt megítélhető, kizárásra kerültek ($n = 11/12048$ [0,091%]) (4. ábra).

5.6. A variánsok osztályozása

A mutációkat három kritérium szerint osztályoztuk: 1) LOF/misszensz, 2) fehérjedomén-lokalizáció, 3) null/hipomorf hatás.

Mind a biológiai predikció (LOF/misszensz), a klinikai (null/hipomorf) és a fehérjedomén-lokalizáció szerinti beosztás az interallelikus-interakciók kiszűrésére és a populációgenetikai módszer validálására szolgált.

5.6.1. Genotípus szerinti beosztás: LOF/misszensz besorolás

A következő variánsokat soroltuk a LOF csoportba: (1) a donor és akceptor konszenzus hasítási helyet érintő mutációkat; (2) a nonszensz mutációkat, kivéve ha a stop kodon az utolsó exonban van; (3) exont érintő nagy deléciókat. Az aminosavcserét okozó mutációk és a kereteltolódást nem okozó deléciók, inszerciók, duplikációk a misszensz csoportba kerültek besorolásra.

5.6.2. Null-hipomorf beosztás

Genotípus szerint null mutációnak ítéltük: (1) a nonszensz mutációk által mediált mRNS lebontást (NMD, nonsense-mediated mRNA decay) (48) okozó mutációkat; (2) azokat a trunkáns/frameshift mutációkat, melyek legalább a fehérje harmadának elvesztését okozzák; (3) a donor (GT) és akceptor (AG) konszenzus hasítási helyet érintő mutációkat, kivéve az utolsó intronban levőket.

Azon variánsokat, melyeket genotípus alapján nem tudtunk besorolni, a társuló fenotípus alapján osztályoztuk. A 17 vizsgált génhez tartozó null és hipomorf státusznak megfelelő fenotípust a 3. táblázat foglalja össze. Ahhoz, hogy fenotípus alapján elkülönítsük a súlyos és az enyhe formát okozó mutációkat, egy határértéket szabtuk meg. Ezt a határértéket a mindkét allélon genotípus szerint null mutációt hordozó betegek diagnózisának, illetve végstádiumának időpontjában betöltött életkor 95-ös percentilise alapján számoltuk (*CAPN3*, *GAA*, *NPHS1*, *NPHS2*). Azon betegségeknél ahol a fenti módszer szerinti határérték nem volt meghatározható (*ASL*, *ATP7B*, *CFTR*, *CTNS*, *DHCR7*, *GALNS*, *GALT*, *IDUA*, *MUT*, *PAH*, *PKHD1*, *PMM2*, *SCL26A4*), a null-hipomorf besorolást szakirodalmi adatok alapján ítéltük meg (3. táblázat).

Amennyiben egy mutáció homozigóta formában, vagy egy genotípus szerint null mutációval összetett heterozigóta formában az esetek több mint 50%-ában enyhe fenotípust eredményezett, a mutációt a hipomorf csoportba soroltuk, máskülönben fenotípus szerint null-nak ítéltük.

3. táblázat: A súlyos és az enyhe fenotípus osztályozása az egyes betegségekben (44).

Gén	Betegség	Súlyos fenotípus (null)	Enyhe fenotípus (hipomorf)
<i>ASL</i>	Arginin aciduria	kezdet \leq 1 élethét (22)	kezdet $>$ 1 élethét (22)
<i>ATP7B</i>	Wilson kór	akut májelégtelenség	nincs akut májelégtelenség
<i>CAPN3</i>	Juvenilis calpainopathia (LGMD 2A)	kezdet \leq 17.4 év (95 th pc) ^a	kezdet $>$ 17.4 év (95 th pc) ^a
<i>CFTR</i>	Cisztás fibrózis	klasszikus kórforma (49)	atipikus kórforma (49)
<i>CTNS</i>	Cisztinózis	1. ESRD \leq 10 év 2. dg \leq 12 hó (50)	1. ESRD $>$ 10 év 2. dg $>$ 12 hó (50)
<i>DHCR7</i>	Smith-Lemli-Opitz szindróma	súlyossági pontszám: \geq 50 (51)	súlyossági pontszám: $<$ 50 (51)
<i>GAA</i>	Pompe kór (II-es típusú Glycogenosis)	dg \leq 4.3 hó (95 th pc) ^a	dg $>$ 4.3 hó (95 th pc) ^a
<i>GALNS</i>	Morquio-szindróma (IV-es típusú Mucopolysaccharidosis)	kezdet \leq 3 év (52)	kezdet $>$ 3 év (52)
<i>GALT</i>	Galaktozémia	klasszikus forma (53)	
<i>IDUA</i>	Hurler- Scheie szindróma (I-es típusú Mucopolysaccharidosis)	Hurler szindróma (54)	Hurler-Scheie szindróma / Scheie szindróma (55)
<i>MUT</i>	Methylmalonsav acidémia	mut0 enzim altípus (56)	mut- enzim altípus (56)
<i>NPHS1</i>	I-es típusú nephrosis szindróma	1. kezdet \leq 3 hó 2. ESRD \leq 7 év (95 pc) ^a	1. kezdet $>$ 3 hó 2. ESRD $>$ 7 év (95 pc) ^a
<i>NPHS2</i>	II-es típusú nephrosis szindróma	ESRD \leq 13 év (95 pc) ^a	ESRD $>$ 13 év (95 pc) ^a
<i>PAH</i>	Fenilketonuria	BH4 hiányos és Phe: \geq 1200 mmol/l (57)	BH4 reszponzív és Phe $<$ 1200 mmol/l (57)
<i>PKHD1</i>	Autoszomális recesszív policisztás vesebetegség	újszülöttkori elhalálozás (58)	1 hónapos túlélés (58)
<i>PMM2</i>	I-es típusú veleszületett glikolizációs zavar	csecsemőkori elhalálozás (59)	1 éves túlélés (59)
<i>SLC26A4</i>	Pendred szindróma (veleszületett siketség és golyva)	mutációk osztályozása a Choi et al. (60) Dai et al. (61) and Soh et al. (62) szakirodalom alapján	

^a súlyos és az enyhe fenotípus közötti határértéket az ESRD kezdete vagy a diagnózis idejének (életkor) 95. percentilise határozta meg, a biallelikus LOF allélt hordozó betegeknél alapján

BH4: tetrahydrobiopterin; dg: diagnózis, ESRD: end-stage renal disease/ végstádiumú veseelégtelenség, hó: hónap; LOF: loss of function, pc: percentilis, Phe: fenilalanin

5.6.3. Fehérjedomén-lokalizáció

A misszensz és a kereteltolódást nem okozó deléciókat és inszerciókat csoportosítottuk aszerint, hogy a kódolt fehérjék melyik funkcionális alegységét érintik. A vizsgált gének által kódolt fehérjék domén szerinti beosztását az 4. táblázatban ismertetem. A *GALT* és az *SLC26A4* által kódolt fehérjékhez nem állt rendelkezésre domén szerinti osztályozás.

4. táblázat: A génekhez tartozó transzkriptum és fehérje-domén beosztás (44).

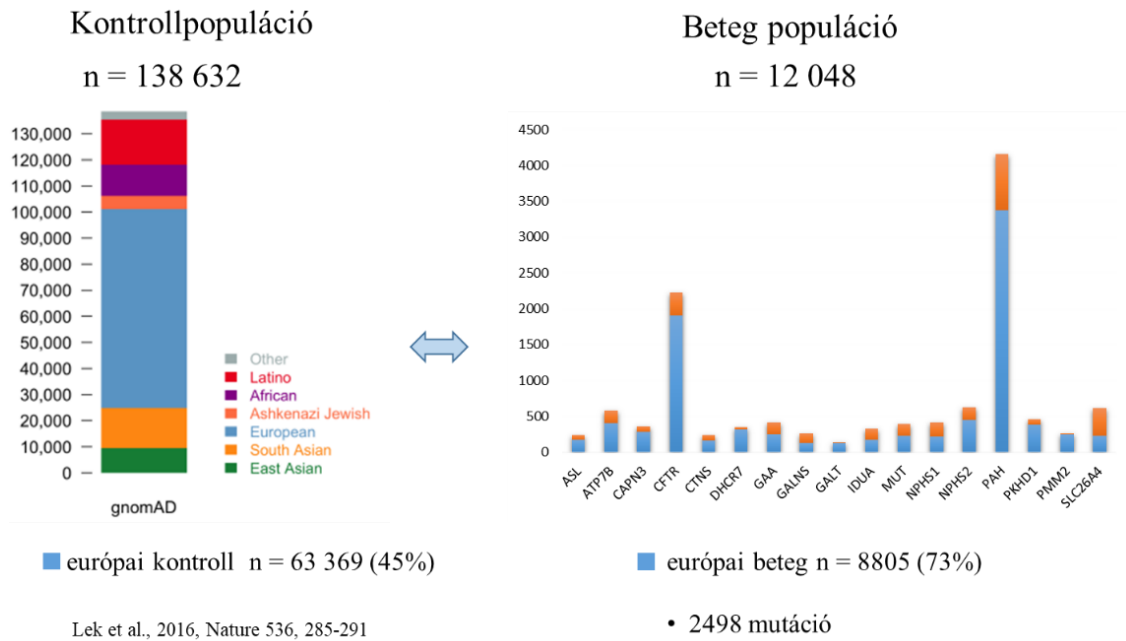
GÉN (referencia)	Transzkriptum (fehérje mérete)	Domén beosztás (AS)
<i>ASL</i> (63)	ENST00000304874.9 (464 AS)	1: 1-110; 2: 111-367; 3: 371-428
<i>ATP7B</i> (64)	ENST00000242839.10 (1465 AS)	1: 1-648; 2: 649-815; 3: 816-910; 4: 911-1020; 5: 1021-1301; 6: 1302-1465
<i>CAPN3</i> (65)	ENST00000397163.8 (821 AS)	1: 1-98; 2: 99-400; 3: 401-653; 4: 654-821
<i>CFTR</i> (66)	ENST0000003084.11 (1480 AS)	1: 1-80; 2: 81-432; 3: 433-584; 4: 590-880; 5: 881-1218; 6: 1219-1382; 7: 1383-1480
<i>CTNS</i> (50, 67)	ENST00000046640.9 (367 AS)	1: 1-128; 2: 128-151; 3: 152-205; 4: 206-228; 5: 229-258; 6: 259-280; 7: 281-318; 8: 319-367
<i>DHCR7</i> (68)	ENST00000355527.8 (475 AS)	1: 1-50; 2: 51-351; 3: 352-410; 4: 411-475
<i>GAA</i> (69)	ENST00000302262.8 (952 AS)	1: 89-135; 2: 136-346; 3: 347-723; 4: 724-818; 5: 819-952
<i>GALNS</i> (70)	ENST00000268695.10 (522 AS)	1: 28-379; 2: 380-481; 3: 484-495
<i>IDUA</i> (71)	ENST00000247933.4 (653 AS)	1: 27-42; 2: 42-396; 3: 397-545; 4: 546-642
<i>MUT</i> (72)	ENST00000274813.3 (750 AS)	1: 1-422; 2: 423-577; 3: 578-750
<i>NPHS1</i> ^a	ENST00000378910.5 (1241 AS)	1: 23-1055; 2: 1056-1076; 3: 1077-1241
<i>NPHS2</i> (73)	ENST00000367615.4 (383 AS)	1: 1-104; 2: 105-125; 3: 125-282; 4: 283-383
<i>PAH</i> (74)	ENST00000553106.1 (452 AS)	1: 1-142; 2: 143-410; 3: 411-452
<i>PKHD1</i> (75)	ENST00000371117.3 (4074 AS)	1: 1-3858; 2: 3859-3879; 3: 3880-4074
<i>PMM2</i> (76)	ENST00000268261.9 (246 AS)	1: 1-90; 2: 95-194; 3: 198-262

^a adatok az Uniprot adatbázisból (<https://www.uniprot.org>); AS: aminosav;

5.7. A beteg populáció európai eredetű betegekre való szűkítése

A beteg populáció többsége európai volt (9010/12048, 75%). A beteg- és a kontrollpopuláció összehasonlítása különböző etnikumok között torzítja az eredményeket. A vizsgálat egyik alappillére, hogy a beteg- és kontrollpopuláció azonos származású legyen.

A kontrollpopulációt a Genome Aggregation Database (gnomAD, <http://gnomAD.broadinstitute.org/>) adatbázisa képezte (43). A szerzők 123 136 monogénes betegségben nem szenvedő személynél végeztek teljes exom-, valamint 15496 embernél teljes genomvizsgálatot. Összesen 138632 ember genotípus adata érhető el az adatbázisban (5. ábra).



5. ábra: A kontroll és a beteg populáció etnikai összetétele: a kontroll és a beteg populáció döntő része európai volt. Ezen kívül latin, afrikai, askenázi zsidó, dél- és észak-ázsiai csoportba osztották az embereket (43).

A gnomAD adatbázisában az európai populációt két csoportra osztották: európai nem finn és finn. Tekintettel arra, hogy a gnomAD adatbázisban és a beteg populációnkban is a kaukázusi személyek voltak a legnagyobb arányban, vizsgálatunkat a beteg- és a kontrollpopulációban egyaránt az európai nem finn személyek által hordozott variánsokra szűkítettük: a gnomAD-ban szereplő 63369 európai nem finn személy variánsait hasonlítottuk össze a beteg populációban található 8805 európai nem finn beteg variánsaival.

Továbbá családonként egy beteget vontunk be (n=233 testvért zártunk ki), így összesen 8805 nem rokon beteg által hordozott 2498 variánst vizsgáltunk (4. ábra).

5.8. Vérokonság hatásának kompenzálása

A vérokonság következményeként a homozigóta betegek száma megnő (77). Önmagában a belterjesség nem befolyásolja egy allél populációs frekvenciáját. Amennyiben azonban két különböző variánst hordozó alpopulációban más a belterjesség mértéke, akkor a beteg populációban eltérő mértékben növeli a belterjesség a variánsok allélfrekvenciáját, és torzítja így a variánsok egymáshoz viszonyított frekvenciáját. Ezen hatás kompenzálása céljából kiszámoltuk variánsenként a beltenyészési együtthatót a következő képlet alapján: (78) $f = 1 - H/2p(1 - p)$, ahol H megfelel a heterozigóta betegek arányának és p megfelel az adott variáns allélfrekvenciájának a beteg populációban. A vérokonság nélküli populációban az adott variáns allélszámát (AC) a következő képlet alapján számoltuk: $AC = 2 * (Bet_{hom} - f * p * N) + Bet_{het}$, ahol Bet_{hom} and Bet_{het} megfelel az adott variánsra homozigóta és heterozigóta betegek számának és N az össz betegszám. Ezzel a kompenzációs módszerrel további 287 variáns került kizárásra, melyek egy alkalommal homozigóta formában fordultak elő (4. ábra).

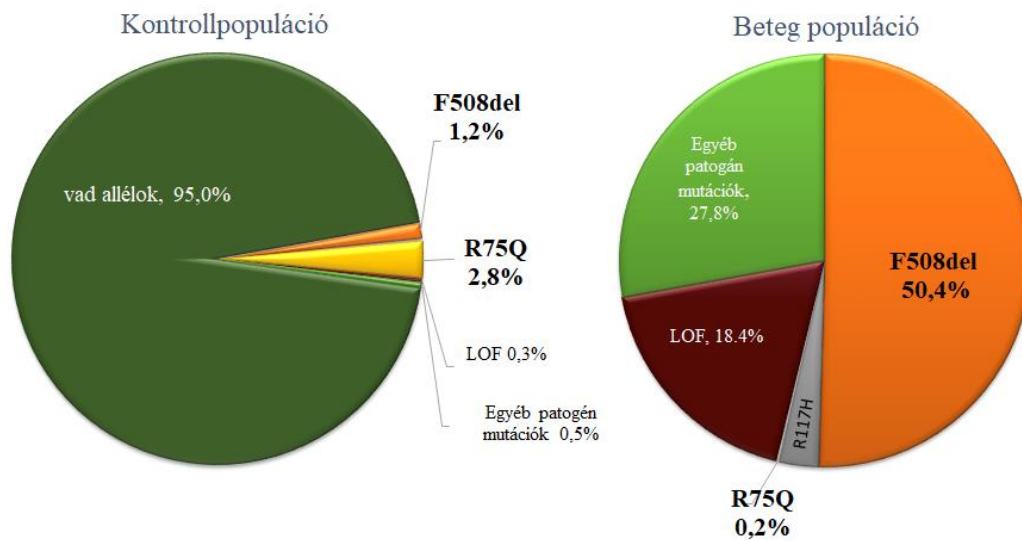
5.9. Patogenitás validálása

Számos olyan variáns szerepel az orvosi irodalomban, amit tévesen patogénként írnak le. Célunk volt kiszűrni a nem patogén mutációkat a HGMD adatbázisban található variánsok közül. A mutációk patogenitását a beteg és az egészséges kontrollpopulációban előforduló allélgyakoriság összevetésével vizsgáltuk.

Patogénnek azt a variánst ítéltük, amelyik szignifikánsan ($p < 0,05$) gyakrabban fordult elő az európai nem-finn beteg populációban, mint az európai nem-finn kontrollcsoportban. Azon variánsokat, amelyek kevesebbszer fordultak elő a beteg populációban, mint a kontrollcsoportban, benignusnak ítéltük, és kizártuk a további vizsgálatból. Továbbá azokat a variánsokat is kizártuk, amelyek nem dúsultak szignifikánsan a beteg populációban.

A benignus és a nem szignifikánsan dúsult variánsok kizárása után azon betegek is kizárásra kerültek, akiknek csak egyik alléljukon maradt így patogén variáns (228/6797). A beteg populációt így a 6569 európai származású, nem finn beteg képezi (4. ábra).

Módszerünk alapelvét a cisztás fibrózis példája jól szemlélteti. A 6. ábrán a *CFTR* mutációinak allélgyakoriságát ábrázoltam a kontroll- és a beteg populációban.



6. ábra: A *CFTR* gén leggyakoribb patogén mutációja az F508del, mely homozigóta formában minden esetben a cisztás fibrózis súlyos formáját eredményezi (79). Ez a mutáció a kontrollpopulációban az allélok 1,2%-án van jelen. A beteg populációban ez a mutáció az allélok 50,4%-án található meg. Az F508del mutáció negyvenszeres dúsulása a beteg populációban jól tükrözi a variáns patogenitását. Azon mutációkat, amelyek a beteg populációban kevesebbszer fordultak elő, mint a kontrollpopulációban, benignusnak ítéltük. Ennek példáját a *CFTR* gén R75Q variánsa szemlélteti. Az R75Q allélfrekvenciája a kontrollpopulációban 2,8%, a beteg populációban 0,2%. Ez mutatja, hogy a beteg populációban való azonosítása csak a véletlen eredménye és nincs szerepe a betegség kialakulásában.

5.10. Intrauterin letalitással járó betegségek kizárása

Azon variánsok, melyek intrauterin letalitást okoznak, a beteg populációban alulreprezentáltak lesznek, torzítva ezzel a LOF-ok arányát a beteg populációban. Célunk volt tehát ezen betegségek kizárása. Ennek megállapítására Fisher-egzakt teszttel összehasonlítottuk a misszensz és a LOF mutációk arányát a beteg- és kontrollpopulációban, illetve Hardy Weinberg tesztet alkalmaztunk a misszensz és LOF mutációkra a beteg populációban. Mindkét tesztnél a szignifikancia szint megítéléséhez Bonferroni korrekciót használtunk: a 0,05-öt elosztottuk a vizsgált gének számával (n=17), így a szignifikancia szint: $p = 0,00294$ volt. Azon betegséget, amelyeknél a tesztek alapján intrauterin letalítás gyanúja felmerült, kizártuk a további vizsgálatból.

5.11. Penetrancia számolás

Kiszámoltuk a beteg populációban szignifikánsan dúsult, tehát populációgenetikai megközelítésben is igazoltan patogén 1936 variáns penetranciáját. Az intrauterin letalitással nem járó AR betegségekben a LOF mutációkat teljesen penetráns variánsnak vélelmeztük. A LOF variánsok beteg populációban megfigyelt dúsulása tükrözi így a teljesen penetráns variánsok dúsulását. Ahhoz, hogy meghatározzuk egy variáns penetranciáját, az adott variáns dúsulását a LOF variánsok dúsulásához viszonyítottuk az alábbi képlet alapján: $\frac{AC_{bet}^V / AC_{gnomAD}^V}{AC_{bet}^{LOF} / AC_{gnomAD}^{LOF}}$, ahol AC_{bet}^V : a vizsgált variáns beteg populációs allélszáma, AC_{gnomAD}^V : a variáns gnomAD-ban előforduló allélszáma, AC_{bet}^{LOF} : LOF mutációk beteg populációs kumulatív allélszáma, AC_{gnomAD}^{LOF} : LOF mutációk gnomAD-ban leírt kumulatív allélszáma.

