

dc_1984_22

*MTA Orvosi Tudományok Osztálya
Doktori Értekezés*



**GENETIKAI POLIMORFIZMUSOK VIZSGÁLATA: A
GÉNVARIÁCIÓK MOLEKULÁRIS HATÁSÁTÓLA
KOMPLEX JELLEGEK ÖRÖKLETES HÁTTERÉIG**

Dr. Rónai Zsolt

*Semmelweis Egyetem,
Biokémiai és Molekuláris Biológiai Intézet Molekuláris Biológiai Tanszék*

2022.

dc_1984_22

TARTALOM

Rövidítésjegyzék.....	3
1 Bevezetés	5
1.1 A genetikai információ megismerése	5
1.1.1 A kezdetek.....	5
1.1.2 A Humán Genom Projekt.....	6
1.1.3 A hiányzó örökölhetőség	10
1.1.4 A genetikai variációk feltérképezése	12
2 Irodalmi áttekintés.....	15
2.1 A genetikai variációk osztályozása	15
2.2 A genetikai variációk biológiai hatása.....	17
2.2.1 A genetikai variációk hatása molekuláris szinten.....	17
2.2.2 A genetikai variációk hatása a fenotípusra.....	19
2.2.2.1 A D4-es dopamin receptor	19
2.2.2.2 A wolframin.....	22
2.2.2.3 A SNAP-25.....	24
2.2.2.4 A 4-es komplement komponens	25
2.3 A genetikai variációk elemzése.....	28
2.3.1 A genotípus meghatározása.....	28
2.3.1.1 A genotipizáló módszerek alapelve, VNTR-ek elemzése	28
2.3.1.2 SNP-k genotipizálása	30
2.3.1.3 CNV-k genotipizálása	35
2.3.1.4 DNS-szekvenálás	36
2.3.2 A genotípus–fenotípus összefüggés kutatása	40
2.3.2.1 Genom szintű asszociáció vizsgálat (GWAS).....	40
2.3.2.2 Kandidáns gén vizsgálat.....	43
2.3.3 A genetikai variációk molekuláris hatásának elemzése.....	45
2.3.3.1 A DNS–fehérje kölcsönhatás vizsgálata	46
2.3.3.2 A miRNS-ek hatásának vizsgálata.....	51
3 Célkitűzések.....	54
4 Módszerek.....	55
4.1 Polimorfizmusok genotipizálása	55
4.1.1 Vizsgálati személyek, mintavétel, DNS-izolálás.....	55
4.1.2 Genotipizálás „hagyományos” PCR-rel.....	55
4.1.3 Genotipizálás real-time PCR-rel.....	56
4.1.4 Automatizált elektroforetikus technikák.....	57
4.1.5 Haplotípus / LD vizsgálata.....	59

4.2	Asszociáció vizsgálatok.....	60
4.2.1	Fenotípus-elemzés	60
4.2.2	Statisztikai elemzés.....	60
4.3	Molekuláris biológiai vizsgálatok	61
4.3.1	Luciferáz riporter konstrukciók készítése	61
4.3.2	Tranziens transzfekció, luciferáz aktivitás mérése	61
5	Eredmények és megbeszélés.....	63
5.1	Polimorfizmusok genotipizálása	63
5.1.1	Ismétlődési variációk és SNP-k elemzése.....	63
5.1.1.1	A dopaminerg rendszer kiválasztott polimorfizmusainak elemzése.....	64
5.1.1.2	A <i>WFS1</i> gén polimorfizmusainak vizsgálata	67
5.1.2	CNV-k genotipizálása.....	71
5.1.2.1	A <i>C4A</i> és <i>C4B</i> génszám elemzése	71
5.1.2.2	A glikogénszintáz-kináz 3 β gént érintő CNV vizsgálata	74
5.1.3	Hatékony elektroforetikus eljárások alkalmazása	75
5.1.3.1	Ultravékony gélelektroforézis	75
5.1.3.2	Kapilláris gélelektroforézis	76
5.1.3.3	DNS-fragmentum analízis mikrofluidikai eszköz alkalmazásával.....	79
5.1.4	Genetikai polimorfizmusok haplotípus analízise	82
5.1.4.1	A <i>DRD4</i> gén SNP-inek haplotípus elemzése	82
5.1.4.2	MiR-SNP-k haplotípusának vizsgálata	83
5.2	Asszociáció vizsgálatok.....	85
5.2.1	Agresszió	86
5.2.2	Cukorbetegség	92
5.2.3	Impulzivitás	96
5.3	Polimorfizmusok molekuláris funkciójának elemzése	98
5.3.1	A SNAP-25 gén polimorfizmusainak funkcionális elemzése	99
5.3.2	A <i>WFS1</i> gén polimorfizmusainak funkcionális vizsgálata.....	104
5.3.2.1	Az 5' régió polimorfizmusainak elemzése	104
5.3.2.2	A <i>WFS1</i> 3' UTR-ében lévő SNP-k funkcionális elemzése	105
6	Kitekintés	111
7	Köszönetnyilvánítás	116
8	Saját közlemények.....	117
8.1	A doktori értekezés alapjául szolgáló közlemények.....	117
8.2	A doktori értekezéshez kapcsolódó közlemények.....	118
9	Irodalomjegyzék.....	120
10	Függelék.....	143