Mivel módszerünk a LOF variánsokhoz viszonyított arányra épül, ezért minden olyan lehetőséget figyelembe vettünk, ami a LOF variánsok beteg populációs allélfrekvenciáját befolyásolja. Azon egyéneknek, akik összetett heterozigóták egy IP és egy LOF variánsra, nem mindig alakul ki a betegség, így az IP variánsoknak köszönhetően a LOF mutációk száma is csökken a beteg populációban. Megvizsgáltuk ezért, hogy a valódi penetranciát mennyire becsüli jól a V/LOF arány. Az alábbi képlet írja le a V/LOF arány és a penetrancia viszonyát.

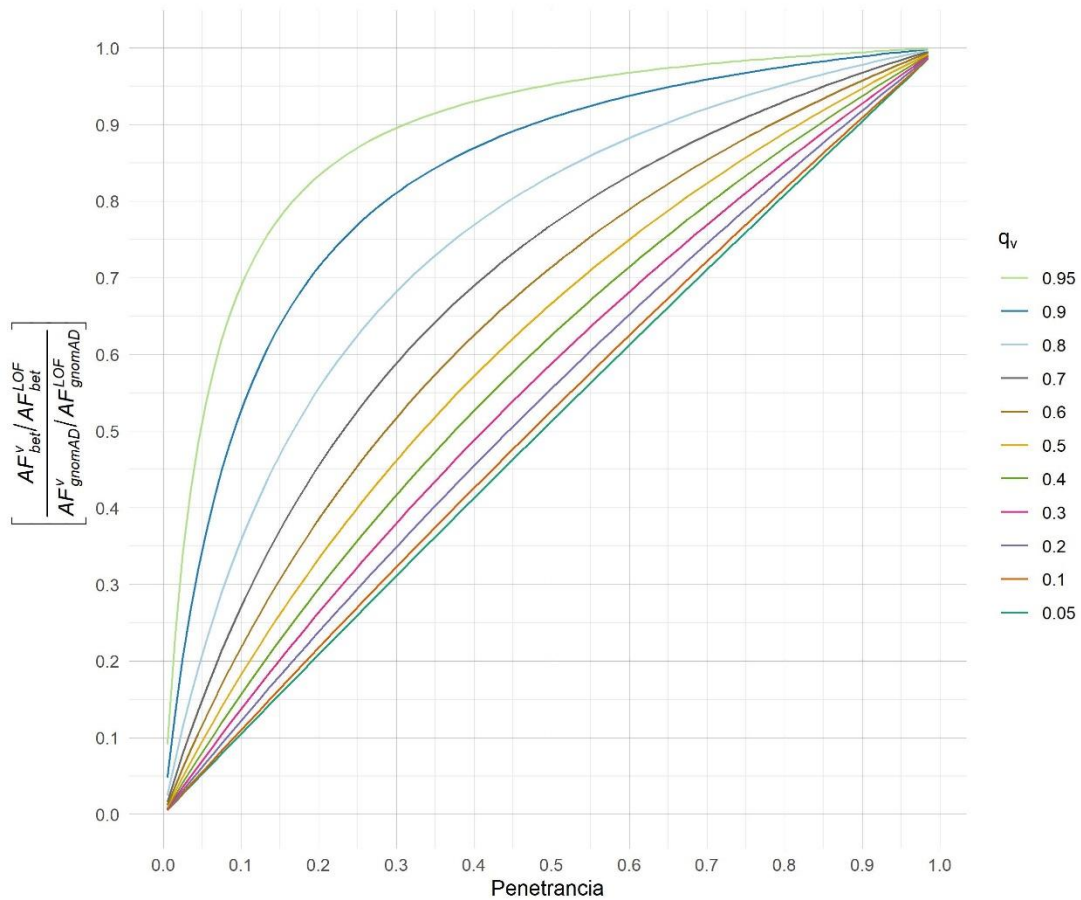
$$\begin{aligned} \frac{V}{LOF} &= \frac{AC_{bet}^v/AC_{bet}^{LOF}}{AC_{gnomAD}^v/AC_{gnomAD}^{LOF}} = \frac{\left(\frac{N_{publ\ bet}(q_v^2 P_v \times 2 + 2q_v P_v q_o P_o + 2q_v P_v q_{LOF})}{N_{publ\ bet}(q_{LOF}^2 \times 2 + 2q_{LOF} q_v P_v + 2q_{LOF} q_o P_o)} \right)}{(q_v/q_{LOF})} \\ &= P_v * \frac{q_v + q_o P_o + q_{LOF}}{q_v P_v + q_o P_o + q_{LOF}} \end{aligned}$$

Az egyenletből leolvasható, hogy a V/LOF hányados akkor korrelál szorosan a penetranciával, ha a q_v alacsony, tehát a vizsgált variáns aránya a többi variánshoz képest alacsony az átlagpopulációban. Mivel $P_v \leq 1$, ezért $q_v + q_o P_o + q_{LOF} \geq q_v P_v + q_o P_o + q_{LOF}$ így $V/LOF \geq P_v$: a V/LOF hányados által becsült penetrancia ezért magasabb, mint a tényleges. Inkomplett penetrancia azonosítása esetén ezért a V/LOF alapján becsült penetranciánál alacsonyabb lehet a tényleges penetrancia akkor, ha a q_v számottevően magasabb.

Ennek szemléltetéséhez az 6. ábrán ábrázoltuk a V/LOF hányados változását a variáns penetranciájának (P_v) függvényében, feltételezve, hogy a többi variáns penetranciája komplett ($P_o=1$). Ekkor a fenti képlet a következőképp rendezhető át (felhasználva, hogy $q_v+q_o+q_{LOF}=1$):

$$P_v * \frac{q_v + q_o + q_{LOF}}{q_v P_v + q_o + q_{LOF}} = \frac{P_v}{q_v P_v + q_o + q_{LOF}} = \frac{P_v}{q_v P_v + q_o + (1 - q_v - q_o)} = \frac{P_v}{q_v(P_v - 1) + 1}$$

Tehát $P_o=1$ mellett a V/LOF arány csak v penetranciájától és a variáns átlagpopulációs relatív allélfrekvenciájától (q_v) függ. A 6. ábrán különböző színekkel jelöltük a különböző q_v értékekhez tartozó görbéket. Az ábráról leolvasható, hogy minél kisebb q_v értéke, annál pontosabb a V/LOF alapján a penetrancia megítélése. Amennyiben a vizsgált variáns relatív allélfrekvenciája nagyobb, mint 20%, akkor számottevő csak a túlbecslés mértéke.



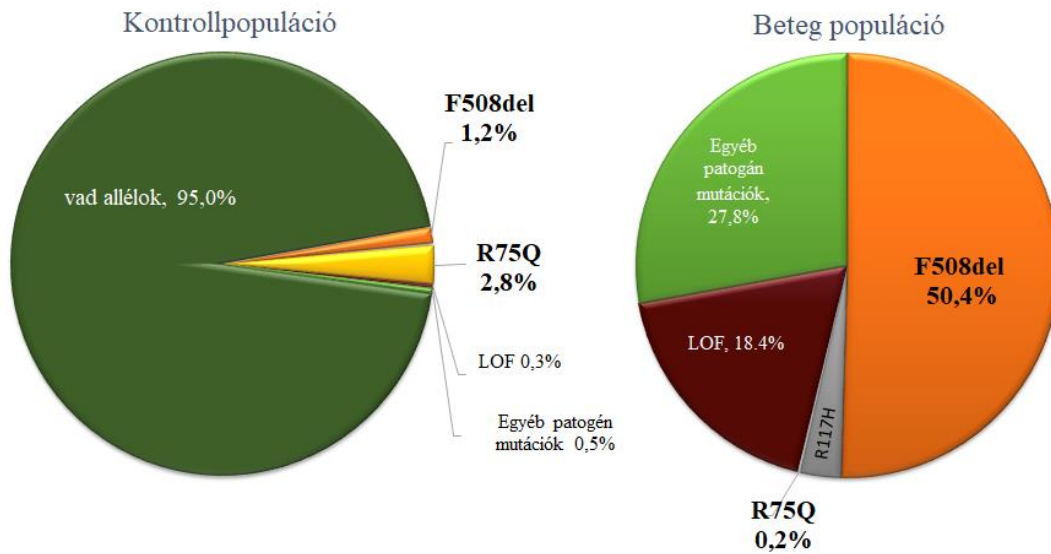
7. ábra: A V / LOF arány és a penetrancia közötti összefüggést a variáns relatív allélgyakoriságának függvényében mutatjuk be a kontrollpopuláció patogén alléljai között ($\frac{q_v}{q_v+q_t+q_{LOF}}$). A kapcsolat feltételezi, hogy az összes többi variáns teljesen penetráns ($P_o = 1$). Abban az esetben, ha az adott variáns relatív allélfrekvenciája nem haladja meg a kontrollpopuláció patogén variánsainak 20%-át ($\frac{q_v}{q_v+q_t+q_{LOF}} < 20\%$) a V/LOF arány jól korrelál a penetranciával. Hasonlóan, más variánsok inkomplett penetranciája ($P_o < 1$) szintén fokozza a túlbecslés mértékét: minél több variáns inkomplett penetráns, és minél kisebb mértékű a penetranciájuk, annál nagyobb mértékű a penetrancia túlbecslése.

Az inkompletten penetráns variánsok azonosítása a következő kritériumok alapján történt: (1) patogének, tehát szignifikánsan ($p < 0,05$) dúsulnak a beteg populációban a kontrollpopulációhoz képest; (2) a penetranciájuk (P) $P = \frac{AC_{bet}^V / AC_{gnomAD}^V}{AC_{bet}^{LOF} / AC_{gnomAD}^{LOF}}$ kevesebb mint 30%; (3) dúsulás mértékének csökkenése tükrözi a variáns penetranciáját, tehát az adott variáns beteg populációban és kontrollpopulációban levő aránya ($AC_{bet}^V / AC_{gnomAD}^V$) szignifikánsan alacsonyabb, mint a LOF variánsok aránya ($AC_{bet}^{LOF} / AC_{gnomAD}^{LOF}$), Fisher-egzakt tesztet alkalmazva.

A szignifikancia szint megállapításához Bonferroni korrekciót alkalmaztunk olyan variánsokra, amelyek szerepeltek a gnomAD-ban ($n=692$), így a 0,05-öt a elosztottuk a gnomAD-ban szereplő variánsok (penetrancia-számítások) számával, így a szignifikancia szint határértéke: $p < 7,22 \times 10^{-5}$ ($=0,05/692$) lett.

Vizsgálatunkhoz felhasznált adatok elérhetőek a https://bitbucket.org/akaposi/incomplete_penetrance linken.

A módszer alapelvét a *CFTR* gén ismert inkompletten penetráns variánsa, az R117H (c.350G>A) példáján a 8. ábra szemlélteti.



8. ábra: Az R117H allélfrekvenciája a kontroll- és a beteg populációban: az R117H variáns a kontrollpopulációban a patogén allélok 0,2%-án van jelen, míg a beteg populációban ez a variáns az allélok 3,1%-án található meg. Az R117H tehát szignifikánsan dúsul ($2,74 \times 10^{-55}$) a beteg populációban, tehát patogén variáns. Összehasonlítva viszont a teljesen penetráns, teljes funkcióvesztést okozó mutációk dúsulási arányával ($\sim 61x$) láthatjuk, hogy az R117H kisebb arányban ($\sim 15x$) dúsul a betegek között. Ez a szignifikánsan ($p = 9,07 \times 10^{-34}$) kisebb arányú dúsulás mutatja, hogy az R117H nem okoz minden esetben betegséget, tehát inkomplett penetráns.

Az R117H variáns Fisher-egzakt tesztjéhez tartozó kontingenciatáblát a 5. táblázat mutatja.

5. táblázat: Fisher-egzakt tesztet alkalmazva az R117H variáns beteg populációban és kontrollpopulációban levő aránya ($AC_{bet}^V / AC_{gnomAD}^V$) szignifikánsan alacsonyabb, mint a LOF variánsok aránya ($AC_{bet}^{LOF} / AC_{gnomAD}^{LOF}$).

	kontroll	beteg	penetrancia	p érték
R117H	311	86	20,19%	$p = 9,07 \times 10^{-34}$
LOF	353	481		

5.12. Asszociációs tesztek

Az interallelikus interakciók azonosítása céljából megvizsgáltuk az inkompletten penetráns variánsok társulásait.

Például: az ismert *NPHS2* R229Q 75 transz-asszociált misszensz alléljai közül, 53 (71%) a 4. doménban helyezkedik el, míg a nem-R229Q asszociált misszensz allélok többsége az 1-3 doménban (376/399, 94,2% $p = 3,4 \times 10^{-34}$).

Ez alapján következtethetünk arra, hogy az R229Q variáns szignifikánsan többször társul a 4. doménban levő mutációkhoz.

Ahhoz, hogy hasonló interallelikus interakciót azonosítsunk, minden egyes társulást megvizsgáltunk.

Egy variáns inkomplett penetranciájának hátterében állhat domináns negatív hatás. Azon R229Q-szerű variánsok, amelyek domináns negatív hatás miatt jelennek meg a beteg populációban, LOF variánssal összetett heterozigóta formában nem vezetnek a betegség kialakulásához. Tehát olyan IP variánsok esetében merül fel a domináns negatív hatás, melyek LOF variánsokkal szignifikánsan kevesebbszer fordultak elő és legalább egy misszensz variánssal szignifikánsan többször társulnak a beteg populációban.

Komplementáló vagy kiegészítő hatás azon variánsoknál várható, melyek nem vezetnek a fehérje lebomlásához. Így azokat az inkompletten penetráns variánsokat tartottuk komplementáció részesének, amelyek legalább egy misszensz variánssal szignifikánsan kevesebbszer fordulnak elő.

5.13. Populációgenetikai módszer validálása

Ahhoz hogy módszerünk specificitását ellenőrizzük, Fisher-egzakt teszttel összehasonlítottuk az inkomplett penetráns és a teljesen penetráns variánsok klinikai (null/hipomorf) és biológiai (misszensz/LOF) jellemzőit. Teljesen penetráns variánsoknak tekintettük azon variánsokat, amelyeknek a penetranciája 70%-ot meghaladta.

Amennyiben populációgenetikai módszerrel interallelikus interakciót igazoltunk, azt *in vitro* vizsgálatokkal tanulmányoztuk tovább.

5.14. Az *NPHS2* társulások vizsgálata FRET módszerrel

A populációgenetikai módszerrel azonosított interallelikus interakciókat experimentálisan vizsgáltuk tovább.

A Förster-féle rezonáns energiaátadás hatékonyságnak vizsgálata

Az oligomerizáció és a patogenitás összefüggésének további megértése céljából összehasonlítottuk a vad és R229Q társulásokat, illetve olyan mutációkat vizsgáltunk melyekkel társulva az R229Q patogén (A297V, A284V, R291W, P341S, F344Lfs*4), illetve a V290M variánst, mely önmagában patogén, azonban az R229Q variánshoz társulva nem. A fent leírt nem patogén társulásokat összehasonlítottuk a nem patogén társulásokkal (vad-vad, R2290-vad, R229Q-V290M). Az F344Lfs*4 variáns az utolsó exonban vezet korai stop kodonhoz, így nem vezet lebomláshoz.

Podocin expresszáló plazmidok felsokszorosítása

Módszerünket Dr. Balogh Eszter „A szteroid-rezisztens nephrosis két új mechanizmusa” című doktori értekezésében korábban már részletezte (37):

A vad podocin variánst kódoló plazmid (pcDNA 3.1 Zeo+) a francia kollaborációs partnereinktől származott. A podocin variánsokat kódoló cDNS a BamHI-XhoI/XhoI-EcoRI hasítási helyek között, a plexMCS és pEGFP-N1 humán expressziós vektorokban helyezkedett el. Az R229Q, A297V, A284V, R291W, V290M, P341S, F344Lfs*4 podocin variánsokat a QuickChange Site-directed Mutagenesis kit alkalmazásával hoztuk létre.

Két hemagglutinin (HA) taget hoztunk létre mutagenesis révén a podocin N terminális részén.

A plazmidokat kompetens *E.Coli* baktériumokba transzformáltuk és 37°C-on, plazmidok által hordozott antibiotikum rezisztencia (Ampicillin/Kanamycin) alapján felsokszorosítottuk, majd plazmid kivonó kittel tisztítottuk és végül Sanger szekvenálással ellenőriztük.

A podocin variánsok tranziens expressziója, immunprecipitációja és peptidelúciója

A fenti humán podocin variánsokat 37°C-on, DMEM-10% FBS-1% Penicillin-Streptomycin összetételű médiumban fenntartott humán embrionális vesesejtekben (HEK293) expresszáltuk, tranziens, kalcium-foszfát módszeren alapuló transzfekcióval. Mycoplasma kontaminációt havonta ellenőriztük. A transzfektált sejtekből 48 órás inkubációs időt követően 150 mM NaCl-ot, 20 mM Tris-t, 1% Triton-X detergenst tartalmazó lízis puffer segítségével, 0,1%-os proteáz inhibitorral való védelemben nyertük ki a fehérjét. A teljes sejtlizátumot 4°C-on először anti-HA monoklonális antitestekkel (1,5 h; clone HA-7, H9658, Sigma Aldrich), majd mágneses gyöngyökkel (0,5 h; Dynabeads Protein G for Immunoprecipitation, Thermo Fisher Sc.) inkubáltuk. A mágneses gyöngyök az immunglobulinok univerzális, Fc régióját kötötték meg. Az így létrejött immunprecipitátumot három alkalommal lízis pufferrel mostuk. A podocin molekulákat a HA ellenes monoklonális antitestről kompetitív elúció módszerével távolítottuk el, lízis pufferben oldott, nagy mennyiségű szintetikus HA peptid (HA Peptide lyophilized powder, B23311, Biotool) hozzáadásával, majd öt lemosást követően az eluátumokat, a megmaradt immunprecipitátumot, valamint teljes sejtlizátumot Western blottal ellenőriztük az alábbiak szerint: SDS-poliakrilamid gélelektroforézist követően nitrocellulóz membránra vittük át a fehérjéket, ezután 5% zsírszegény tejporos blokkolást majd anti-HA monoklonális ellenanyaggal történő festést végeztünk, szekunder ellenanyagként HRP-konjugált antitesteket (sc-2005, Santa Cruz Biotechnology) használtunk. A detekció elektrokemiluminszcens oldattal (sc-2048, Western Blotting Limunol Reagent) történt. A mintákat a kísérlet folytatásáig -80°C-on tároltuk.

Natív podocinok jelölése fluorofórokkal és az azt követő fluoreszcencia-spektroszkópia

A következő lépésben az egyes podocin variánsokat fluoreszcens Alexa molekulákkal festettük meg. A vizsgálni kívánt podocinpárok egyikét donor, (Alexa Fluor 488, A10254, Thermo Fisher), a másikat akceptor (Alexa fluor 555, A20346, Thermo Fisher) 4 nmol maleimid festékekkel jelöltük meg azok tiol-csoportjain, fénytől védve, 4°C-on, éjszaka inkubáltuk, majd kettesével elegyítettük a podocin variánsokat dimerizálódás céljából. A FRET mérés két különböző festéssel jelölt 4 µM-os koncentrációjú podocin között történt, az éjszakai inkubációt követően. A be nem kötött festékmolekulákat PD SpinTrap G-25 szűrővel eltávolítottuk. A két podocin variáns között a bekötött festékmolekulák (fluorofórok) révén mérhető fluoreszcens rezonancia-energia transzfer keletkezik. Ehhez szükséges a festékmolekulák kellő közelsége (<10nm). Az így kapott FRET-hatékonyság, melyet a donor élettartam csökkenéséből számítottuk ki, jellemzi az oligomerizálódó podocin variánsok között fennálló távolságot. Valamennyi mérést három alkalommal végeztük el, legalább két különböző expresszióból származó fehérjék között.

5.15. Statisztikai elemzések

A patogenitás, intrauterin letalitással járó betegségek, penetrancia, asszociációs tesztek meghatározására, illetve a populációgenetikai módszer validálása céljából Fisher-exact tesztet alkalmaztunk.

A FRET hatékonyság összehasonlítását Mann-Whitney U- próbával vizsgáltuk.

5.16. Klinikai irányelv létrehozása az *NPHS2* R229Q asszociációk értelmezésére

Különböző megközelítések alapján (családvizsgálatok, klinikai, populációgenetika, biológiai- és biofizikai) létrehoztunk egy klinikai útmutatót, mely segíti az *NPHS2* R229Q variánssal való társulások patogénitásának eldöntését. A családvizsgálatok adatait, az *in vitro* és *in silico* vizsgálatok eredményeit a korábban a kutatócsoportunk által publikált tanulmányokból használtuk fel (36, 37). Populációgenetikai számítások a kigyűjtött beteg-adatbázisból történtek. A fent részletezett módszer alapján összesen 545 biallelikus *NPHS2* variánst hordozó beteget gyűjtöttünk ki, akik közül 106 beteg egyik allélján R229Q variáns szerepelt. Őket a következőképpen jelöljük: [R229Q];[mut]. A további 439 beteg allélján nem szerepelt az R229Q variáns [mut];[mut].

Az útmutatóban feltüntettük az összes ismert R229Q transzasszociált allélt és az ezekhez tartozó adatokat. Azon variánsok patogének az R229Q variánssal, amelyek szignifikánsan dúsulnak az [R229Q];[mut] hordozó betegek között, a többi beteghez képest ([mut];[mut]). Ezzel a módszerrel: [AF[R229Q];[mut]]/(AF[mut];[mut]) azon variánsok azonosíthatóak, amelyek domináns negatív hatással vannak az R229Q variánssra: A284V, A288T, R291W, A297V, E310K, E310V, L327F, Q328R és F344Lfs*4.

A módszerrel azt is megállapítottuk, mely variánsokkal társulva nem patogén az R229Q variáns. Ezen társulások azonosítása csak véletlen, így alulreprezentáltak lesznek a beteg populációban. A módszer a gyakori variánsok esetén informatív, ezért ahhoz, hogy egy variáns alulreprezentáltsága a 106 [R229Q];[mut] beteg között szignifikáns ($p < 0,05$) legyen, legalább 32-szer kellett előfordulnia a 878 [mut];[mut] allél között. Ezek alapján három variáns bizonyult alulreprezentáltnak: R138Q, L156Ffs*11 és R168H.

Tekintettel arra, hogy az R229Q-val patogén variánsok kumulatív allélfrekvenciája az európai kontrollpopulációban (gnomAD) 5/126,000-hez, azon variánsok, melyek a kontrollpopulációban relatív magas európai allélfrekvenciával rendelkeznek és nem jelennek meg, vagy alulreprezentáltak az [R229Q];[mut] beteg populációban, várhatóan nem patogének az R229Q variánssal társulva: I263L, E264Q, G273W, R291Q, V319L, Q320P, K377R, R286Tfs*17, V290M, A317Lfs*31.

6. Eredmények

6.1. Nem patogén variánsok azonosítása

Populációgenetikai számítások során egy nem patogén variáns tévesen IP variánsként azonosítható, ezért célunk volt a nem patogén HGMD variánsok kizárása a vizsgálatainkból. Azokat a variánsokat, amelyeknek az allélfrekvenciája csökkent a beteg populációban a kontrollpopulációhoz képest, nem patogénnek ítéltük. A 17 AR betegségben összesen 25 ilyen variánst találtunk, amelyek egyike sem volt LOF variáns (4. ábra, 6. táblázat). Összesen 111 variáns patogenitása nem ítélt meg egyértelműen az allélfrekvenciájuk változása alapján, mivel dúsultak a beteg populációban, de nem szignifikáns mértékben (4. ábra). Ezen variánsokat kizártuk a további vizsgálatokból. A nem patogén variánsokat különböző programokkal (PolyPhen-2, Mutation Taster, SIFT) is megvizsgáltuk, melyek a variánsok patogenitását az aminosav konzerváltsága és az oldallánc-változás alapján ítélik meg.

6. táblázat: A 25 nem patogén variáns, melyek a beteg populációban ritkábban fordulnak elő, mint a kontrollpopulációban (44). A táblázatban feltüntettük a variánsok allélfrekvenciáját beteg- és kontrollpopulációban, illetve a programok predikcióit.

Gén	variáns		európai AF		PolyPhen-2	Mut Taster	SIFT	
	nukleotid	aminosav	betegpopuláció ACvariáns/ ACössz. variáns (%)	gnomAD ACvariáns/ ACössz. variáns (%)				
1	<i>CAPN3</i>	c.984C>T	C328C	1/352 (0.28)	1303/126542 (1.03)		dc	
2	<i>CFTR</i>	c.91C>T	R31C	1/3342 (0.03)	186/126228 (0.15)	ben 0.057	dc	del 0.01
3	<i>CFTR</i>	c.224G>A	R75Q	6/3342 (0.18)	3596/126364 (2.85)	ben 0.404	dc	del 0.01
4	<i>CFTR</i>	c.443T>C	I148T	1/3342 (0.03)	191/126106 (0.15)	ben 0.024	dc	tol 0.34
5	<i>CFTR</i>	c.509G>A	R170H	1/3342 (0.03)	47/126250 (0.04)	pos dam 0.635	dc	del 0.01
6	<i>CFTR</i>	c.650A>G	E217G	1/3342 (0.03)	160/126658 (0.13)	ben 0.005	dc	tol 0.16
7	<i>CFTR</i>	c.1516A>G	I506V	1/3342 (0.03)	58/126578 (0.05)	pos dam 0.582	dc	tol 0.06
8	<i>CFTR</i>	c.1523T>G	F508C	4/3342 (0.12)	218/123520 (0.18)	prob dam 0.974	dc	del 0
9	<i>CFTR</i>	c.1584G>A	splicing site	9/3342 (0.27)	3064/126290 (2.43)		dc	
10	<i>CFTR</i>	c.2506G>T	D836Y	1/3342 (0.03)	50/126558 (0.04)	pos dam 0.684	dc	tol 0.1
11	<i>CFTR</i>	c.2735C>T	S912L	1/3342 (0.03)	46/126678 (0.04)	ben 0.024	pm	tol 0.64
12	<i>CFTR</i>	c.2900T>C	L967S	2/3342 (0.06)	132/126622 (0.1)	pos dam 0.458	dc	del 0.01
13	<i>CFTR</i>	c.3485G>T	R1162L	1/3342 (0.03)	164/126100 (0.13)	prob dam 0.998	dc	del 0
14	<i>IDUA</i>	c.314G>A	R105Q	1/220 (0.45)	19851/126176 (15.73)	ben 0.007		tol 1
15	<i>IDUA</i>	c.1081G>A	A361T	1/220 (0.45)	19415/122736 (15.82)	ben 0.004		tol 1
16	<i>NPHS1</i>	c.791C>G	P264R	2/218 (0.92)	1939/125382 (1.55)	ben 0.313	pm	tol 0.21
17	<i>NPHS1</i>	c.1223G>A	R408Q	4/218 (1.83)	7961/126654 (6.29)	prob dam 0.997		del 0.01
18	<i>PKHD1</i>	c.1736C>T	T579M	3/580 (0.52)	4511/126308 (3.57)	prob dam 0.968	pm	del 0.03
19	<i>PKHD1</i>	c.2489A>G	N830S	1/580 (0.17)	8736/126490 (6.91)	ben 0.033		tol 0.07
20	<i>PKHD1</i>	c.3407A>G	Y1136C	3/580 (0.52)	1612/126674 (1.27)	ben 0.013	pm	tol 0.08
21	<i>PKHD1</i>	c.5125C>T	L1709F	3/580 (0.52)	658/126596 (0.52)	prob dam 1	dc	del 0.01
22	<i>PKHD1</i>	c.8606C>A	T2869K	5/580 (0.86)	1955/126258 (1.55)	prob dam 0.97	pm	del 0
23	<i>PKHD1</i>	c.9415G>T	D3139Y	1/580 (0.17)	757/126108 (0.6)	ben 0.035	pm	tol 0.11
24	<i>PKHD1</i>	c.11525G>T	A3842L	1/580 (0.17)	3049/126226 (2.42)	ben 0.195	pm	tol 0.33
25	<i>PMM2</i>	c.590A>C	E197A	1/496 (0.2)	3472/126702 (2.74)	ben 0.022	dc	tol 0.15

AC allele count/allél szám; AF allélfrekvencia; ben, benignus; dc, disease causing/betegséget okozó; dam, damaging/kóros; del, deleterious/káros; Mut Taster, Mutation Taster; tol, tolerálható pm, polimorfizmus; pos, possible/lehetséges; pr, probably/valószínűleg

7. táblázat: A beteg populációban nem szignifikánsan dúsult 111 variáns (44). A táblázatban feltüntettük a variánsok allélfrekvenciáját a beteg- és kontrollpopulációban, illetve a programok predikcióit.

	gén	variáns	Európai betegek AF		PolyPhen-2 score	MutTast	SIFT
			beteg populáció	gnomAD			
1	<i>ATP7B</i>	c.122A>G (p.Asn41Ser)	1/552 (0.181%)	54/126572 (0.0427%)	pr dam 0.995	dc	del 0.01
2	<i>ATP7B</i>	c.347T>C (p.Ile116Thr)	1/552 (0.181%)	21/126578 (0.0166%)	pos dam 0.864	pm	del 0
3	<i>ATP7B</i>	c.562C>T (p.Gln188*)	1/552 (0.181%)	1/15002 (0.00667%)		dc	
4	<i>ATP7B</i>	c.1995G>A (p.Met665Ile)	2/552 (0.362%)	216/126368 (0.171%)	ben 0.013	dc	del 0.04
5	<i>ATP7B</i>	c.2605G>A (p.Gly869Arg)	2/552 (0.362%)	159/126726 (0.125%)	pr dam 0.997	dc	del 0
6	<i>ATP7B</i>	c.3106G>A (p.Val1036Ile)	1/552 (0.181%)	13/126198 (0.0103%)	ben 0.128	dc	tol 0.07
7	<i>ATP7B</i>	c.3243+5G>A	1/552 (0.181%)	22/126240 (0.0174%)		dc	
8	<i>ATP7B</i>	c.3547_3548delGC (p.Ala1183Tyrfs*2)	1/552 (0.181%)	1/15008 (0.00666%)			
9	<i>CAPN3</i>	c.640G>A (p.Gly214Ser)	1/352 (0.284%)	25/126696 (0.0197%)	pos dam 0.766	dc	del 0.01
10	<i>CAPN3</i>	c.1746-20C>G	4/352 (1.14%)	595/126710 (0.47%)		pm	
11	<i>CFTR</i>	c.11C>A (p.Ser4*)	1/3342 (0.0299%)	1/14992 (0.00667%)		dc	
12	<i>CFTR</i>	c.125C>T (p.Ser42Phe)	2/3342 (0.0598%)	29/111202 (0.0261%)	ben 0.232	dc	del 0.01
13	<i>CFTR</i>	c.166G>A (p.Glu56Lys)	1/3342 (0.0299%)	3/111210 (0.0027%)	ben 0.01	dc	del 0.02
14	<i>CFTR</i>	c.200C>T (p.Pro67Leu)	1/3342 (0.0299%)	8/126388 (0.00633%)	pr dam 0.983	dc	del 0
15	<i>CFTR</i>	c.220C>T (p.Arg74Trp)	2/3342 (0.0598%)	40/126376 (0.0317%)	ben 0.001	dc	tol 0.19
16	<i>CFTR</i>	c.262_263delTT (p.Leu88Ilefs*22)	3/3342 (0.0898%)	37/126278 (0.0293%)			
17	<i>CFTR</i>	c.358G>A (p.Ala120Thr)	1/3342 (0.0299%)	17/126378 (0.0135%)	pos dam 0.635	dc	del 0.04
18	<i>CFTR</i>	c.377G>A (p.Gly126Asp)	1/3342 (0.0299%)	3/111338 (0.00269%)	pr dam 0.929	dc	del 0
19	<i>CFTR</i>	c.476T>C (p.Leu159Ser)	1/3342 (0.0299%)	3/125054 (0.0024%)	pr dam 0.914	dc	tol 0.1
20	<i>CFTR</i>	c.489+3A>G	3/3342 (0.0898%)	45/122898 (0.0366%)		dc	
21	<i>CFTR</i>	c.596A>G (p.His199Arg)	1/3342 (0.0299%)	1/111662 (0.000896%)	pr dam 0.95	dc	del 0.02
22	<i>CFTR</i>	c.638G>A (p.Gly213Glu)	1/3342 (0.0299%)	7/126662 (0.00553%)	pr dam 0.97	dc	del 0.01
23	<i>CFTR</i>	c.794T>G (p.Met265Arg)	1/3342 (0.0299%)	4/124204 (0.00322%)	ben 0.285	dc	del 0
24	<i>CFTR</i>	c.861_865delCTTAA (p.Asn287Lysfs*19)	1/3342 (0.0299%)	1/111056 (0.0009%)			
25	<i>CFTR</i>	c.1001G>A (p.Arg334Gln)	2/3342 (0.0598%)	17/126496 (0.0134%)	ben 0.006	dc	tol 0.2
26	<i>CFTR</i>	c.1007T>A (p.Ile336Lys)	1/3342 (0.0299%)	1/111496 (0.000897%)	pr dam 0.984	dc	del 0
27	<i>CFTR</i>	c.1021_1022dupTC (p.Phe342Hisfs*28)	1/3342 (0.0299%)	8/126494 (0.00632%)			
28	<i>CFTR</i>	c.1043T>A (p.Met348Lys)	1/3342 (0.0299%)	25/126440 (0.0198%)	ben 0.095	dc	del 0.01
29	<i>CFTR</i>	c.1046C>T (p.Ala349Val)	1/3342 (0.0299%)	28/126384 (0.0222%)	ben 0.143	dc	del 0.02

30	CFTR	c.1052C>G (p.Thr351Ser)	1/3342 (0.0299%)	23/126444 (0.0182%)	ben 0.143	dc	del 0.02
31	CFTR	c.1054C>T (p.Arg352Trp)	1/3342 (0.0299%)	2/126426 (0.00158%)	ben 0.023	dc	del 0
32	CFTR	c.1125A>C (p.Leu375Phe)	1/3342 (0.0299%)	11/111140 (0.0099%)	pr dam 0.928	dc	del 0
33	CFTR	c.1135G>T (p.Glu379*)	1/3342 (0.0299%)	1/111336 (0.000898%)		dc	
34	CFTR	c.1327G>T (p.Asp443Tyr)	2/3342 (0.0598%)	56/121460 (0.0461%)	pos dam 0.689	dc	del 0.01
35	CFTR	c.1394C>A (p.Thr465Asn)	1/3342 (0.0299%)	1/111608 (0.000896%)	ben 0.365	dc	del 0
36	CFTR	c.1438G>T (p.Gly480Cys)	1/3342 (0.0299%)	0/14990 (0%)	pr dam 0.999	dc	del 0
37	CFTR	c.1466C>A (p.Ser489*)	1/3342 (0.0299%)	2/126598 (0.00158%)		dc	
38	CFTR	c.1475C>T(p.Ser492Phe)	1/3342 (0.0299%)	1/111614 (0.000896%)	pos dam 0.694	dc	del 0
39	CFTR	c.1679G>C (p.Arg560Thr)	1/3342 (0.0299%)	5/111106 (0.0045%)	pr dam 0.993	dc	del 0
40	CFTR	c.1724T>A (p.Phe575Tyr)	1/3342 (0.0299%)	3/111242 (0.0027%)	pr dam 0.964	dc	del 0.02
41	CFTR	c.1837G>A (p.Ala613Thr)	1/3342 (0.0299%)	1/15000 (0.00667%)	tol 0.08	dc	ben 0.28
42	CFTR	c.2017G>T (p.Gly673*)	1/3342 (0.0299%)	1/110540 (0.000905%)		dc	
43	CFTR	c.2052delA (p.Lys684Asnfs*38)	1/3342 (0.0299%)	15/124874 (0.012%)			
44	CFTR	c.2417A>G (p.Asp806Gly)	1/3342 (0.0299%)	1/86326 (0.00116%)	ben 0.25	dc	del 0
45	CFTR	c.2421A>G (p.Ile807Met)	2/3342 (0.0598%)	56/101356 (0.0553%)	ben 0.285	dc	del 0.01
46	CFTR	c.2557A>T (p.Ile853Phe)	1/3342 (0.0299%)	1/111624 (0.000896%)	ben 0.001	pm	tol 0.32
47	CFTR	c.2563G>A (p.Val855Ile)	1/3342 (0.0299%)	13/126602 (0.0103%)	ben 0	pm	tol 0.68
48	CFTR	c.2813T>G (p.Val938Gly)	1/3342 (0.0299%)	3/111678 (0.00269%)	pos dam 0.524	dc	del 0
49	CFTR	c.3041A>G (p.Tyr1014Cys)	2/3342 (0.0598%)	14/126526 (0.0111%)	pr dam 0.994	dc	del 0
50	CFTR	c.3067_3072delATAGTG (p.Ile1023_Val1024del)	1/3342 (0.0299%)	1/111492 (0.000897%)			
51	CFTR	c.3154T>G (p.Phe1052Val)	3/3342 (0.0898%)	113/126118 (0.0896%)	pr dam 0.975	dc	del 0
52	CFTR	c.3205G>A (p.Gly1069Arg)	2/3342 (0.0598%)	28/126400 (0.0222%)	ben 0.011	dc	tol 0.31
53	CFTR	c.3209G>A (p.Arg1070Gln)	1/3342 (0.0299%)	5/111414 (0.00449%)	pos dam 0.572	dc	tol 0.09
54	CFTR	c.3222T>A (p.Phe1074Leu)	1/3342 (0.0299%)	1/111482 (0.000897%)	pos dam 0.482	dc	del 0
55	CFTR	c.3274T>C (p.Tyr1092His)	1/3342 (0.0299%)	8/126570 (0.00632%)	pos dam 0.842	dc	tol 0.1
56	CFTR	c.3275A>G (p.Tyr1092Cys)	1/3342 (0.0299%)	1/111538 (0.000897%)	pr dam 0.919	dc	del 0.03
57	CFTR	c.3371_3373delAAG (p.Glu1124del)	1/3342 (0.0299%)	1/111256 (0.000899%)			
58	CFTR	c.3389G>C (p.Gly1130Ala)	1/3342 (0.0299%)	3/111360 (0.00269%)	pr dam 0.993	dc	del 0.02
59	CFTR	c.3409A>G (p.Met1137Val)	1/3342 (0.0299%)	6/126416 (0.00475%)	ben 0.225	dc	del 0
60	CFTR	c.3415A>G (p.Ile1139Val)	1/3342 (0.0299%)	28/126434 (0.0221%)	pos dam 0.482	dc	tol 0.17
61	CFTR	c.3718-24G>A	1/3342 (0.0299%)	9/126256 (0.00713%)		pm	
62	CFTR	c.3719T>G (p.Val1240Gly)	1/3342 (0.0299%)	1/111274 (0.000899%)	pr dam 0.931	dc	del 0
63	CFTR	c.3808G>A (p.Asp1270Asn)	2/3342 (0.0598%)	15/126386 (0.0119%)	pos dam 0.618	dc	del 0