RÖVIDÍTÉSJEGYZÉK

ADHD	figyelemhiányos hiperaktivitás (attention deficit hyperactivity disorder)
ANCOVA	variancia–kovariancia-elemzés (analysis of covariance)
ANOVA	variancia-analízis (analysis of variance)
BAC	bakteriális mesterséges kromoszóma (bacterial artificial chromosome)
bp	bázispár
CNV	ismétlődési szám variáció (copy number variation)
COPD	krónikus obstruktív tüdőbetegség (chronic obstructive pulmonary disease)
C_T	küszöbciklus (threshold cycle)
Dam	DNS-adenin-metiltranszferáz
DAT	dopamin transzporter
DBH	dopamin- β -hidroxiláz
DRD4	D4-es dopamin receptor
EMSA	electrophoretic mobility shift assay
ER	endoplazmás retikulum
EtBr	etídium bromid
FRET	fluoreszcencia rezonancia energia átadás (fluorescence resonance energy transfer)
GSK3B	glikogénszintáz-kináz 3 β
GWAS	genomszintű asszociációvizsgálat (genome wide association study)
HEK	humán embrionális vese (Human embryonic kidney)
HTR2A	szerootonin receptor 2A (5-hydroxytryptamine receptor)
HWE	Hardy–Weinberg-egyensúly
in / del	inzerció / deléció polimorfizmus
kb	kilobázispár
LD	kapcsoltsági egyensúly hiánya (linkage disequilibrium)
MAF	ritka allél gyakorisága (minor allele frequency)
MHC	fő hisztokompatibilitási komplex (major histocompatibility complex)
miR	mikro-RNS

miRNS	mikro-RNS
miR-SNP	mikro-RNS kötőhelyében elhelyezkedő SNP
NCBI	National Center for Biotechnology Information
NGS	újgenerációs szekvenálás (next generation sequencing)
nt	nukleotid
OMIM	Humán gének és genetikai rendellenességek on-line katalógusa (Online Mendelian Inheritance in Man)
OR	esélyhányados (odds ratio)
PCR	polimeráz lánreakció (polymerase chain reaction)
PVP	polivinil-pirrolidon
RFLP	restrikciós fragmentum hosszúság polimorfizmus (restriction fragment length polymorphism)
RISC	RNS-indukált csendesítő komplex (RNA-induced silencing complex)
rSAP	rekombináns alkalikus foszfatáz (recombinant shrimp alkaline phosphatase)
SBE	primerextenzió (single base extension)
SNAP-25	szinaptoszóma-asszociált fehérje-25 (synaptosome associated protein-25)
SNARE	soluble N-ethylmaleimide sensitive factor attachment protein receptor
SNP	egyponos nukleotid polimorfizmus (single nucleotide polymorphism)
STR	rövid ismétlődési polimorfizmus (short tandem repeat)
T1DM	1-es típusú diabetes mellitus
T2DM	2-es típusú diabetes mellitus
TBE	Tris–borát–EDTA pufferoldat (89 mM tris, 89 mM borát, 2 mM Na ₂ EDTA)
TH	tirozin-hidroxiláz
UPR	selejtfehérje válasz (unfolded protein response)
UTR	nem fehérjekódoló génszakasz (untranslated region)
VNTR	tandem ismétlődési polimorfizmus (variable number of tandem repeats)
WES	exom-szekvenálás (whole exome sequencing)
WFS1	wolframin ER transzmembrán glikoprotein
WGS	genom-szekvenálás (whole genome sequencing)