64	<i>CFTR</i>	c.3873G>C (p.Gln1291His)	1/3342 (0.0299%)	2/111268 (0.0018%)	pr dam 0.966	dc	del 0
65	<i>CFTR</i>	c.3889dupT (p.Ser1297Phefs*5)	1/3342 (0.0299%)	5/109962 (0.00455%)			
66	<i>CFTR</i>	c.3908delA (p.Asn1303Thrfs*25)	1/3342 (0.0299%)	1/110514 (0.000905%)			
67	<i>CFTR</i>	c.4129G>C (p.Asp1377His)	1/3342 (0.0299%)	1/111422 (0.000897%)	pr dam 0.979	dc	del 0
68	<i>CFTR</i>	c.4333G>A (p.Asp1445Asn)	2/3342 (0.0598%)	37/126322 (0.0293%)	ben 0.02	dc	del 0
69	<i>DHCR7</i>	c.89G>C (p.Gly30Ala)	1/594 (0.168%)	18/126718 (0.0142%)	pr dam 0.999	dc	del 0.05
70	<i>DHCR7</i>	c.854_856delTCT (p.Phe285del)	1/594 (0.168%)	1/14988 (0.00667%)			
71	<i>GALNS</i>	c.499T>G (p.Phe167Val)	1/134 (0.746%)	227/126672 (0.179%)	pos dam 0.6	dc	tol 0.09
72	<i>GALNS</i>	c.1127G>A (p.Arg376Gln)	1/134 (0.746%)	261/124678 (0.209%)	ben 0.004	pm	tol 0.6
73	<i>IDUA</i>	c.1727+4C>T	1/220 (0.455%)	3/14942 (0.0201%)		dc	
74	<i>NPHS1</i>	c.1610C>T (p.Thr537Met)	1/218 (0.459%)	31/101392 (0.0306%)	pr dam 0.982	dc	del 0.01
75	<i>NPHS2</i>	c.83A>T (p.Lys28Met)	1/434 (0.23%)	1/14950 (0.00669%)	pos dam 0.822	pm	del 0.01
76	<i>NPHS2</i>	c.467delT (p.Leu156Cysfs*25)	1/434 (0.23%)	35/83688 (0.0418%)			
77	<i>PAH</i>	c.136G>A (p.Gly46Ser)	3/5070 (0.0592%)	20/126650 (0.0158%)	pr dam 1	dc	tol 0.13
78	<i>PAH</i>	c.158G>A(p.Arg53His)	4/5070 (0.0789%)	86/126626 (0.0679%)	ben 0.006	dc	del 0.01
79	<i>PAH</i>	c.169G>A (p.Glu57Lys)	1/5070 (0.0197%)	3/126212 (0.00238%)	ben 0.003	dc	tol 0.28
80	<i>PAH</i>	c.250G>T (p.Asp84Tyr)	1/5070 (0.0197%)	2/111664 (0.00179%)	pos dam 0.883	dc	del 0
81	<i>PAH</i>	c.275C>T (p.Thr92Ile)	1/5070 (0.0197%)	1/111664 (0.000896%)	ben 0.039	pm	tol 0.29
82	<i>PAH</i>	c.355C>T (p.Pro119Ser)	1/5070 (0.0197%)	1/111592 (0.000896%)	pr dam 0.975	dc	del 0.01
83	<i>PAH</i>	c.434A>T (p.Asp145Val)	2/5070 (0.0394%)	12/126680 (0.00947%)	pr dam 1	dc	del 0
84	<i>PAH</i>	c.463C>T (p.Arg155Cys)	1/5070 (0.0197%)	1/111606 (0.000896%)	pr dam 1	dc	del 0
85	<i>PAH</i>	c.500A>T (p.Asn167Ile)	2/5070 (0.0394%)	9/126638 (0.00711%)	pos dam 0.714	dc	del 0
86	<i>PAH</i>	c.506G>A (p.Arg169His)	1/5070 (0.0197%)	11/124992 (0.0088%)	ben 0.025	dc	del 0.01
87	<i>PAH</i>	c.841C>T (p.Pro281Ser)	1/5070 (0.0197%)	1/111500 (0.000897%)	pr dam 1	dc	del 0
88	<i>PAH</i>	c.859C>G (p.Leu287Val)	1/5070 (0.0197%)	2/111284 (0.0018%)	pos dam 0.472	dc	del 0.03
89	<i>PAH</i>	c.868C>T (p.His290Tyr)	1/5070 (0.0197%)	1/111292 (0.000899%)	pr dam 0.996	dc	del 0
90	<i>PAH</i>	c.926C>A (p.Ala309Asp)	1/5070 (0.0197%)	1/111544 (0.000897%)	pr dam 0.999	dc	del 0
91	<i>PAH</i>	c.964G>A (p.Ala322Thr)	1/5070 (0.0197%)	4/111492 (0.00359%)	pr dam 0.993	dc	del 0.01
92	<i>PAH</i>	c.965C>G (p.Ala322Gly)	1/5070 (0.0197%)	5/126506 (0.00395%)	pr dam 0.996	dc	del 0
93	<i>PAH</i>	c.982A>G (p.Thr328Ala)	1/5070 (0.0197%)	3/111410 (0.00269%)	pr dam 1	dc	del 0
94	<i>PAH</i>	c.994G>A (p.Gly332Arg)	1/5070 (0.0197%)	1/15006 (0.00666%)	pr dam 1	dc	del 0
95	<i>PAH</i>	c.1066-7C>A	1/5070 (0.0197%)	1/111290 (0.000899%)		pm	
96	<i>PAH</i>	c.1130A>G (p.Tyr377Cys)	1/5070 (0.0197%)	1/111590 (0.000896%)	pr dam 1	dc	tol 0.36
97	<i>PAH</i>	c.1161C>A (p.Tyr387*)	1/5070 (0.0197%)	0/14996 (0%)		dc	

98	<i>PAH</i>	c.1199+1G>A	1/5070 (0.0197%)	1/111584 (0.000896%)		dc	
99	<i>PAH</i>	c.1200-8G>A	1/5070 (0.0197%)	1/111684 (0.000895%)	pr dam 0.913	dc	del 0
100	<i>PAH</i>	c.1262T>C (p.Ile421Thr)	1/5070 (0.0197%)	1/111692 (0.000895%)	pos dam 0.82	dc	tol 0.13
101	<i>PAH</i>	c.1340C>A (p.Ala447Asp)	1/5070 (0.0197%)	3/126500 (0.00237%)	pr dam 0.982	dc	del 0
102	<i>PKHD1</i>	c.652G>A (p.Glu218Lys)	1/580 (0.172%)	28/126212 (0.0222%)	ben 0.21	pm	del 0.01
103	<i>PKHD1</i>	c.5410C>T (p.Arg1804Cys)	1/580 (0.172%)	21/126138 (0.0166%)	ben 0.276	pm	tol 0.26
104	<i>PKHD1</i>	c.5768A>T (p.Gln1923Leu)	1/580 (0.172%)	172/125990 (0.137%)	ben 0.035	pm	tol 0.14
105	<i>PKHD1</i>	c.7675G>C (p.Val2559Leu)	1/580 (0.172%)	212/126354 (0.168%)	ben 0.001	pm	tol 0.49
106	<i>PKHD1</i>	c.7921A>G (p.Thr2641Ala)	1/580 (0.172%)	28/126528 (0.0221%)	pr dam 0.962	pm	del 0
107	<i>PKHD1</i>	c.8521A>G (p.Met2841Val)	1/580 (0.172%)	49/125822 (0.0389%)	ben 0.016	pm	tol 0.05
108	<i>PKHD1</i>	c.9241A>G (p.Ile3081Val)	1/580 (0.172%)	21/126424 (0.0166%)	pr dam 0.999	dc	del 0.03
109	<i>SLC26A4</i>	c.1003T>C (p.Phe335Leu)	1/350 (0.286%)	121/126468 (0.0957%)	pr dam 0.999	dc	tol 0.34
110	<i>SLC26A4</i>	c.1790T>C (p.Leu597Ser)	3/350 (0.857%)	949/126404 (0.751%)	pr dam 0.995	dc	del 0
111	<i>SLC26A4</i>	c.1826T>G (p.Val609Gly)	1/350 (0.286%)	79/126202 (0.0626%)	ben 0		tol 0.37

AF, allélfrekvencia; ben, benignus; dc, disease causing/betegséget okozó; dam, damaging/kóros; del, deleterious/káros; Mut Tast, Mutation Taster; tol, tolerálható; pm, polimorfizmus; pos, possible/lehetséges; pr, probably/valószínűleg;

6.2. Intrauterin letalitást okozó betegségek kizárása

A penetrancia számításához feltételeztük, hogy a LOF variánsok teljesen penetránsak. Azon géneket, melyek LOF variánsai in utero letalitást okoznak, célunk volt azonosítani és kizárni a további vizsgálatokból.

HWE tesztel vizsgáltuk a misszensz és LOF mutációk társulását. A vizsgálat során a *DHCR7* gén esetében jelentős diszequilibriumot tapasztaltunk. HWE teszt eredményeit a 8. táblázatban részletezzük.

8. táblázat: HWE teszt eredményei: A *DHCR7* gén esetében szignifikáns diszkekvilibrium van (44).

Gén		[missz];[missz]	[missz];[LOF]	[LOF];[LOF]	p
<i>ASL</i>	figy	44	41	7	0.544
	várt	45.22	38.55	8.22	
<i>ATP7B</i>	figy	134	103	10	0.0706
	várt	139.31	92.37	15.31	
<i>CAPN3</i>	figy	22	74	43	0.291
	várt	25.04	67.91	46.04	
<i>CFTR</i>	figy	982	379	33	0.616
	várt	984.51	373.97	35.51	
<i>CTNS</i>	figy	5	18	8	0.337
	várt	6.32	15.35	9.32	
<i>DHCR7</i>	figy	90	176	24	1.84x10 ⁻⁶
	várt	109.25	137.48	43.25	
<i>GAA</i>	figy	19	24	11	0.503
	várt	17.79	26.40	9.79	
<i>GALNS</i>	figy	37	24	1	0.186
	várt	38.72	20.54	2.72	
<i>GALT</i>	figy	69	10	0	0.548
	várt	69.31	9.367	0.31	
<i>IDUA</i>	figy	15	39	28	0.826
	várt	14.51	39.96	27.51	
<i>MUT</i>	figy	53	51	8	0.363
	várt	55.02	46.95	10.02	
<i>NPHS1</i>	figy	27	49	13	0.223
	várt	29.80	49.93	15.80	
<i>NPHS2</i>	figy	117	53	7	0.747
	várt	116.34	54.31	6.34	
<i>PAH</i>	figy	1502	310	19	0.503
	várt	1499.53	314.92	16.53	
<i>PKHD1</i>	figy	87	129	43	0.679
	várt	88.61	125.76	44.61	
<i>PMM2</i>	figy	227	16	0	0.596
	várt	227.26	15.47	0.265	
<i>SLC26A4</i>	figy	108	45	9	0.149
	várt	105.12	50.75	6.12	

missz, misszensz; figy, megfigyelt;

Összehasonlítottuk a misszensz és a LOF mutációk arányát a beteg és kontrollpopulációban.

9. táblázat: misszensz/LOF arány összehasonlítása a beteg és kontrollpopulációban: *NPHS2*, *ASL*, *PKHD1*, *CAPN3* és *CFTR* génekben a misszensz variánsok allélfrekvenciája szignifikánsan alacsonyabb. Az *NPHS1*, *IDUA*, *CTNS* és *DHCR7* gén esetében a LOF mutációk allélfrekvenciája szignifikánsan alacsonyabb (44).

Gén	misszensz / LOF allélok (arány)		arány (beteg/kontroll)	p (limit: 0.00294)
	európai beteg populáció	európai kontrollpopuláció		
<i>NPHS2</i>	320/72 (4.44)	4727/67 (70.55)	0.06	1.63x10 ⁻⁴⁵
<i>ASL</i>	141/60 (2.35)	495/48 (10.31)	0.23	9.25x10 ⁻¹²
<i>PKHD1</i>	308/220 (1.4)	528/209 (2.53)	0.55	1.02x10 ⁻⁶
<i>CAPN3</i>	136/173 (0.79)	162/120 (1.35)	0.58	0.00131
<i>GALT</i>	160/10 (16)	407/15 (27.13)	0.59	0.257
<i>PAH</i>	3870/413 (9.37)	1178/81 (14.54)	0.64	0.00381
<i>GALNS</i>	101/26 (3.88)	124/22 (5.64)	0.69	0.267
<i>CFTR</i>	2485/481 (5.17)	2506/353 (7.1)	0.73	2.72 x10 ⁻⁵
<i>SLC26A4</i>	268/64 (4.19)	416/105 (3.96)	1.06	0.792
<i>GAA</i>	129/127 (1.02)	69/89 (0.78)	1.31	0.189
<i>MUT</i>	163/69 (2.36)	84/48 (1.75)	1.35	0.201
<i>ATP7B</i>	387/127 (3.05)	390/191 (2.04)	1.49	0.00332
<i>PMM2</i>	473/16 (29.56)	696/45 (15.47)	1.91	0.0312
<i>NPHS1</i>	110/81 (1.36)	73/173 (0.42)	3.22	6.29x10 ⁻⁹
<i>IDUA</i>	74/112 (0.66)	35/187 (0.19)	3.53	5.27x10 ⁻⁸
<i>CTNS</i>	36/34 (1.06)	12/51 (0.24)	4.5	0.000129
<i>DHCR7</i>	360/225 (1.6)	190/919 (0.21)	7.74	9.45x10 ⁻⁷⁶

Azt tapasztaltuk, hogy a *DHCR7* gén LOF mutációi jelentősen alulreprezentáltak a beteg populációban (7,7x, $p = 9,45 \times 10^{-76}$) (9. táblázat). Ezen eredmények alapján arra következtetünk, hogy a biallelikus LOF mutációk in utero letalitást okoznak, ezért a beteg populációban nem jelennek meg. Ezt a felvetést az irodalomban több tanulmány vizsgálta (80, 81).

A fenti eredmények alapján a *DHCR7* gént kizártuk a további vizsgálatokból.

Az *NPHS1*, *IDUA* és *CTNS* LOF alléljai szintén szignifikánsan alulreprezentáltak voltak, viszont ezen gének esetében a HWE teszt diszekvilibriumot nem mutatott. Ennek háttérében az állhat, hogy az *NPHS1* és *IDUA* génben három olyan LOF mutáció van, amelyek a kontrollpopulációban való allélfrekvenciája, a kontroll személyek földrajzi

összetételéből fakadóan torzítja az eredményeinket: az *NPHS1* Fin major L41Dfs*50 (gnomAD AF:27/195), a Fin minor *NPHS1*, R1109* (gnomAD AF:19/195), és az *IDUA* Q70* variánsa (gnomAD AF:85/207). Ez a három variáns gyakori a svéd populációban. Mivel a svédek aránytalanul nagy számban vannak jelen a gnomAD európai-nem finn csoportjában (43), ezért magas ezen variánsok allélfrekvenciája a kontrollcsoportban. Az *NPHS2*, *ASL*, *PKHD1*, *CAPN3* és *CFTR* gén misszensz mutációi alulreprezentáltak voltak a beteg populációban. Ebből következtethetünk arra, hogy ezen génekben vannak IP variánsok.

6.3. Inkompletten penetráns variánsok azonosítása

Inkompletten penetránsnak tartottuk azt a variánst, amely: 1) szignifikánsan ($p < 0,05$) dúsul a beteg populációban a kontrollhoz képest, 2) az adott variáns dúsulása a LOF variánsok dúsulásához képest szignifikánsan ($p < 7,22 \times 10^{-05}$) alacsonyabb 3) illetve penetranciája 30% alatti.

Számos patogén variáns, különösen a ritka variánsok egyszer sem fordulnak elő a gnomAD adatbázisban, ezért alacsony azon variánsoknak a száma, melyek penetranciáját frekvencia-csökkenésük alapján meg lehet ítélni.

Az 1936 mutáció közül 85 volt elég gyakori a kontrollpopulációban ahhoz, hogy teoretikusan szignifikánsan csökkenhessen a relatív allélfrekvenciájuk a beteg populációban. Ezek közül 25 (29,4%) bizonyult inkompletten penetránsnak, közöttük a három korábban igazolt IP variáns: (*CFTR* R117H (9), L997F (82) és *NPHS2* R229Q (36) és 22 új. Ezeket a variánsokat a 10. táblázatban mutatom be.

IP variánsokat azon génekből igazoltunk, melyeknél a misszensz variánsok allélfrekvenciája alacsonyabb volt a beteg populációban a LOF mutációkhoz képest (9. táblázat).

10. táblázat: A 25 inkompleten penetráns variáns (44). A táblázatban feltüntettük a variánsok penetranciáját az $\frac{AC_{bet}^V / AC_{gnomAD}^V}{AC_{bet}^{LOF} / AC_{gnomAD}^{LOF}}$ képletnek megfelelően, illetve a variánst hordozó betegek nemzetiségét és a variáns klinikai megítélését (hipomorf/null).

	gén	variáns	Európai AF		LOF _{BET} / LOF _{gnomAD} (Európai AC)	dúsulás betegpop. vs. kontrollpop. (p)	penetrancia (p)	nemzetiség	klinikai beosztás
			betegpop	gnomAD					
1	ASL	c.35G>A R12Q	6/218 (2.75%)	252/126330 (0.2%)	60/48	13.8x (6.56x10 ⁻⁶)	1.9% (9.15x10 ⁻³²)	CAN, DK, AT, IT, USA	hip
2		c.392C>T T131M	2/218 (0.92%)	94/126120 (0.07%)		12.3x (0.012)	1.70% (4.09x10 ⁻¹⁹)	TR, USA	na
3	CAPN3	c.1029+3A>AG	1/352 (0.28%)	17/126348 (0.01%)	173/120	21.1x (0.049)	0.0408% (5.43x10 ⁻⁶)	CZE, DK, GB	na
4		c.2393C>A A798E	5/352 (1.42%)	26/126636 (0.02%)		69.2x (2.55x10 ⁻⁸)	13.33% (4.90x10 ⁻⁶)	GB, ES	hip
5	CFTR	c.349C>T R117C	8/3342 (0.24%)	49/126402 (0.04%)	481/353	6.2x (1.03x10 ⁻⁴)	11.98% (5.67x10 ⁻¹¹)	DK, FR, IT, ES	hip
6		c.350G>A R117H	86/3342 (2.66%)	311/126396 (0.25%)		10.8x (2.74x10 ⁻⁵⁵)	20.29% (9.07x10 ⁻³⁴)	FRA, DE, IT, NL, ES, USA	hip
7		c.2991G>C L997F	26/3342 (0.81%)	315/126490 (0.25%)		3.24 (3.74x10 ⁻⁷)	6.06% (6.96x10 ⁻⁶⁴)	FR, DE, IT, ES, TR, USA	hip
8	GAA	c.-32-13T>G	159/442 (35.97%)	609/114256 (0.53%)	127/89	67.49x (1.11x10 ⁻²³⁰)	18.30% (1.70x10 ⁻²⁵)	AUS, AUT, CAN, DEU, GR, IT, NL	hip
9	GALNS	c.850T>G F284V	1/134 (0.75%)	43/1123588 (0.03%)	26/22	21.45x (0.0466)	1.968% (9.22x10 ⁻⁹)	ENG, TUR	hip
10	NPHS2	c.686G>A R229Q	81/434 (18.66%)	4537/126346 (3.59%)	72/67	5.2x (1.16x10 ⁻³³)	1.66% (1.37x10 ⁻⁷⁶)	USA, AS, CZE, FR, GB, DE, TR	hip
11	PAH	c.117C>G F39L	24/5070 (0.47%)	21/126678 (0.02%)	413/81	28.55x (1.82x10 ⁻²²)	22.41% (7.75x10 ⁻⁰⁶)	AS, AT, DK, DE, PL, SK, ES	hip
12		c.194T>C I65T	78/5070 (1.56%)	68/126520 (0.05%)		28.99x (6.1x10 ⁻⁷¹)	22.47% (8.11x10 ⁻¹³)	AU, AT, GB, FR, DE, IT, PL, PT, SK	hip
13		c.261C>A S87R	5/5070 (0.1%)	17/126672 (0.01%)		7.35x (0.00128)	5.77% (1.86x10 ⁻⁰⁹)	AU, DK, PL, ES	na
14		c.533A>G E178G	14/5070 (0.28%)	19/126496 (0.02%)		18.38x (6.41x10 ⁻¹²)	14.45% (3.46x10 ⁻⁰⁷)	FR, PL, SLO, ES, TR	hip
15		c.688G>A V230I	15/5070 (0.3%)	28/126564 (0.02%)		13.37x (3.25x10 ⁻¹¹)	10.51% (2.23x10 ⁻¹¹)	FR, IT, USA	na
16		c.734T>C V245A	27/5070 (0.53%)	113/126482 (0.09%)		5.96x (4.90x10 ⁻¹²)	4.69% (1.17x10 ⁻⁴⁵)	AT, DK, DE, PL, SK, SLO, ES	hip
17		c.805A>C I269L	4/5070 (0.08%)	16/126598 (0.001%)		6.24x (0.00648)	4.90% (2.20x10 ⁻⁰⁹)	IT, PL, PT, ES	hip
18		c.896T>G F299C	15/5070 (0.3%)	22/126374 (0.02%)		16.69x (2.55x10 ⁻¹²)	13.37% (2.27x10 ⁻⁰⁸)	AU, USA	null
19		c.1139C>T T380M	14/5070 (0.28%)	43/126558 (0.03%)		8.13x (2.50x10 ⁻⁰⁸)	6.39% (1.31x10 ⁻¹⁹)	AU, IT, PL, ES, TR	hip
20		c.1241A>G Y414C	114/5070 (2.31%)	88/126708 (0.07%)		33.23x (1.04x10 ⁻¹⁰⁸)	25.41% (3.32x10 ⁻¹³)	AU, AT, FR, DE, IT, PL, SK, SLO, ES	hip
21	PKHD1	c.4870C>T R1624W	5/580 (0.86%)	28/126656 (0.02%)	220/209	39x (4.13x10 ⁻⁷)	16.96% (4.39x10 ⁻⁰⁵)	FR, NL, USA	na
22		c.5498C>T S1833L	2/580 (0.34%)	37/126386 (0.003%)		11.78x (0.0138)	5.14% (3.15x10 ⁻⁹)	ES, USA	hip
23		c.6992T>A I2331K	8/580 (1.38%)	47/126480 (0.04%)		37.12x (1.81x10 ⁻¹⁰)	16.17% (1.05x10 ⁻⁷)	DK, DE, USA	hip
24		c.7264T>G C2422G	3/580 (0.52%)	78/126580 (0.06%)		8.39x (0.00619)	3.65% (2.43x10 ⁻¹⁸)	AT, DE	na
25		c.10036T>C C3346R	2/580 (0.34%)	39/126522 (0.03%)		11.19x (0.0152)	4.87% (8.72x10 ⁻¹⁰)	ES	null

AC, allél count/allélszám; AF, allélfrekvencia; AT, Ausztria; AU, Ausztrália; bet, beteg; CAN, Kanada; CZE, Csehország; DE, Németország; DK, Dánia; GB, Nagy Britannia; ES, Spanyol; FR, Franciország; GR, Görögország; HR, horvát; hip, hipomorf; HU, Magyarország; IT, Olaszország; LOF, loss-of-function; na, nincs adat; NL, Hollandia; PL, Lengyelország; PT, Portugália; SK, Szlovákia, SLO, Szlovénia; SRB, Szerbia; TR, Törökország; USA, Amerikai Egyesült Államok

6.4. Populációgenetikai módszer validálása

A biológiai és a klinikai beosztás során igazolódott, hogy az IP variánsok hipomorfak, míg a teljesen penetráns variánsok döntően null mutációk, alátámasztva ezzel módszerünk jó specificitását. Hasonlóan a klinikai beosztás eredményeihez, a biológiai beosztás is jó specificitást mutatott: míg az IP variánsok közül egy sem, addig a teljesen penetráns variánsok 34,4%-a volt LOF mutáció (11. táblázat).