1 BEVEZETÉS

1.1 A genetikai információ megismerése

1.1.1 A kezdetek

Az öröklődés mechanizmusának, a genetikai információ átadásának megismerése több mint másfél évszázados múltra tekint vissza. Johann Gregor Mendel Ágoston-rendi szerzetes a brno-i Szent Tamás-kolostorban végezte el azokat a kísérleteket, melyekkel leírta az öröklődés alapvető törvényszerűségeit. Eredményeit 1865 februárjában Brnoban két tudományos konferencián is bemutatta, 1866-ban pedig egy tudományos publikációban (*Versuche über Pflanzenhybriden* – „Növényhibridizációs kísérletek”), valamint egy a kor neves botanikusának, Carl Nägelinek írt levelében foglalta össze. Kortársai azonban nem ismerték fel eredményeinek kiemelkedő jelentőségét, amihez hozzájárulhatott az is, hogy Mendel maga is kétségekkel értékelt megfigyeléseit, mivel a zöldborsón végzett kísérleteket hölgyemálon nem tudta reprodukálni. Ennek azonban egyszerű oka volt: a hölgyemál – a pitypangokhoz hasonlóan – nem csak ivaros, hanem „apomixissel” (testi sejtekből létrejött magvakkal) is szaporodik, és az így keletkező növények genetikailag az anyával azonosak [1].

Mendel kísérleteinek – a növényi hibridek létrehozásán túlmutató – általános jelentőségére a XX. század elején derült fény. Három botanikus, Carl Correns, Erich von Tschermak és Hugo de Vries egymástól függetlenül újra megfigyelte a Mendel által leírt öröklődési alapelveket, és talán még ennél is jelentősebb, hogy William Bateson rámutatott arra, hogy ezek a szabályok nem csak növényekben, hanem állatokban is megfigyelhetők. Munkásságuk alapján a szakirodalomban megjelent a genetikai számos alapfogalma: a „gén”, az „allél”, a „homo- és heterozigóta”, valamint a „mutáció” [2]. Thomas Hunt Morgan ecetmuslicán végzett kísérletei alapján rávilágított arra, hogy Mendel elmélete és a kromoszómák segítségével történő öröklődés nem két egymásnak ellentmondó teória, a kromoszómák funkciójának feltárásában végzett munkásságát 1933-ban orvosi Nobel-díjjal ismerték el [3, 4].

A XX. század közepén kiderült, hogy a genetikai kapcsolatban áll az anyagcsere folyamatokkal is: George W. Beadle és Edward L. Tatum leírta a B₆-vitamin génfüggő termelődését [5]. Ez a megfigyelés megalapozta az „egy gén – egy enzim” hipotézist, az azonban még mindig ismertlen volt, hogy a gének milyen molekulák, a legtöbb biológus azt feltételezte: fehérjék.

1944-ben Oswald Avery, Colin MacLeod és Maclyn McCarty kísérlete rávilágított a DNS szerepére, amikor kísérleteikben *Streptococcus pneumoniae*-ből származó tisztított DNS-sel másik baktériumtörzset fertőzővé alakítottak, mégis Francis Crick és James Watson a dezoxiribonukleinsav kettős spirál szerkezetére vonatkozó felfedezése (1953) vezetett csak el a DNS szerepének széleskörű felismeréséhez [1].

1956-ban Arthur Kornberg leírta a DNS-szintézis mechanizmusát, és azonosította az első DNS-függő DNS-polimerázt (a bakteriális DNS-polimeráz I-et), munkáját három évvel később Nobel-díjjal jutalmazták [6]. A DNS elsődleges szerkezetének, azaz információtartalmának megismerése az 1970-es évektől kezdett elérhetővé válni. 1973-ban jelent meg az első DNS-szekvenáláson alapuló közlemény, melyben Walter Gilbert és Allan Maxam leírták a lac-operon represszor kötő helyében lévő 24 bp hosszúságú szakasz bázissorrendjét [7], négy évvel később pedig a specifikus kémiai hasítás segítségével működő DNS-szekvenálási eljárást [8]. Ugyanebben az évben vált ismertté a már tradicionálisnak nevezett, de még mind a mai napig alapműszernek számító láncterminátor didezoxinukleotid segítségével megszakított szintézisen alapuló Sanger-szekvenálás [9].

1983-ban Kary Mullis megalkotta a molekuláris biológia egyik alapműszerét: a polimeráz láncreakció (PCR) [10] a *Thermus aquaticus*-ból származó hőstabil DNS-polimerázzal (Taq-polimeráz) [11] alapjaiban változtatta meg a genetikai információ megismerésének lehetőségét. Ez az eljárás a Sanger-szekvenálással együtt a Humán Genom Projekt metodikai alappilléreinek mondható.

1.1.2 A Humán Genom Projekt

A teljes emberi genom szekvenálásának és a bennünk található (fehérjekódoló) gének azonosításának lehetősége és gondolata az 1980-as évek során merült föl. Bár a feladat ekkor még – elméleti, technikai és anyagi szempontból egyaránt – megfoghatatlannak tűnt, egyre többen vetették fel, hogy az emberek közötti genetikai különbségek megismerése mind elméleti, mind gyakorlati orvosi szempontból új távlatokat jelentene a betegségek – pl. a rosszindulatú daganatok – megértésében [12, 13]. Ekkor már néhány organizmus – *Escherichia coli*, *Saccharomyces cerevisiae* – genomja ismert volt, ezek nagysága azonban együtt is az emberi genetikai

információ csupán kb. 1/8-át érte el. A Humán Genom Program nagy nemzetközi összefogásként 1990-ben indult útjának. Az USA-ban, Nagy-Britanniában, Franciaországban és Japánban működő központokhoz később német és kínai centrumok csatlakoztak. Megalapították a Humán Genom szervezetet (Human Genome Organization – HUGO), ami a genomkutatásban résztvevő csoportok nemzetközi szervezésének és összehangolásának fóruma lett [14]. A kutatás vezetői 1996-ban meghozták a projekt egyik talán legalapvetőbb döntését: minden 1000 bázispárnál hosszabb megismert szekvencia 24 órán belül nyilvános adatbázisba kerül, hogy annak további elemzése, jelentésének feltárása minél előbb elkezdődhessen [15].

Amellett, hogy a projekt kezdetén a technológia – mind az informatikai, mind a molekuláris biológiai, DNS-szekvenáló eszköztár – gyerekcipőben járt, hamar világossá vált egy másik, elméleti jellegű kihívás is. A humán genom igen sok, különböző kiterjedésű ismétlődő szekvenciát tartalmaz, ami az egyes szekvenciariészletek helyes összeillesztését jelentősen megnehezíti. A projekt vezető kutatói emiatt a klónról klónra történő hierarchikus (Hierarchical shotgun) szekvenálás stratégiáját alkalmazták. A módszer lényege, hogy a genomot hosszú, átfedő szakaszokra hasítják fel, melyekből könyvtárat hoznak létre: az átlagosan 170 kb hosszú DNS-szakaszokat BAC (bacterial artificial chromosome) vektorba klónozzák. A könyvtárakból ezt követően kiválogatásra kerülnek azok a klónok, melyek tartalmaznak megfelelő hosszúságú átfedő szekvenciákat, mivel ezek nélkül a végső szekvencia összeállítása nem lenne lehetséges. Az ismert fizikai elhelyezkedésű BAC-klónokat ezt követően további rövidebb fragmentumra hasítják, melyeket két oldalról megszekvenálnak. A kb. 10^5 bp nagyságú szekvenciák kisebb darabokból történő összeillesztése nem jelent nehézséget. A BAC-klónok fizikai elhelyezkedése pedig – azok átfedő szakaszai alapján – ismert, így ezzel az eljárással a nagy kiterjedésű emberi genom pontos szekvenciája is biztonságosan meghatározható bizonyult. A módszer azonban a megfelelő BAC-klónok, illetve könyvtárak létrehozása miatt jelentős munka- és időigényű előkészítő lépéseket tesz szükségessé [16].