11. táblázat: A populációgenetikai megközelítés validálása klinikai és biológiai értékeléssel. Mind a klinikai (hipomorf/null), mind a biológiai megközelítés (misszensz/LOF) összhangban volt a populációgenetikai módszer eredményével (44). Az IP variánsok döntően hipomorf és misszensz variánsok.

	hipo	null	p (hipo vs. null)	na
ip (25)	16	2	5,12x10 ⁻¹⁰	7
kp (1660)	177	756		727
	missz	LOF	p (LOF vs. missz)	más
ip (25)	23	0	2,86x10 ⁻⁰⁵	2
kp (1660)	924	572		164

hipo, hipomorf; ip, inkompletten penetráns; kp, teljesen penetráns; LOF, loss-of-function; missz, misszensz; na, nincs adat

6.5. Interallelikus interakciók azonosítása

Speciális társulásokat kerestünk az inkompletten penetráns variánsoknál igazolva, hogy az alulreprezentáltság háttérében interallelikus interakció áll.

A vizsgált IP variánsok közül egyetlen, az *NPHS2* R229Q variánsa felelt meg azoknak a kritériumainknak, melyek alapján az inkomplett penetrancia háttérében domináns negatív hatás áll. Domináns negatív hatás-függő variáns az amelyik: (1) gyakrabban társul misszensz variánsokhoz, mivel interallelikus interakció olyan fehérje variánsok között lehetséges, melyek nem vezetnek a fehérje teljes lebomlásához; (2) olyan variánsokhoz társul, amelyek egy adott fehérje-doménben helyezkednek el; (3) gyakrabban társul bizonyos specifikus variánsokhoz, melyek domináns negatív hatással vannak rá. Vizsgálatunk során az *NPHS2* R229Q többször társult misszensz variánsokkal (12. táblázat). Olyan variánsokkal fordult elő a beteg populációban melyek döntően a 4-es

doménban lokalizálódtak. Hét olyan variánst igazoltunk, melyek domináns negatív hatással vannak rá: A284V, A288T, A297V, E310K, E310V, Q328R, F344Lfs*4

Speciális társulásait a 12. táblázatban részletezem.

12. táblázat: Az *NPHS2* R229Q variáns: asszociációs tesztek

Az *NPHS2* R229Q variánsa szignifikánsan gyakrabban társult misszensz mutációkkal mint LOF mutációkkal, illetve szignifikánsan gyakrabban társul olyan mutációkhoz amelyek a 4-es fehérje doménban vannak, igazolva ezzel az interallelikus interakció szerepét (44). Hét variánssal társult szignifikánsan többször.

	R229Q-asszociált allélok ($\Sigma=89$)	nem-R229Q-asszociált allélok ($\Sigma=527$)	p
misszensz	75	399	6,05x10 ⁻⁵
LOF	2	94	
osztályozatlan	12	34	
1-3 domén (AA 1-282)	22	376	3,4x10 ⁻³⁴
4 domén (AA 283-382)	53	23	
osztályozatlan	14	128	
A284V	34	10	4,41x10 ⁻²³
A288T	5	0	0,000057
A297V	6	0	7,8x10 ⁻⁶
E310K	7	0	1x10 ⁻⁶
E310V	4	0	0,0004
Q328R	4	0	0,0004
F344Lfs*4	6	0	7,8x10 ⁻⁶

AA, aminosav;

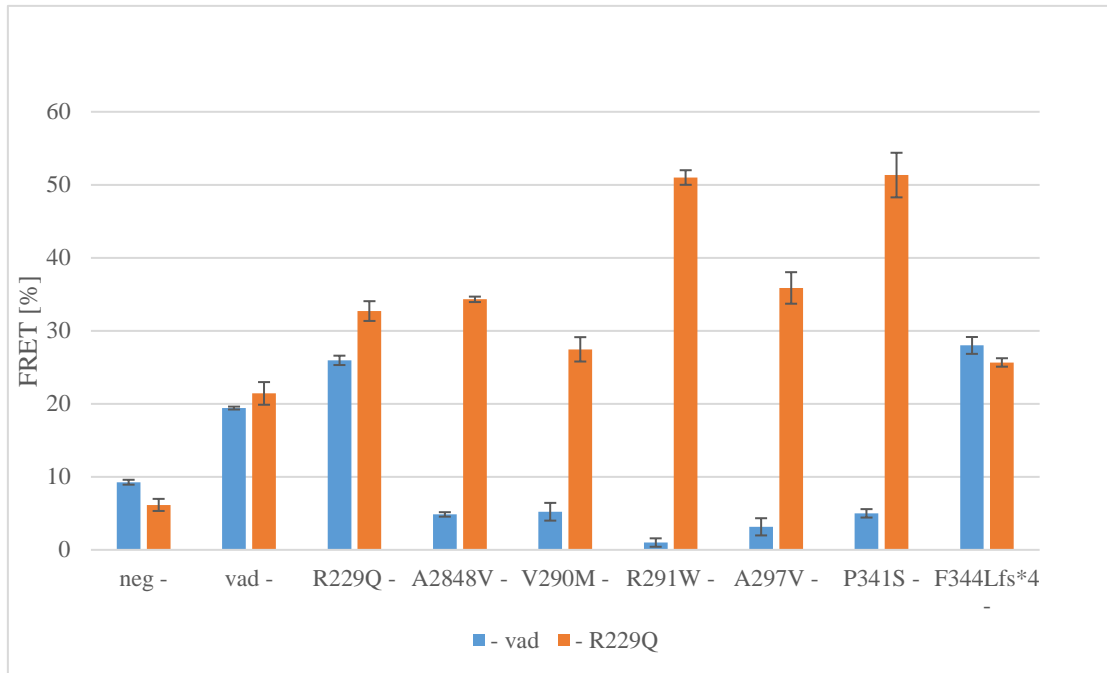
6.6. A podocin variánsok vizsgálatának eredményei

A vad és R229Q podocin homo- és heterooligomerjeiben mért FRET hatékonyságot vizsgáltuk.

Összehasonlítva a vad és R229Q homo- és heterooligomereket, az R229Q társulások esetében szignifikánsan magasabb ($p = 2,7 \times 10^{-07}$), 20% feletti FRET hatékonyságot mértünk (9. ábra). Ezen különbség alapján következtethetünk arra, hogy az R229Q társulások esetén, a korábbi *in silico* számításokkal összhangban, valóban megváltozik az oligomerek konformációja.

Összehasonlítva a természetben előforduló patogén társulásokat (R229Q-A284V, R229Q-A297V, R229Q-R291W, R229Q-P341S, R229Q-F344Lfs*4) a nem patogén társulásokkal (vad-vad, R229Q-vad, R229Q-V290M) azt láttuk, hogy a patogén társulások esetén szignifikánsan magasabb volt a FRET hatékonyság, mint a nem patogén asszociációk esetében ($p = 0,0029$) (10. ábra).

Mind az R229Q variáns, mind a domináns negatív hatást kifejtő variánsok tehát növelik a FRET hatékonyságot, ami az oligomerek szerkezetében lévő változást tükrözi.



9. ábra: FRET-hatékonyság az R229Q és vad podocinok homo-és heterooligomerjeiben. Az R229Q társulások FRET-hatékonysága nagyobb a vad homo- és heterooligomerek FRET hatékonyságánál ($p = 2,7 \times 10^{-07}$).



10. ábra: Az R229Q variáns patogén és nem patogén társulásai
A patogén társulások (R229Q-A284V, R229Q-A297V, R229Q-R2291W, R229Q-P341S, R229Q-F344Lfs*4) esetében szignifikánsan magasabb FRET hatékonyságot

tapasztaltunk ($p = 0,0029$), a nem patogén társulásokkal szemben (R2290-vad. R229Q-R229Q, R229Q-V290M)

6.7. Klinikai irányelv az *NPHS2* R229Q asszociációk patogenitásának értelmezésére

Az adatbázisunkban elérhető *NPHS2* variánsokat hordozó beteg populáció alapján kiszámoltuk mely variánsokkal társulva patogén, illetve benignus az R229Q variáns. Ezen eredményeket kiegészítve a már kutatócsoportunk által igazolt eredményekkel (36, 37) egy útmutatót hoztunk létre, mely segíti a genetikai tanácsadást.

A fent részletezett módszerrel megvizsgáltuk az *NPHS2* R229Q variánsával társuló variánsainak előfordulását az [R229Q];[mut] betegek és az R229Q variánst nem hordozó [mut];[mut] betegek között. A domináns negatív hatás bizonyítható a mutációk [R229Q];[mut] betegek közötti szignifikáns dúsulása alapján, a [mut];[mut] betegekhez képest (14 táblázat).

Ezzel a módszerrel a következő variánsok domináns negatív hatását bizonyítottuk: A284V, A288T, R291W, A297V, E310K, E310V, L327F, Q328R, és F344Lfs*4 (13. táblázat). Ezen domináns negatív mutációk átlagpopulációs előfordulási aránya alacsony, a várt előfordulási arányuk az R229Q-val társulva egy populációban sem haladja meg a 1.3×10^{-6} -t.

A fenti módszerrel az R229Q-val nem patogén variánsok is felderíthetőek. Azon variánsok, melyek szignifikánsan alulreprezentáltak az [R229Q];[mut] beteg között, a [mut];[mut] betegekhez képest, valószínűtlen, hogy az R229Q-val patogének legyenek. Három ilyen variáns bizonyult szignifikánsan alulreprezentáltnak az R229Q variánssal: R138Q, L156Ffs*11 és R168H (13. táblázat).

Azon variánsoknak is valószínűtlen a patogenitása az R229Q variánssal, melyeknek magas az allélfrekvenciája az európai kontrollpopulációban és egyszer sem társul az R229Q variánshoz, vagy alulreprezentált a társulásuk a beteg populációban : I263L, E264Q, G273W, R291Q, V319L, Q320P, K377R, R286Tfs*17, V290M, A317Lfs*31 (13. és 14. táblázat).

13. táblázat: A benignus és patogén R229Q-asszociációk, populációgenetikai módszer és családvizsgálatok alapján (83).

R229Q variánsal társuló mutációk			AC betegpop. [RQ];[mut] Σ=106 allél 106 bet.	AC betegpop. [mut];[mut]. Σ=878 allél 439 bet.	dúsulás a [RQ];[mut] vs. [mut];[mut]. betegpop. $[AF_{[RQ];[mut]}/(AF_{[mut];[mut]})] (P^t)$	F_{max} of pts w [RQ];[mut] (régio) ^s	Életkor dg/ESRD (év)	R229Q-val asszoc.
nukleotid	aminosav	PolyPhen-2 score						
c.413G>A	R138Q	0.999	6	281 (182beteg)	0.18x (3.9×10^{-7})	1/12k (Eu)	3.8-34/34-40	ben
c.412C>T	R138*	ND	0	24 (15beteg)	0x (0.16)	1/780k (Eu)	NA	ben
c.467dupT	L156Ffs*11	ND	0	59 (41beteg)	0x (0.002)	1/65k (Eu)	NA	ben
c.503G>A	R168H	0.999	0	33 (21beteg)	0x (0.04)	1/650k (Eu)	NA	ben
c.534+1G>T	abnormal transcript	ND	0	4 (3beteg)	0x (1)	0	NA	ben
c.714G>T	R238S	0.998	1	6 (4beteg)	1.4x (0.5)	1/570k (SAs)	ND	ben
c.851C>T	A284V	0.998	34	10 (5beteg)	28x (6.6×10^{-23})	1/1.5M (Eu)	0-39/8-50	pat
c.855_856delAA	R286Tfs*17	ND	3	23 (17beteg)	1.1x (0.75)	1/97k (Eu)	20-27/30-35	ben
c.862G>A	A288T	0.999	5	0	NA - dúsul (1.6×10^{-5})	1/1.5M (Eu)	12-28/9-40	pat
c.868G>A	V290M	0.989	2	22 (18beteg)	0.75x (1)	1/55k (Eu)	1.5-2.3/ND	ben
c.871C>T	R291W	0.978	5	3 (3beteg)	14x (0.0007)	1/15M (EA)	0-3.5/10-34	pat
c.890C>T	A297V	0.627	6	0	NA - dúsul (2×10^{-5})	1/1.4M (Afr)	4.3-34/3.6-31	pat
c.928G>A	E310K	0.936	7	0	NA - dúsul (2×10^{-7})	0	0.6-30/3.2	pat
c.929A>C	E310A	0.837	2	0	NA - dúsul (0.01)	0	4.1-18/34	pat
c.929A>T	E310V	0.960	4	0	NA - dúsul (0.0001)	0	0.2-5/8.5-11.9	pat
c.979C>T	L327F	0.997	4	0	NA - dúsul (0.0001)	1/780k (Eu)	3-25/29	pat
c.983A>G	Q328R	0.108	4	0	NA - dúsul (0.0001)	0	6-10/ND	pat
c.1032delT	F344Lfs*4	ND	6	0	NA - dúsul (2×10^{-5})	0	0.6-17.6/ND	pat
c.376_377insT	K126Ifs*7	ND	1	0	NA - dúsul (0.1)	1/570k (SAs)	ND	ben
c.451+3A>T		ND	1	1	8.3 (0.2)	1/630k (Latin)	14.8/ND	ben
c.643C>T	Q215*	ND	1	7	1.2 (0.6)	1/590k (Eu)	0.8/43.5	ben
c.810G>T	L270F	0.925	1	0	NA - dúsul (0.1)	0	38/ND	pat
c.826_833dup CACTCACT	A279Tfs*17	ND	1	0	NA - dúsul (0.1)	0	8/ND	ben

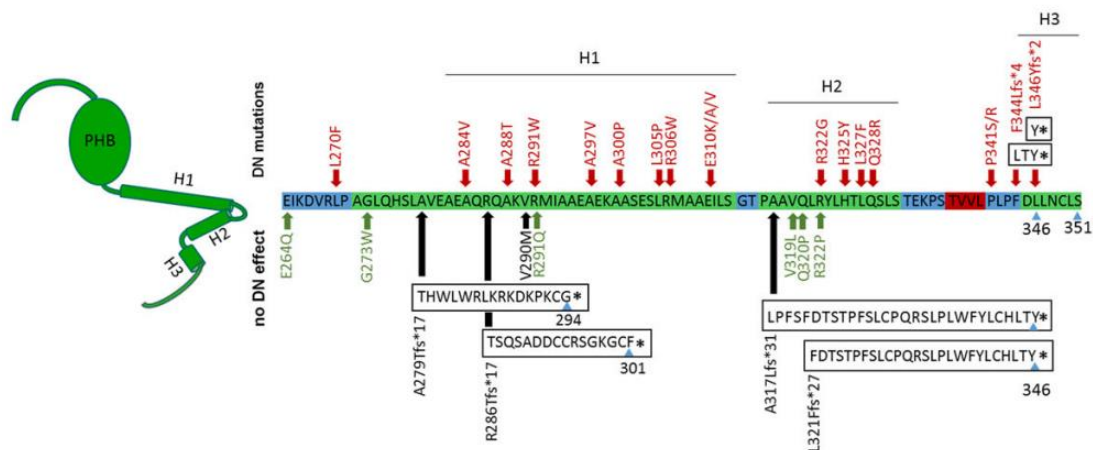
c.841G>A	E281K	0.991	1	0	NA - dúsul (0.1)	0	testvér: I. 3/ND	pat
c.883G>A	A295T	1.0	1	2 (1bet)	4.1x (0.3)	0	4/ND	pat
c.898G>C	A300P	0.953	1	0	NA - dúsul (0.1)	0	testvér: I. 60/ND II. 16.2/ND	pat
c.914T>C	L305P	0.996	1	0	NA - dúsul (0.1)	0	ND	pat
c.916A>T	R306W	0.985	1	0	NA - dúsul (0.1)	0	testvér I. 14.5/ND II. ND/ND	pat
c.961delC	L321Ffs*27	ND	1	0	NA - dúsul (0.1)	0	13.3/ND	ben (?)
c.964C>G	R322G	1.0	1	2 (1bet)	4.1x (0.3)	1/210k (Eu)	34/38	pat
c.965G>C	R322P	1.0	1	0	NA - dúsul (0.1)	1/140k (SAs)	1.6/ND	ben (?)
c.973C>T	H325Y	0.885	1	0	NA - dúsul (0.1)	0	18.8/21.2	pat (?)
c.1021C>T	P341S	0.995	1	0	NA - dúsul (0.1)	0	8/ND	pat
c.1022C>G	P341R	0.997	1	0	NA - dúsul (0.1)	0	6/19	pat
c.1036delC	L346Yfs*2	ND	1	0	NA - dúsul (0.1)	0	3 testvér 15-28/29-32	pat

AC, allélszám; AF, allélfrekvencia; Afr, Afrika; asszoc, asszociáció; bet, beteg; dg, diagnózis; EA, East Asia/Kelet Ázsia; ESRD, end-stage renal disease/végstádiumú veseelégtelenség; Eu, Európa; F_{max}, legmagasabb várt allélfrekvencia; k, ezer; M, millió; mut, mutáció; NA, nem alkalmazható; ND, no data/nincs adat; pat, patogén; RQ, R229Q; Sam, South America/Dél Amerika; Sas, South Asia/Dél Ázsia; SSNS, szteroid-szenzitív nephrosis szindróma; †Fisher-exact teszt szignifikancia szint; §gnomAD (43), (?), bizonytalan predikció;

14. táblázat: Gyakori 3' végű variánsok, melyek alulreprezentáltak az európai [p.R229Q];[mut] betegek között, az európai kontrollpopulációhoz képest: R229Q variánssal társulva nem patogének (83).

		PolyPhen-2 score	allélszám		p [‡] vs. DN mutációk
			európai betegek	európai kontrollpop.	
Domináns negatív mutációk		na	54	5	
Gyakori 3' végű variánsok DN hatás nélkül	I263L	0.009	0	7	1x10 ⁻⁶
	E264Q	0.962	0	31	2.9x10 ⁻¹⁹
	G273W	0.968	0	3	0.001
	R286Tfs*17	na	3	18	9.9x10 ⁻¹¹
	V290M	0.989	2	32	2.4x10 ⁻¹⁷
	R291Q	0.605	0	10	8.8x10 ⁻⁹
	A317Lfs*31	na	0	3	0.001
	V319L	0.002	0	3	0.001
	Q320P	0.473	0	3	0.001
	K377R	0.003	0	22	2.2x10 ⁻¹⁵

DN, domináns negatív; na, nincs adat; pop, populáció; [†]gnomAD (43), [‡]Fisher-exact teszt – szignifikancia szint. A félkövér betűvel jelölt variánsok homozigóta formában patogének.



A 11. ábrán a podocin struktúrájának megfelelően, lokalizáció alapján tüntettük fel az R229Q variánssal patogén/domináns negatív és benignus variánsokat.

11. ábra: A domináns negatív és benignus variánsok lokalizációja (83): A podocin szakaszt a 264-es aminosav résztől a 351-es aminosav részig tüntettük fel. A podocin struktúráit különböző színnel jelöltük: (loop) hurok szerkezetet - kézzel, hélixet – zölddel, TVV motívumot – pirossal. A domináns negatív hatású mutációkat pirossal jelöltük a podocin szakasz felső részén, lokalizációjuknak megfelelően. Fekete színnel jelöltük a szekvencia alsó részén azon variánsokat, melyek homozigóta formában patogének, de az R229Q-val társulva nem okoznak betegséget. Zöld színnel jelöltük azon variánsokat, melyek sem homozigóta formában, sem R229Q-val társulva nem okoznak betegséget.

A bevezetőben részletezett korábbi eredményeink, családvizsgálatok és populációgenetikai módszer összegzéseként, az a variáns patogén az R229Q variánssal, amely megfelel a következő kritériumoknak (83):

1. A variáns megfelel a patogenitás standard kritériumának, azaz evolúciósan konzervált aminosavat érint.
2. Az érintett aminosav az oligomerizációért felelős régióban helyezkedik el (aminosavak 270-351) és az aminosavcsere következtében megváltozik az oldallánc szerkezete, mérete vagy polaritása.
3. Az oligomerizációt nem gátolja meg. Ahogy ez az A279Tfs*17 variáns esetében történik, mely mindkét oligomerizációs részt érinti, így várhatóan nem patogén az R229Q variánshoz társulva. Az L321Ffs*27 és a A317Lfs*31 variánsok a második oligomerizációs részt érintve szintén nagy valószínűséggel benignusak az R229Q variánssal.
4. A [R229Q];[mut] személyek várt frekvenciája a kontrollpopulációban kevesebb mint $1:10^6$
5. Dúsul a [R229Q];[mut] genotípusú beteg között, a [mut];[mut] genotípusú betegekkel szemben.
6. [R229Q];[mut] asszociáció szegregálódik a családban.
7. A klinikum megfelel az R229Q-asszociált nephropathiának, kifejezett ödéma nélkül, szövettanilag FSGS és lassú progresszió jellemzi (ESRD 10-50 éves életkor között).

7. Megbeszélés

Jelen dolgozatban, a korábban kutatócsoportunk által igazolt inkompletten penetráns *NPHS2* R229Q variáns példájára alapozva, célunk volt további inkompletten penetráns variánsokat és interallelikus interakciókat azonosítani gyakori AR betegségekben. Korábban az R229Q variáns interallelikus interakcióit populációgenetikai megközelítésben azonosította a kutatócsoportunk (36). Ezen módszer kiterjesztése és pontosítása révén szerettünk volna további variánsokat azonosítani. Ezt az tette lehetővé, hogy az irodalomban növekvő számban rendelkezésre állnak nagy esetszámú kohorszvizsgálatok, illetve kontroll populációként elérhetővé vált a gnomAD adatbázis.

Az általunk vizsgált gének nem szerepelnek kóroki génként a gnomAD kontroll populációjában exom vagy genom szekvenálással azonosított multifaktoriális gének között. Azon személyeket, akik súlyos gyermekkori betegségben szenvednek, kizárták a vizsgálatból. Ezen okok miatt a gnomAD allélgyakorisága egy megbízható kontroll csoport a vizsgált AR betegségekben.

A patogén mutációkat tartalmazó HGMD adatbázis segítségével a szakirodalomból, manuálisan létrehoztunk egy adatbázist mely összesen 17 gyakori AR betegségben szenvedő 12,048 beteg genotípus és fenotípus adatait tartalmazza. Kizárólag olyan tanulmányokat vettünk figyelembe, ahol a gén teljes szekvenálása történt, és nem célzott szűrés egy adott mutációra. A szakirodalom gyors elemzésére egy számítógépes algoritmust dolgoztunk ki.

A populációgenetikai módszer lehetővé tette, hogy összehasonlítsuk a variánsok allélfrekvenciáját az általunk létrehozott beteg populációban és a gnomAD kontrollpopulációjában. Ezáltal azonosítottuk a HGMD-ben tévesen patogénnek leírt variánsokat, kiszűrtük azt a gént, melyben a LOF mutációk in utero letalitást okoznak, meghatároztuk a variánsok penetranciáját és interallelikus interakciókat vizsgáltunk az inkomplett penetrancia hátterében. Módszerünk alapelve az volt, hogy a variánsok LOF variánsokhoz viszonyított, beteg- és kontroll populációs gyakorisága (relatív allélfrekvenciája) alapján ítéltük meg a variáns penetranciáját.

A populációgenetikai megközelítésnek azonban számos korlátozása és kihívása van, amit figyelembe kellett vennünk az algoritmus kidolgozása során:

[1] Az AR betegségekért felelős mutációk különböző földrajzi eloszlása torzíthatja az eredményeket. Ennek a problémának a kiküszöbölésére, csak a nem-finn európai személyek adatait gyűjtöttük, illetve csak olyan gyakori betegségeket vizsgáltunk, amelyek legalább 100 különböző variánssal és magas kumulatív LOF allélszámmal rendelkeztek.

Azonban még a legnagyobb referencia adatbázisban is, amely több mint 63.000 európai ember szekvencia adatait tartalmazza, van aránytalanság: a gnomAD populációjának 20%-át svédek teszik ki, annak ellenére, hogy az európai lakosság 2%-át képviselik (43). Ennek következményeként a skandináv variánsok gyakrabban előfordulhatnak a kontrollpopulációban, így tévesen inkompletten penetráns variánsnak ítélnélhetjük őket. Ennek ellenére, még az *NPHSI* Finn major és Finn minor, ismerten gyakori skandináv variánsok sem jelentek meg az IP variánsok között.

[2] A beltenyészet következtében a homozigóta betegek túlzott gyakorisága torzíthatja a mutációk allélfrekvenciáját a beteg populációban. Ennek tudatában, kiszámoltuk a beltenyésztési együtthatót minden egyes variáns esetében (78), így korrigálva a homozigóta betegek számát.

[3] A HGMD adatbázisában több olyan nem patogén variáns van, amiket tévesen patogénként publikáltak. Ezen variánsok tévesen IP variánsnak ítélnélhetők populációgenetikai módszerrel. Újraértékeltek a HGMD-ben szereplő összes variáns patogenitását és csak azokat vizsgáltuk, amelyek szignifikánsan dúsultak a beteg populációban. Ha egy variáns kevesebbszer fordult elő a beteg populációban, mint a kontrollcsoportban nem-patogén variánsnak ítéltük (n=25) és kizártuk a további vizsgálatokból. Ezen variánsok döntően *CFTR* és *PKHD1* gének mutációi voltak (19/25), egyik sem volt LOF mutáció. Azokat a mutációkat is kizártuk a vizsgálatból, amelyek nem szignifikánsan dúsultak a beteg populációban (111/2211, 5,02%), annak ellenére, hogy néhány közülük LOF mutáció volt (42/111, 37,8%). Ezen variánsok patogenitása nem igazolható populációgenetikai módszerrel.

[4] Egy variáns relatív allélfrekvenciáját torzíthatják a beteg populáción belül más IP variánsok, illetve a kontrollpopulációban nem felismert patogén variánsok. Ezek kiküszöbölésére a penetrancia megítéléséhez a variánsok allélfrekvenciáját a LOF

variánsokéhoz hasonlítottuk, melyek teljesen penetráns variánsok, és nem lehetnek interallelikus interakciók áldozatai.

Amennyiben a LOF variánsok in utero letalitást okoznak, alulreprezentáltak lesznek a beteg populációban, ezzel torzítva a LOF mutációk arányát. Ahhoz, hogy kiküszöböljük ezt a torzító hatást, azokat a betegségeket, ahol a LOF mutációk in utero letalitást okoznak, kizártuk a vizsgálatból. Ennek kiszűrésére két módszert alkalmaztunk: összehasonlítottuk a misszensz és a LOF mutációk arányát a beteg és kontrollpopulációban, illetve a beteg populáción belül a misszensz és a LOF mutációk Hardy-Weinberg ekvilibrumát vizsgáltuk meg. Mindkét megközelítés igazolta azt, hogy a Smith Lemli Opitz szindróma LOF variánsai alulreprezentáltak a beteg populációban ($7,7x$, $p = 9,45 \times 10^{-76}$), rámutatva, hogy a biallelikus LOF *DHCR7* mutációk in utero letalitáshoz vezetnek az esetek legalább 87%-ában. Az irodalomban korábban már felmerült ez a hatás bizonyos *DHCR7* gén LOF mutációi esetében, illetve hasonlóan magas halálozási arányt számoltak más vizsgálatok is (80, 81, 84, 85). A *DHCR7* gén variánsait kizártuk a további vizsgálatokból [5].

Az inkomplett penetrancia csak akkor mutatható ki, ha egy variáns kellően gyakori a kontrollpopulációban ahhoz, hogy szignifikánsan kevesebbszer forduljon elő a beteg populációban. A szignifikancia szintet Bonferroni korrekcióval ítéltük meg ($p < 7,22 \times 10^{-05}$). Összesen 692/1936 (35,74%) variáns szerepelt a gnomAD adatbázisában, melyből 85 variáns felelt meg a fent említett frekvencia kritériumnak.

A 85 gyakori variáns magas allélgyakorisága részben ellensúlyozta az alapító hatásukból eredő torzítást: a gyakori variánsok jellemzően Európa kolonizációjának elején jelentek meg, így jól eloszlanak az európai származású személyek között.

A 85 variáns közül 25 (29,4%) bizonyult inkomplett penetránsnak, köztük a három korábban már ismert IP variáns: *NPHS2* R229Q, *CFTR* R117H és L997F.

Az *NPHS2* R229Q variánsra számított penetrancia (1,4%) tökéletesen megfelel a hiányzó betegek korábbi becslésének (36).

Az *CFTR* R117H variáns penetranciáját (20,3%), a korábban becsült penetranciánál (3,1%) (9) magasabbnak számoltuk. A klasszikus fenotípusú CF betegek alcsoportjában alacsony penetranciájúnak találtuk (2,6%). A *CFTR* L997F variánsát korábban már leírták olyan homozigóta és összetett heterozigóta egyéneknél, akik nem voltak érintettek (82).

Módszerünk ellenőrzése céljából összehasonlítottuk a klinikai (null/hipomorf) és biológiai (misszensz/LOF) jellemzőit az inkompletten és a teljesen (>70%) penetráns variánsok között. A jó korreláció alapján a populációgenetikai módszerünk, a fent felsorolt limitáló tényezők ellenére, jó specificitást mutatott.

A 25 azonosított IP variáns közül 22-ről ezidáig nem volt ismert, hogy inkompletten penetráns. A variánsok döntően (22/25) nagyon szignifikánsan ($p < 10^{-4}$), és/vagy jelentősen dúsultak ($\geq 10x$) a beteg populációban, erősen alátámasztva kóroki szerepük. Az IP variánsok fele (13/25) nagyon alacsony penetranciával rendelkezik ($\leq 10\%$), így klinikai szempontból releváns az ismeretük.

A legtöbb új IP variánshoz kapcsolódó fenotípus összhangban van az inkomplett penetranciájukkal: *ASL* R12Q (22, 63) *CAPN3* A798E (86) *CFTR* R117C (87) *GAA* c.-32-13T>G (88), *GALNS* F284V, *PAH* I65T (89) T380M (90) és a *PKHDI* I2331K (91) variánsokat különösen enyhe fenotípussal rendelkező betegeknél írták le. Ezen variánsok inkomplett penetranciája döntően hipomorf jellegükből adódik.

Egy variáns alulreprezentáltsága adódhat abból is, hogy a betegség megjelenése nem jellemző, nehezítve ezzel a megfelelő klinikai és genetikai diagnózist. Igaz ez két *PKHDI* gén esetében igazolt IP variánsra (S1833L; I2331K). Leírtak olyan beteget, aki összetett heterozigóta volt a *PKHDI* V1741M és S1833L variánsaira, panaszai cholangitissel kezdődtek 38 évesen, majd kolesztatikus májfibrózissal és cisztás medulláris vesebetegséggel diagnosztizálták (92). Hasonló, nem specifikus fenotípust mutatott egy olyan beteg, aki a *PKHDI* I2331K és T36M mutációt hordozta és 5 évesen kis méretű veséi voltak (jobb: -1.8 SD, bal: -2.1 SD), normál vesefunkció mellett (93).

Bizonyos mutációk funkcionális hatása szintén összhangban van az inkomplett penetranciájukkal: a gyakori *GAA* c.-32-13T>G variánst hordozó betegek esetében normálisan átíródó mRNS-t detektáltak fibroblasztokban (94), illetve a *PAH* I65T variáns esetében *in vitro* vizsgálattal kimutatták, hogy interallelikus interakció áldozata (20).

Az azonosított *PAH* IP variánsok BIOPKU (<http://www.biopku.org/>) curated by N. Blau)) adatbázisában elérhető enzim aktivitása, szintén reziduális aktivitásra utalnak: F39L: 49%, I65T: 25%, S87R: 25%, E178G: 39%, V230I: 63%, V245A: 50%, T380M: 28%, Y414C: 36%.

Ezzel szemben a homozigóta formában súlyos PKU fenotípust okozó (89) F299C variáns enzimaktivitása jelentősen alacsonyabb (3%), valószínűtlenné téve, hogy ezen variáns

inkompletten penetráns, és rámutatva arra, hogy a populációgenetikai megközelítés is tévedhet. Fontos ezért az eredményeket óvatosan, más megközelítésekkel együtt értékelni.

Vizsgálatunk során 1936 variáns szisztematikus elemzésével megállapítottuk, hogy a leggyakoribb 85 variáns közül 25 inkompletten penetráns. Az IP variánsok ennyire magas aránya váratlan volt.

Az interallelikus interakciók azonosításához nagyszámú összetett heterozigóta beteg szükséges, így a ritka interallelikus interakciók könnyen észrevehetetlenek maradnak módszerünkkel. Interallelikus interakció háttérében állhat interallelikus komplementáció vagy domináns negatív hatás. Ezen jelenségek azonosítása céljából megvizsgáltuk az inkompletten penetráns variánsok társulásait.

Domináns negatív hatást az inkomplett penetrancia háttérében egy variánsnál, az *NPHS2* R229Q esetében találtunk. Számítógépes algoritmusunk eredményességét igazolja, hogy korábban ezen variáns inkomplett penetranciáját családvizsgálatokkal, populációgenetikai számításokkal, *in vitro* és *in silico* vizsgálatokkal igazoltuk (36, 37). Eredményeink alapján azonban az interallelikus interakciók összességében ritka okai az inkomplett penetranciának.

Az R229Q interallelikus interakcióinak molekuláris szintű megértése céljából korábban kutatócsoportunk fehérjelokalizációs kísérletekkel vizsgálta a különböző podocin társulásokat, illetve a FRET-mérés lehetőséget adott a heterooligomerek szerkezetének közvetett vizsgálatára.

Eddigi eredményeink alapján tehát tudtuk, hogy a podocin kizárólag a C-terminális régióban dimerizálódik, valamint szerkezeti modellek alapján láttuk, hogy az R229Q rigidebbé teszi a C-terminális helikális régió és a globuláris régió kapcsolatát, ami hajlamossá teszi az abnormális heterodimerek képződésére az oligomerizációs régiót érintő podocin variánsokkal. Ezen vizsgálatok adták az alapját a patogén és nem patogén oligomerek további vizsgálatának (37).

Jelen dolgozatban az oligomerizáció és a patogenitás összefüggésének további megértése céljából FRET- hatékonyság módszerével olyan podocin variánsokat vizsgáltunk, melyekkel társulva az R229Q variáns betegséghez vezet (A297V, A284V, R291W,

P341S, F344Lfs*4), illetve a V290M podocin variánst, mely önmagában patogén, azonban az R229Q variánshoz társulva nem.

Korábbi fehérjelokalizációs vizsgálatok alapján tudjuk, hogy az R229Q podocin az A297V, A284V és F344Lfs*4 podocin variánssal együtt expresszálva a citoplazmában reked és nem jut ki a sejtmembránba, ezzel ellentétben a V290M podocin variánssal való koexpresszió során a membránban lokalizálódott.

A FRET-mérés során azt láttuk, hogy az R229Q homo- és heterooligomerek szignifikánsan ($p = 2,7 \times 10^{-07}$) nagyobb FRET jelet adtak (>20%), mint a vad heterooligomerek. Ezen eredmények igazolták, hogy az R229Q megváltoztatja a oligomerek konformációját. Az, hogy a patogén társulások szignifikánsan nagyobb FRET jelet adtak ($p = 0,0029$), mint a nem patogén társulások, alátámasztja, hogy az interallelikus interakció háttérében egy kóros konformáció áll.

A R229Q variáns patogén és nem patogén társulásainak megkülönböztetése fontos, viszont kihívást jelent a klinikumban: az *NPHS2* asszociált SRNS/FSGS fenotípusa nem specifikus. Az *NPHS2* variánsai a 25 év alatt SRNS/FSGS-el diagnosztizált betegek csupán 6,7%-át (120 / 1783) okozzák, és az R229Q társulások csak az FSGS miatt gondozott betegek 1,18% -ában kórosiak (21/1783) (95). Tehát 100 SRNS/FSGS beteget kell szűrni ahhoz, hogy egy patogén R229Q asszociációt azonosítsunk.

A nem specifikus fenotípus miatt az *NPHS2* gént gyakran szekvenálják a klinikai gyakorlatban. Az R229Q magas, több százalékos gyakorisága több populációban (Európában, Dél-Ázsiában és Dél-Amerikában) (gnomAD) (43) megteremti az alapját annak, hogy benignus társulásait véletlenszerűen azonosítsák a betegekben. Fontos ezért az R229Q benignus és patogén társulásainak az elkülönítése a klinikai gyakorlatban.

A fent ismertetett vizsgálatok lehetővé tették az interallelikus interakció molekuláris alapjának megértését az *NPHS2* R229Q variánsa esetében (36, 37).

A családvizsgálatok, a populációgenetikai számítások, kiegészítve az *in silico* és *in vitro* vizsgálatokkal az esetek többségében alkalmasak az R229Q variáns asszociációk patogenitásának megítélésére.

Genetikai tanácsadást ezáltal segítve, létrehoztunk egy klinikai útmutatót mely tartalmazza a gyakori R229Q asszociációk jellemzőit és patogenitását, illetve irányelvet mutat a ritka társulások megítélésére.

Az irányelvek közé tartozik az a megállapítás, miszerint azon R229Q variánshoz társuló variáns patogén, mely az oligomerizációs részen elhelyezkedő aminosavakat érinti (270 – 351 aminosavak).

Vannak azonban kivételek ez alól, mint ahogy azt az eredményekben részletezett 10. ábra mutatja: nem minden L270-S351 podocin szakaszon levő variáns patogén az R229Q variánssal. Fehérjeszerkezeti modelleken vizsgálva az aminosavcserék töltése, mérete, oldallánc szerkezete, illetve térszerkezetben való elhelyezkedése befolyásolja a patogenitásukat. A H1 és a H2 hélixek mentén a podocin dimereket képez, és így egy szuper hélixet alkot. Azon aminosavak, amelyeknek az oldalláncai kifelé, és nem az interakciós felszín felé tekintenek, azoknak az aminosavaknak a cseréje nem lesz patogén. Igaz ez a V290M variánssra, mely annak ellenére, hogy az első oligomerizációs részen helyezkedik el, nincs domináns negatív hatással az R229Q variánssra.

A G273W és V319L benignus hatása, a V290M-hez hasonlóan, azzal magyarázható, hogy nem a dimerizációs/oligomerizációs rész fele tekintenek (83).

Kivétel ezalól az R291Q variáns, mely bár az interakciós részre esik, nem befolyásolja érdemben az interakciós készsége. Ezen aminosavcsere esetében az Arg-Gln csere nem jár érdemi szerkezeti változással, ellenben az R291W variánssal, mely a gyűrűs oldallánc megjelenését jelenti egy egyszerű szénlánc helyett, így fejt ki domináns negatív hatást.

Azon trunkáns, L270-S351 régióban levő mutációk sem patogének az R229Q variánssal, melyek nem képeznek oligomereket. Ilyenek az R286Tfs*17 és a A317Lfs*31 variánsok, melyek vagy a H1 (p.R286Tfs*17) vagy a H2 (A317Lfs*31) régiót érintik, ezáltal érintve mindkét, vagy csak a második oligomerizációs részt (11. ábra). Ennek megfelelően az R286Tfs*17 podocin monomereket alkot, míg az A317Lfs*31 podocin várhatóan csak dimereket. Ez lehet a magyarázata annak, hogy nem gátolják az R229Q podocin sejtmembránba való kijutását (37).

Az F344Lfs*4 és az L346Yfs*2 variánsok a második oligomerizációs részt érintik, de nem gátolják az oligomerizációt, így kifejtethetnek domináns negatív hatást az R229Q-val szemben.

A patogenitás megállapításában az elsődleges populációgenetikai megközelítés mellett a családvizsgálatok, a biokémiai megítélés és a klinikum is fontos támpontot ad.

8. Következtetések

1. Létrehoztunk egy populációgenetikai algoritmust, mely képes magas specificitással azonosítani az inkompleten penetráns variánsokat és az interallelikus interakciókat.
2. Az autoszomális recesszív betegségekért felelős gyakori variánsok 29%-át (25/85) inkompleten penetránsnak találtuk. Az *NPHS2* R229Q variáns volt ezek között az egyetlen interallelikus interakció-függő.
3. Igazoltuk kísérletesen az *NPHS2* R229Q variáns interallelikus interakciói háttérében a kóros oligomerizáció szerepét. A PHB domének között mért FRET hatékonyság alapján az R229Q podocin oligomerek szerkezete eltér a vad oligomerekétől. A patogén R229Q oligomerek szerkezete a benignusakétól.
4. Létrehoztunk egy irányelvet, mely segít megítélni az R229Q asszociációk patogenitását a klinikai gyakorlatban

9. Összefoglalás

Munkánk során egy új, populációgenetikai módszert dolgoztunk ki az inkompletten penetráns variánsok azonosítására. A korábban azonosított *NPHS2* R229Q példáján, a jelenség hátterében interallelikus interakciókat kerestünk.

Összesen 17 autoszomális recesszíven öröklődő gént vizsgáltunk: *ASL*, *ATP7B*, *CAPN3*, *CFTR*, *CTNS*, *DHCR7*, *GAA*, *GALNS*, *GALT*, *IDUA*, *MUT*, *NPHS1*, *NPHS2*, *PAH*, *PKHD1*, *PMM2*, *SLC26A4*.

A beteg populációt manuálisan hoztuk létre, ezáltal összesen 12048 beteg geno- és fenotípus adatát és 3296 variánst elemeztünk. A variánsok analíziséhez egy számítógépes algoritmust fejlesztettünk.

Ahhoz, hogy meghatározzuk a patogén variánsok (n=1936) penetranciáját, az adott mutáció allélfrekvenciáját a teljesen penetráns, LOF mutációk kumulatív allélfrekvenciájához viszonyítottuk a beteg- és a kontroll-populációban a következő

képlet alapján: $\frac{AC_{bet}^V / AC_{gnomAD}^V}{AC_{bet}^{LOF} / AC_{gnomAD}^{LOF}}$ (V/LOF arány).

Az interallelikus interakciók azonosítása és módszerünk validálása céljából a variánsokat osztályoztuk biológiai predikció, klinikum illetve fehérje-doménba való lokalizáció alapján.

A 85 kellően gyakori variáns közül 25 (29%) variáns penetranciája volt szignifikánsan csökkent ($p < 7,22 \times 10^{-5}$). A módszer specificitását támasztja alá a variánsok biológiai és a klinikai megközelítésével való jó korreláció.

Egy variáns esetében igazoltunk interallelikus interakciót (*NPHS2* R229Q). A variáns társulásait FRET módszerrel vizsgáltuk tovább, szignifikáns különbséget tapasztalva a vad és R229Q homo- és heterooligomerek ($p = 2,7 \times 10^{-07}$) és a nem patogén és patogén társulások között ($p = 0,0029$). Igazoltuk így, hogy a domináns negatív hatás a megváltozott konformáció következménye.

Eddigi eredményeinket összesítve létrehoztunk egy, az R229Q társulások klinikai megítélését segítő irányelvet.

10. Summary

We aimed to identify incompletely penetrant (IP) variants and interallelic interactions in autosomal recessive disorders by a population-genetic approach

Altogether we analysed 17 frequent autosomal recessive disorders: *ASL*, *ATP7B*, *CAPN3*, *CFTR*, *CTNS*, *DHCR7*, *GAA*, *GALNS*, *GALT*, *IDUA*, *MUT*, *NPHS1*, *NPHS2*, *PAH*, *PKHD1*, *PMM2*, *SLC26A4* and developed a computerized algorithm.

We have gathered genotype and clinical information from 12048 patients and analysed 3296 variants. We calculated the relative allele frequency of each pathogenic variant (n=1936) to the LOF variants of the corresponding gene in the patient and the general population and estimated the penetrance of each variant by calculating their ratio:

$$\frac{AC_{pt}^V / AC_{pt}^{LOF}}{AC_{gnomAD}^V / AC_{gnomAD}^{LOF}} \text{ (V/LOF ratio)}.$$

To assess the interallelic interactions and to validate the population-genetic algorithm, we classified all variants based on biological, clinical assessments and domain localization.

We found 25 variants, 29% of the frequent 85 variants, to be underrepresented in the patient population ($p < 7.22 \times 10^{-5}$). The good specificity of the method is indicated by the fact that none of the 25 IP variants was LOF and 16 were hypomorphic based on the phenotype.

Among them, only the *NPHS2* R229Q variant was subject to interallelic interactions.

We found a markedly different fluorescent resonance energy transfer (FRET) efficiency between the wt and R229Q homo- and heterooligomers ($p = 2,7 \times 10^{-07}$) and between the pathogenic associations of R229Q compared with non-pathogenic associations ($p = 0,0029$). We thus proved experimentally that the interallelic interactions of *NPHS2* R229Q result from an altered conformation of the oligomers.

We created a guideline to help the clinical assessment of *NPHS2* R229Q associations based on population-genetic, biochemical and clinical aspects.

11. Irodalomjegyzék

1. Zlotogora J. (2003) Penetrance and expressivity in the molecular age. *Genetics in Medicine*, 5: 347-352.
2. Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, Phillips AD, Shaw K, Stenson PD, Cooper DN, Tyler-Smith C, Genomes Project C. (2012) Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet*, 91: 1022-1032.
3. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, Albers CA, Zhang ZD, Conrad DF, Lunter G, Zheng H, Ayub Q, DePristo MA, Banks E, Hu M, Handsaker RE, Rosenfeld JA, Fromer M, Jin M, Mu XJ, Khurana E, Ye K, Kay M, Saunders GI, Suner MM, Hunt T, Barnes IH, Amid C, Carvalho-Silva DR, Bignell AH, Snow C, Yngvadottir B, Bumpstead S, Cooper DN, Xue Y, Romero IG, Genomes Project C, Wang J, Li Y, Gibbs RA, McCarroll SA, Dermitzakis ET, Pritchard JK, Barrett JC, Harrow J, Hurles ME, Gerstein MB, Tyler-Smith C. (2012) A systematic survey of loss-of-function variants in human protein-coding genes. *Science*, 335: 823-828.
4. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. (2010) A map of human genome variation from population-scale sequencing. *Nature*, 467: 1061-1073.
5. Burdon KP, Wirth MG, Mackey DA, Russell-Eggitt IM, Craig JE, Elder JE, Dickinson JL, Sale MM. (2004) A novel mutation in the Connexin 46 gene causes autosomal dominant congenital cataract with incomplete penetrance. *J Med Genet*, 41: e106.
6. Saini S, Robinson PN, Singh JR, Vanita V. (2012) A novel 7 bp deletion in PRPF31 associated with autosomal dominant retinitis pigmentosa with incomplete penetrance in an Indian family. *Exp Eye Res*, 104: 82-88.
7. Mathias A, Moss AJ, Lopes CM, Barsheshet A, McNitt S, Zareba W, Robinson JL, Locati EH, Ackerman MJ, Benhorin J, Kaufman ES, Platonov PG, Qi M, Shimizu W, Towbin JA, Michael Vincent G, Wilde AA, Zhang L, Goldenberg I.

- (2013) Prognostic implications of mutation-specific QTc standard deviation in congenital long QT syndrome. *Heart Rhythm*, 10: 720-725.
8. Cooper DN, Krawczak M, Polychronakos C, Tyler-Smith C, Kehrer-Sawatzki H. (2013) Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Hum Genet*, 132: 1077-1130.
 9. Thauvin-Robinet C, Munck A, Huet F, Genin E, Bellis G, Gautier E, Audrezet MP, Ferec C, Lalau G, Georges MD, Claustres M, Bienvenu T, Gerard B, Boisseau P, Cabet-Bey F, Feldmann D, Clavel C, Bieth E, Iron A, Simon-Bouy B, Costa C, Medina R, Leclerc J, Hubert D, Nove-Josserand R, Sermet-Gaudelus I, Rault G, Flori J, Leroy S, Wizla N, Bellon G, Haloun A, Perez-Martin S, d'Acromont G, Corvol H, Clement A, Houssin E, Binquet C, Bonithon-Kopp C, Alberti-Boulme C, Morris MA, Faivre L, Goossens M, Roussey M, Collaborating Working Group on RH, Girodon E. (2009) The very low penetrance of cystic fibrosis for the R117H mutation: a reappraisal for genetic counselling and newborn screening. *J Med Genet*, 46: 752-758.
 10. Soriano A, Manna R. (2012) Familial Mediterranean fever: new phenotypes. *Autoimmun Rev*, 12: 31-37.
 11. van Spronsen FJ. (2010) Phenylketonuria: a 21st century perspective. *Nat Rev Endocrinol*, 6: 509-514.
 12. Srivastava S, Molinari E, Raman S, Sayer JA. (2017) Many Genes-One Disease? Genetics of Nephronophthisis (NPHP) and NPHP-Associated Disorders. *Front Pediatr*, 5: 287.
 13. Hoefele J, Wolf MT, O'Toole JF, Otto EA, Schultheiss U, Deschenes G, Attanasio M, Utsch B, Antignac C, Hildebrandt F. (2007) Evidence of oligogenic inheritance in nephronophthisis. *J Am Soc Nephrol*, 18: 2789-2795.
 14. Katsanis N, Ansley SJ, Badano JL, Eichers ER, Lewis RA, Hoskins BE, Scambler PJ, Davidson WS, Beales PL, Lupski JR. (2001) Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science*, 293: 2256-2259.
 15. Badano JL, Katsanis N. (2002) Beyond Mendel: an evolving view of human genetic disease transmission. *Nat Rev Genet*, 3: 779-789.

16. Davies JC, Turner MW, Klein N, London MBLCFSG. (2004) Impaired pulmonary status in cystic fibrosis adults with two mutated MBL-2 alleles. *Eur Respir J*, 24: 798-804.
17. Frangolias DD, Ruan J, Wilcox PJ, Davidson AG, Wong LT, Berthiaume Y, Hennessey R, Freitag A, Pedder L, Corey M, Sweezey N, Zielenski J, Tullis E, Sandford AJ. (2003) Alpha 1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *Am J Respir Cell Mol Biol*, 29: 390-396.
18. Buscher R, Eilmes KJ, Grasemann H, Torres B, Knauer N, Sroka K, Insel PA, Ratjen F. (2002) beta2 adrenoceptor gene polymorphisms in cystic fibrosis lung disease. *Pharmacogenetics*, 12: 347-353.
19. Zhang D, Bai L. (2016) Interallelic interaction and gene regulation in budding yeast. *Proc Natl Acad Sci U S A*, 113: 4428-4433.
20. Shen N, Heintz C, Thiel C, Okun JG, Hoffmann GF, Blau N. (2016) Co-expression of phenylalanine hydroxylase variants and effects of interallelic complementation on in vitro enzyme activity and genotype-phenotype correlation. *Mol Genet Metab*, 117: 328-335.
21. Nadler HL, Chacko CM, Rachmeler M. (1970) Interallelic complementation in hybrid cells derived from human diploid strains deficient in galactose-1-phosphate uridyl transferase activity. *Proc Natl Acad Sci U S A*, 67: 976-982.
22. Balmer C, Pandey AV, Rufenacht V, Nuoffer JM, Fang P, Wong LJ, Haberle J. (2014) Mutations and polymorphisms in the human argininosuccinate lyase (ASL) gene. *Hum Mutat*, 35: 27-35.
23. McInnes RR, Shih V, Chilton S. (1984) Interallelic complementation in an inborn error of metabolism: genetic heterogeneity in argininosuccinate lyase deficiency. *Proc Natl Acad Sci U S A*, 81: 4480-4484.
24. Mekahli D, Liutkus A, Ranchin B, Yu A, Bessenay L, Girardin E, Van Damme-Lombaerts R, Palcoux JB, Cachat F, Lavocat MP, Bourdat-Michel G, Nobili F, Cochat P. (2009) Long-term outcome of idiopathic steroid-resistant nephrotic syndrome: a multicenter study. *Pediatr Nephrol*, 24: 1525-1532.
25. Noone DG, Iijima K, Parekh R. (2018) Idiopathic nephrotic syndrome in children. *Lancet*, 392: 61-74.
26. Santin S, Bullich G, Tazon-Vega B, Garcia-Maset R, Gimenez I, Silva I, Ruiz P, Ballarin J, Torra R, Ars E. (2011) Clinical utility of genetic testing in children and

- adults with steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol*, 6: 1139-1148.
27. Buscher AK, Kranz B, Buscher R, Hildebrandt F, Dworniczak B, Pennekamp P, Kuwertz-Broking E, Wingen AM, John U, Kemper M, Monnens L, Hoyer PF, Weber S, Konrad M. (2010) Immunosuppression and renal outcome in congenital and pediatric steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol*, 5: 2075-2084.
 28. Akchurin O, Reidy KJ. (2015) Genetic causes of proteinuria and nephrotic syndrome: impact on podocyte pathobiology. *Pediatr Nephrol*, 30: 221-233.
 29. Ha TS. (2017) Genetics of hereditary nephrotic syndrome: a clinical review. *Korean J Pediatr*, 60: 55-63.
 30. Yu SM, Nissaisorakarn P, Husain I, Jim B. (2018) Proteinuric Kidney Diseases: A Podocyte's Slit Diaphragm and Cytoskeleton Approach. *Front Med (Lausanne)*, 5: 221.
 31. Schwarz K, Simons M, Reiser J, Saleem MA, Faul C, Kriz W, Shaw AS, Holzman LB, Mundel P. (2001) Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin. *J Clin Invest*, 108: 1621-1629.
 32. Grahammer F, Schell C, Huber TB. (2013) Molecular understanding of the slit diaphragm. *Pediatr Nephrol*, 28: 1957-1962.
 33. Boute N, Gribouval O, Roselli S, Benessy F, Lee H, Fuchshuber A, Dahan K, Gubler MC, Niaudet P, Antignac C. (2000) NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome. *Nat Genet*, 24: 349-354.
 34. Somlo S, Mundel P. (2000) Getting a foothold in nephrotic syndrome. *Nat Genet*, 24: 333-335.
 35. Huber TB, Simons M, Hartleben B, Sernetz L, Schmidts M, Gundlach E, Saleem MA, Walz G, Benzing T. (2003) Molecular basis of the functional podocin-nephrin complex: mutations in the NPHS2 gene disrupt nephrin targeting to lipid raft microdomains. *Hum Mol Genet*, 12: 3397-3405.
 36. Tory K, Menyhard DK, Woerner S, Nevo F, Gribouval O, Kerti A, Straner P, Arrondel C, Cong EH, Tulassay T, Mollet G, Perczel A, Antignac C. (2014) Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. *Nature Genetics*, 46: 299-+.

37. Straner P, Balogh E, Schay G, Arrondel C, Miko A, L'Aune G, Benmerah A, Perczel A, D KM, Antignac C, Mollet G, Tory K. (2018) C-terminal oligomerization of podocin mediates interallelic interactions. *Biochim Biophys Acta Mol Basis Dis*, 1864: 2448-2457.
38. Hinkes BG, Mucha B, Vlangos CN, Gbadegesin R, Liu J, Hasselbacher K, Hangan D, Ozaltin F, Zenker M, Hildebrandt F, Arbeitsgemeinschaft fur Paediatriche Nephrologie Study G. (2007) Nephrotic syndrome in the first year of life: two thirds of cases are caused by mutations in 4 genes (NPHS1, NPHS2, WT1, and LAMB2). *Pediatrics*, 119: e907-919.
39. Bouchireb K, Boyer O, Gribouval O, Nevo F, Huynh-Cong E, Moriniere V, Campait R, Ars E, Brackman D, Dantal J, Eckart P, Gigante M, Lipska BS, Liutkus A, Megarbane A, Mohsin N, Ozaltin F, Saleem MA, Schaefer F, Souлами K, Torra R, Garcelon N, Mollet G, Dahan K, Antignac C. (2014) NPHS2 mutations in steroid-resistant nephrotic syndrome: a mutation update and the associated phenotypic spectrum. *Hum Mutat*, 35: 178-186.
40. Tsukaguchi H, Sudhakar A, Le TC, Nguyen T, Yao J, Schwimmer JA, Schachter AD, Poch E, Abreu PF, Appel GB, Pereira AB, Kalluri R, Pollak MR. (2002) NPHS2 mutations in late-onset focal segmental glomerulosclerosis: R229Q is a common disease-associated allele. *J Clin Invest*, 110: 1659-1666.
41. Kerti A, Csohany R, Wagner L, Javorszky E, Maka E, Tory K. (2013) NPHS2 homozygous p.R229Q variant: potential modifier instead of causal effect in focal segmental glomerulosclerosis. *Pediatr Nephrol*, 28: 2061-2064.
42. Machuca E, Hummel A, Nevo F, Dantal J, Martinez F, Al-Sabban E, Baudouin V, Abel L, Grunfeld JP, Antignac C. (2009) Clinical and epidemiological assessment of steroid-resistant nephrotic syndrome associated with the NPHS2 R229Q variant. *Kidney Int*, 75: 727-735.
43. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, DeFlaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G,

- Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Exome Aggregation C. (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536: 285-291.
44. Miko A, Kaposi A, Schnabel K, Seidl D, Tory K. (2021) Identification of incompletely penetrant variants and interallelic interactions in autosomal recessive disorders by a population-genetic approach. *Hum Mutat*, 42: 1473-1487.
45. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, El Houdaigui B, Fatima R, Gall A, Garcia Giron C, Grego T, Guijarro-Clarke C, Haggerty L, Hemrom A, Hourlier T, Izuogu OG, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Gonzalez Martinez J, Marugan JC, Maurel T, McMahan AC, Mohanan S, Moore B, Muffato M, Oheh DN, Paraschas D, Parker A, Parton A, Prosovetskaia I, Sakthivel MP, Salam AIA, Schmitt BM, Schuilenburg H, Sheppard D, Steed E, Szpak M, Szuba M, Taylor K, Thormann A, Threadgold G, Walts B, Winterbottom A, Chakiachvili M, Chaubal A, De Silva N, Flint B, Frankish A, Hunt SE, GR II, Langridge N, Loveland JE, Martin FJ, Mudge JM, Morales J, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Cunningham F, Yates AD, Zerbino DR, Flicek P. (2021) Ensembl 2021. *Nucleic Acids Res*, 49: D884-D891.
46. den Dunnen JT. (2017) Describing Sequence Variants Using HGVS Nomenclature. *Methods Mol Biol*, 1492: 243-251.
47. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PE. (2008) Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. *Hum Mutat*, 29: 6-13.
48. Baker KE, Parker R. (2004) Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol*, 16: 293-299.
49. Noone PG, Knowles MR. (2001) 'CFTR-opathies': disease phenotypes associated with cystic fibrosis transmembrane regulator gene mutations. *Respir Res*, 2: 328-332.

50. Kalatzis V, Nevo N, Cherqui S, Gasnier B, Antignac C. (2004) Molecular pathogenesis of cystinosis: effect of CTNS mutations on the transport activity and subcellular localization of cystinosin. *Hum Mol Genet*, 13: 1361-1371.
51. Witsch-Baumgartner M, Fitzky BU, Ogorelkova M, Kraft HG, Moebius FF, Glossmann H, Seedorf U, Gillessen-Kaesbach G, Hoffmann GF, Clayton P, Kelley RI, Utermann G. (2000) Mutational spectrum in the Delta7-sterol reductase gene and genotype-phenotype correlation in 84 patients with Smith-Lemli-Opitz syndrome. *Am J Hum Genet*, 66: 402-412.
52. Montano AM, Tomatsu S, Gottesman GS, Smith M, Orii T. (2007) International Morquio A Registry: clinical manifestation and natural course of Morquio A disease. *J Inherit Metab Dis*, 30: 165-174.
53. Bosch AM. (2006) Classical galactosaemia revisited. *J Inherit Metab Dis*, 29: 516-525.
54. Muenzer J, Wraith JE, Clarke LA, International Consensus Panel on M, Treatment of Mucopolysaccharidosis I. (2009) Mucopolysaccharidosis I: management and treatment guidelines. *Pediatrics*, 123: 19-29.
55. Thomas JA, Beck M, Clarke JT, Cox GF. (2010) Childhood onset of Scheie syndrome, the attenuated form of mucopolysaccharidosis I. *J Inherit Metab Dis*, 33: 421-427.
56. Jansen R, Ledley FD. (1990) Heterozygous mutations at the mut locus in fibroblasts with mut0 methylmalonic acidemia identified by polymerase-chain-reaction cDNA cloning. *Am J Hum Genet*, 47: 808-814.
57. Guttler F, Azen C, Guldberg P, Romstad A, Hanley WB, Levy HL, Matalon R, Rouse BM, Trefz F, de la Cruz F, Koch R. (1999) Relationship among genotype, biochemical phenotype, and cognitive performance in females with phenylalanine hydroxylase deficiency: report from the Maternal Phenylketonuria Collaborative Study. *Pediatrics*, 104: 258-262.
58. Denamur E, Delezoide AL, Alberti C, Bourillon A, Gubler MC, Bouvier R, Pascaud O, Elion J, Grandchamp B, Michel-Calemard L, Missy P, Zaccaria I, Le Nagard H, Gerard B, Loirat C, Societe Francaise de F, Barbet J, Beaufrere AM, Berchel C, Bessieres B, Boudjemaa S, Buenerd A, Carles D, Clemenson A, Dechelotte P, Devisme L, Dijoud F, Esperandieu O, Fallet C, Gonzales M, Hillion Y, Jacob B, Joubert M, Kermanach P, Lallemand A, Laquerriere A, Laurent N,

- Liprandi A, Loeuillet L, Loget P, Martinovic J, Menez F, Narcy F, Roux JJ, Rouleau-Dubois C, Sinico M, Tantau J, Wann AR. (2010) Genotype-phenotype correlations in fetuses and neonates with autosomal recessive polycystic kidney disease. *Kidney Int*, 77: 350-358.
59. de Lonlay P, Seta N, Barrot S, Chabrol B, Drouin V, Gabriel BM, Journal H, Kretz M, Laurent J, Le Merrer M, Leroy A, Pedespan D, Sarda P, Villeneuve N, Schmitz J, van Schaftingen E, Matthijs G, Jaeken J, Korner C, Munnich A, Saudubray JM, Cormier-Daire V. (2001) A broad spectrum of clinical presentations in congenital disorders of glycosylation I: a series of 26 cases. *J Med Genet*, 38: 14-19.
60. Choi BY, Stewart AK, Madeo AC, Pryor SP, Lenhard S, Kittles R, Eisenman D, Kim HJ, Niparko J, Thomsen J, Arnos KS, Nance WE, King KA, Zalewski CK, Brewer CC, Shawker T, Reynolds JC, Butman JA, Karniski LP, Alper SL, Griffith AJ. (2009) Hypo-functional SLC26A4 variants associated with nonsyndromic hearing loss and enlargement of the vestibular aqueduct: genotype-phenotype correlation or coincidental polymorphisms? *Hum Mutat*, 30: 599-608.
61. Dai P, Stewart AK, Chebib F, Hsu A, Rozenfeld J, Huang D, Kang D, Lip V, Fang H, Shao H, Liu X, Yu F, Yuan H, Kenna M, Miller DT, Shen Y, Yang W, Zelikovic I, Platt OS, Han D, Alper SL, Wu BL. (2009) Distinct and novel SLC26A4/Pendrin mutations in Chinese and U.S. patients with nonsyndromic hearing loss. *Physiol Genomics*, 38: 281-290.
62. Soh LM, Druce M, Grossman AB, Differ AM, Rajput L, Bitner-Glindzicz M, Korbonits M. (2015) Evaluation of genotype-phenotype relationships in patients referred for endocrine assessment in suspected Pendred syndrome. *Eur J Endocrinol*, 172: 217-226.
63. Sampaleanu LM, Vallee F, Thompson GD, Howell PL. (2001) Three-dimensional structure of the argininosuccinate lyase frequently complementing allele Q286R. *Biochemistry*, 40: 15570-15580.
64. Deguti MM, Genschel J, Cancado EL, Barbosa ER, Bochow B, Mucenic M, Porta G, Lochs H, Carrilho FJ, Schmidt HH. (2004) Wilson disease: novel mutations in the ATP7B gene and clinical correlation in Brazilian patients. *Hum Mutat*, 23: 398.
65. Sorimachi H, Imajoh-Ohmi S, Emori Y, Kawasaki H, Ohno S, Minami Y, Suzuki K. (1989) Molecular cloning of a novel mammalian calcium-dependent protease

- distinct from both m- and mu-types. Specific expression of the mRNA in skeletal muscle. *J Biol Chem*, 264: 20106-20111.
66. Sheppard DN, Welsh MJ. (1999) Structure and function of the CFTR chloride channel. *Physiol Rev*, 79: S23-45.
 67. Kalatzis V, Cherqui S, Antignac C, Gasnier B. (2001) Cystinosin, the protein defective in cystinosis, is a H(+)-driven lysosomal cystine transporter. *EMBO J*, 20: 5940-5949.
 68. Witsch-Baumgartner M, Loffler J, Utermann G. (2001) Mutations in the human DHCR7 gene. *Hum Mutat*, 17: 172-182.
 69. Sugawara K, Saito S, Sekijima M, Ohno K, Tajima Y, Kroos MA, Reuser AJ, Sakuraba H. (2009) Structural modeling of mutant alpha-glucosidases resulting in a processing/transport defect in Pompe disease. *J Hum Genet*, 54: 324-330.
 70. Rivera-Colon Y, Schutsky EK, Kita AZ, Garman SC. (2012) The structure of human GALNS reveals the molecular basis for mucopolysaccharidosis IV A. *J Mol Biol*, 423: 736-751.
 71. Maita N, Tsukimura T, Taniguchi T, Saito S, Ohno K, Taniguchi H, Sakuraba H. (2013) Human alpha-L-iduronidase uses its own N-glycan as a substrate-binding and catalytic module. *Proc Natl Acad Sci U S A*, 110: 14628-14633.
 72. Lempp TJ, Suormala T, Siegenthaler R, Baumgartner ER, Fowler B, Steinmann B, Baumgartner MR. (2007) Mutation and biochemical analysis of 19 probands with mut0 and 13 with mut- methylmalonic aciduria: identification of seven novel mutations. *Mol Genet Metab*, 90: 284-290.
 73. Huber TB, Hartleben B, Kim J, Schmidts M, Schermer B, Keil A, Egger L, Lecha RL, Borner C, Pavenstadt H, Shaw AS, Walz G, Benzing T. (2003) Nephrin and CD2AP associate with phosphoinositide 3-OH kinase and stimulate AKT-dependent signaling. *Mol Cell Biol*, 23: 4917-4928.
 74. Bercovich D, Elimelech A, Zlotogora J, Korem S, Yardeni T, Gal N, Goldstein N, Vilensky B, Segev R, Avraham S, Loewenthal R, Schwartz G, Anikster Y. (2008) Genotype-phenotype correlations analysis of mutations in the phenylalanine hydroxylase (PAH) gene. *J Hum Genet*, 53: 407-418.
 75. Gunay-Aygun M, Tuchman M, Font-Montgomery E, Lukose L, Edwards H, Garcia A, Ausavarat S, Ziegler SG, Piwnicka-Worms K, Bryant J, Bernardini I, Fischer R, Huizing M, Guay-Woodford L, Gahl WA. (2010) PKHD1 sequence

- variations in 78 children and adults with autosomal recessive polycystic kidney disease and congenital hepatic fibrosis. *Mol Genet Metab*, 99: 160-173.
76. Silvaggi NR, Zhang C, Lu Z, Dai J, Dunaway-Mariano D, Allen KN. (2006) The X-ray crystal structures of human alpha-phosphomannomutase 1 reveal the structural basis of congenital disorder of glycosylation type 1a. *J Biol Chem*, 281: 14918-14926.
 77. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Malik Sharif S, Karbani G, Ahmed M, Bond J, Clayton D, Inglehearn CF. (2006) Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. *Am J Hum Genet*, 78: 889-896.
 78. Jonhston HRK, Bronya JB; Sherman, Stephanie L. (2018) Population Genetics. In: Emery and rimoin's principles and practice of medical genetics and genomics : clinical principles and applications. (R. E. Reed E. Pyeritz Ed. 7th edition. ed.). San Diego, CA: Elsevier.
 79. Kelley KA, Stamm S, Kozak CA. (1992) Expression and chromosome localization of the murine cystic fibrosis transmembrane conductance regulator. *Genomics*, 13: 381-388.
 80. Lazarin GA, Haque IS, Evans EA, Goldberg JD. (2017) Smith-Lemli-Opitz syndrome carrier frequency and estimates of in utero mortality rates. *Prenat Diagn*, 37: 350-355.
 81. Nowaczyk MJ, Waye JS, Douketis JD. (2006) DHCR7 mutation carrier rates and prevalence of the RSH/Smith-Lemli-Opitz syndrome: where are the patients? *Am J Med Genet A*, 140: 2057-2062.
 82. Derichs N, Schuster A, Grund I, Ernsting A, Stolpe C, Kortge-Jung S, Gallati S, Stuhmann M, Kozlowski P, Ballmann M. (2005) Homozygosity for L997F in a child with normal clinical and chloride secretory phenotype provides evidence that this cystic fibrosis transmembrane conductance regulator mutation does not cause cystic fibrosis. *Clin Genet*, 67: 529-531.
 83. Miko A, D KM, Kaposi A, Antignac C, Tory K. (2018) The mutation-dependent pathogenicity of NPHS2 p.R229Q: A guide for clinical assessment. *Hum Mutat*, 39: 1854-1860.

84. Nowaczyk MJ, Heshka T, Eng B, Feigenbaum AJ, Wayne JS. (2001) DHCR7 genotypes of cousins with Smith-Lemli-Opitz syndrome. *Am J Med Genet*, 100: 162-163.
85. Nowaczyk MJ, Nakamura LM, Wayne JS. (2001) DHCR7 and Smith-Lemli-Opitz syndrome. *Clin Invest Med*, 24: 311-317.
86. Hauerlev S, Sveen ML, Duno M, Angelini C, Vissing J, Krag TO. (2012) Calpain 3 is important for muscle regeneration: evidence from patients with limb girdle muscular dystrophies. *BMC Musculoskelet Disord*, 13: 43.
87. Monestrol Id. ((2014). WS23.5 *Journal of Cystic Fibrosis*, abstract.) Clinical importance of homozygous R117C (c.349C>T) CFTR mutation. . *Journal of Cystic Fibrosis*, abstract. .
88. Kroos MA, Pomponio RJ, Hagemans ML, Keulemans JL, Phipps M, DeRiso M, Palmer RE, Aulsems MG, Van der Beek NA, Van Diggelen OP, Halley DJ, Van der Ploeg AT, Reuser AJ. (2007) Broad spectrum of Pompe disease in patients with the same c.-32-13T->G haplotype. *Neurology*, 68: 110-115.
89. Okano Y, Eisensmith RC, Dasovich M, Wang T, Guttler F, Woo SL. (1991) A prevalent missense mutation in Northern Europe associated with hyperphenylalaninaemia. *Eur J Pediatr*, 150: 347-352.
90. Zschocke J, Graham CA, Stewart FJ, Carson DJ, Nevin NC. (1994) Non-phenylketonuria hyperphenylalaninaemia in Northern Ireland: frequent mutation allows screening and early diagnosis. *Hum Mutat*, 4: 114-118.
91. Bergmann C, Senderek J, Windelen E, Kupper F, Middeldorf I, Schneider F, Dornia C, Rudnik-Schoneborn S, Konrad M, Schmitt CP, Seeman T, Neuhaus TJ, Vester U, Kirfel J, Buttner R, Zerres K, Apn. (2005) Clinical consequences of PKHD1 mutations in 164 patients with autosomal-recessive polycystic kidney disease (ARPKD). *Kidney Int*, 67: 829-848.
92. Rossetti S, Torra R, Coto E, Consugar M, Kubly V, Malaga S, Navarro M, El-Youssef M, Torres VE, Harris PC. (2003) A complete mutation screen of PKHD1 in autosomal-recessive polycystic kidney disease (ARPKD) pedigrees. *Kidney Int*, 64: 391-403.
93. Szabo T, Orosz P, Balogh E, Javorszky E, Mattyus I, Bereczki C, Maroti Z, Kalmar T, Szabo AJ, Reusz G, Varkonyi I, Marian E, Gombos E, Orosz O, Madar L, Balla G, Kappelmayer J, Tory K, Balogh I. (2018) Comprehensive genetic

- testing in children with a clinical diagnosis of ARPKD identifies phenocopies. *Pediatr Nephrol*, 33: 1713-1721.
94. Dardis A, Zanin I, Zampieri S, Stuani C, Pianta A, Romanello M, Baralle FE, Bembi B, Buratti E. (2014) Functional characterization of the common c.-32-13T>G mutation of GAA gene: identification of potential therapeutic agents. *Nucleic Acids Res*, 42: 1291-1302.
95. Sadowski CE, Lovric S, Ashraf S, Pabst WL, Gee HY, Kohl S, Engelmann S, Vega-Warner V, Fang H, Halbritter J, Somers MJ, Tan W, Shril S, Fessi I, Lifton RP, Bockenhauer D, El-Desoky S, Kari JA, Zenker M, Kemper MJ, Mueller D, Fathy HM, Soliman NA, Group SS, Hildebrandt F. (2015) A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. *J Am Soc Nephrol*, 26: 1279-1289.

12. Saját publikációk jegyzéke

12.1. A disszertációhoz kapcsolódó közlemények jegyzéke

Ágnes Mikó, Ambrus Kaposi, Karolina Schnabel, Dániel Seidl, Kálmán Tory

Identification of incompletely penetrant variants and interallelic interactions in autosomal recessive disorders by a population-genetic approach

HUMAN MUTATION 42: 11 pp. 1473-1487. (2021)

IF: 4,878*

Ágnes Mikó, Dóra K. Menyhárd, Ambrus Kaposi, Corinne Antignac, Kálmán Tory

The mutation-dependent pathogenicity of NPHS2 p.R229Q: A guide for clinical assessment

HUMAN MUTATION 39: 12 pp. 1854-1860. (2018)

IF: 4,453

Pál Stráner, Eszter Balogh, Gusztáv Schay, Christelle Arrondele, **Ágnes Mikó**, Gerda L'Aunéb, Alexandre Benmerah, András Perczel, Dóra K. Menyhárd, Corinne Antignac, Géraldine Mollete, Kálmán Tory

C-terminal oligomerization of podocin mediates interallelic interactions

BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR BASIS OF DISEASE 1864:

7 pp. 2448-2457. (2018)

IF: 4,328

12.2. A disszertációtól független közlemények jegyzéke

Mikó Ágnes, Lóth Szendile, Müller Judit, Lotz Bence, Rossitto Patrizio, Szabolcs Andrea, Benyó Gábor, Jávorszky Eszter, Tory Kálmán, Dezsőfi Antal

Arthrogryposis–renalis diszfunkció–cholestasis szindróma [Arthrogryposis–renal dysfunction–cholestasis syndrome]

ORVOSI HETILAP 163: 2 pp. 74-78. (2022)

IF: 0,540**

Szeri Flora, **Miko Agnes**, Navasiolava Nastassia, Kaposi Ambrus, Verschuere Shana, Li Qiaoli, F Terry Sharon, Boraldi Federica, Uitto Jouni, van de Wetering Koen, Ludovic Martin, Daniela Quaglino, Olivier M Vanakker, Tory Kalman, Aranyi Tamas

The pathogenic p.R391G ABCC6 displays incomplete penetrance implying the necessity of an interacting partner for the development of pseudoxanthoma elasticum

Paper: **DOI: 10.1101/2020.11.26.20236489, 20 p. (2021)**

Csak repozitóriumban hozzáférhető közlemény

13. Köszönetnyilvánítás

Legnagyobb köszönettel és hálával témavezetőmnek, Dr. Tory Kálmánnak tartozom, aki bevezetett a kutatás világába és méltónak talált a közös munkára. Alapossága és precizitása, amellyel a munkáját végzi igazi inspiráció és követendő példa számomra az élet minden területén. Mentorként tekintek rá, és remélem a jövőben is a szárnyai alatt fejlődhetek, nem csak szakmailag, hanem emberileg is.

Köszönet Dr. Kaposi Ambrusnak, aki mindig fáradhatatlan lelkesedéssel állt hozzá az algoritmus kidolgozásához és a statisztikai megoldásokhoz. Kifogyhatatlan ötleteivel és energiájával egy színfolt volt a csapatban.

Köszönöm Prof. Dr. Tulassay Tivadarnak, hogy felvett a Doktori Iskolába, Prof. Dr. Szabó Attilának, hogy lehetőséget biztosított ahhoz, hogy ezt a kutatást elkezdhessük és végigvihessük.

Köszönettel tartozom Bernáth Máriának, akinek gondos labormunkájára mindig számíthattam. Barátsága és jó tanácsai a munkám alatt meghatározóak voltak.

Külön köszönet Dr. Balogh Eszternek, aki mérhetetlen türelmével és jó tanácsaival segített mind kutatómunkám, mind a klinikumban töltött idő alatt. Őszinte barátsága, önzetlen szeretete páratlan és egy életen át tartó.

Megszeretném köszönni Dr. Schay Gusztávnak a közös munkát, illetve a rendkívül sok türelmet és magyarázatot.

Köszönet Dr. Schnabel Karolinának és Dr. Seidl Dánielnek, hogy a TDK munkájuk során kitartó lelkesedéssel és szorgalommal segítették a kutatócsoport munkáját. A közös munka mindig üdítő volt számomra velük. Dánielnek külön köszönet a statisztikában való segítségéért!

Köszönet a munkacsoport tagjainak, Keszthelyt T. Magdolnának, Jávorszky Eszternek, Kétszeri Maténak és Légrádi Reginának akikkel öröm volt együtt dolgozni és a mindennapok általuk családias hangulatban teltek.

Köszönöm a laborból Dr. Pap Domonkosnak és Dr. Veres-Székely Apornak a sok technikai segítséget.

Köszönöm továbbá a laborban kutató valamennyi munkatársamnak a munkámhoz nyújtott segítséget.

Köszönöm a családomnak, hogy mindvégig bíztattak és türelemmel többször végighallgatták a előadásaimra való felkészülést, még akkor is ha nem sokat értettek a populációgenetikához ☺. Köszönöm, hogy osztoznak velem az örömeimben.