Az NIH (National Institute of Health) és az Egyesült Államok Energia Hivatala (U.S. Department of Energy) által vezetett „állami” projekt mellett Craig Venter 1998-ban létrehozott egy magánvállalatot (Celera Genomics). A csoport a BAC klónok fizikai térképezését szükségtelennek vélte, elképzelésük szerint a teljes genom (whole genome shotgun) szekvenálásával jelentős költség és idő takarítható meg. Ezen eljárás során a teljes genomot már az első lépésben

különböző hosszúságú (2, 10 és 50 kb) átfedő szakaszokra hasítják, melyeket közvetlenül megszekvenálnak. A technika során alapvető, hogy a létrehozott fragmentumok legalább 5–10× lefedjék a genom minden szakaszát. Bár az eljárás rövid, és viszonylag kevés ismétlődést tartalmazó genomok meghatározása során valóban megbízható eredményt ad, a humán genom esetén azonban az egyes szakaszok helyes összeillesztése komoly nehézségekbe ütközik. A Celera Genomics kutatói az általuk kitűzött terveket nem tudták teljes mértékben megvalósítani, a szekvenálási eredmények feldolgozása során a Humán Genom Projekt eredményeit is felhasználták. Ennek ellenére az általuk javasolt technika – különösen az informatikai eszközök fejlődésével karöltve – a genomok meghatározásának fontos eszközévé vált [16].

Mind a hierarchikus, mind a teljes genom meghatározásán alapuló módszer nélkülözhetetlen kulcslépése a DNS-szekvencia meghatározása: ezen technika jelentős fejlődése nélkül az emberi genom megismerése nem lett volna lehetséges. Az eredeti Sanger-szekvenálás során az egyes nukleotidok meghatározása négy különböző reakcióban, radioaktív foszforral jelölt lánc-terminátor molekulák alkalmazásával történt. A reakciótermékeket poliakrilamid gélen választották el, és az autoradiográfiás módszerrel előhívott filmeket manuálisan elemezték, ami nem csak rendkívül sok időt vett igénybe, de a hibalehetőség is meglehetősen magas volt [17]. A kapilláris elektroforézis, valamint a fluoreszcens festékkel jelölt didezoxinukleotidok, illetve a lézer indukálta fluoreszcencia (LIF) elvén alapuló detektálás alkalmazásának bevezetése ugrászerű előrelépést jelentett. Bár a kapilláris elektroforézis módszerét már korábban kidolgozták, alkalmazását DNS-szekvenálásra 1990-ben írták le [18]. A technika önmagában is nagyságrendekkel emelte a szekvenálás gyorsaságát és megbízhatóságát, emellett több lényeges elméleti és módszertani újítás további jelentős előrelépést jelentett. Bevezetésre került a nem keresztkött poliakrilamid elválasztómátrix alkalmazása, mely folyékony halmazállapotának köszönhetően a kapillárisba pumpálható, illetve onnan kimosható, így az ömlesztett kvarc (fused silica) kapillárisok több egymást követő elválasztásra is használhatók [19]. Bár a hagyományos gélelektroforézishez képest a kapilláris elektroforézis megkérdőjelezhetetlen előrelépést jelentett, egyetlen kapilláris alkalmazása az áteresztőképességet a sok sávós gélhez képest nem növeli. Az igazi változást így a több, párhuzamos kapillárist tartalmazó berendezések megalkotása jelentette. A humán genom szekvenciájának megismerésében az Applied Biosystems cég 96 kapillárist tartalmazó „3700-as” eszköze kulcsfontosságú szerepet játszott: ebben a berendezésben

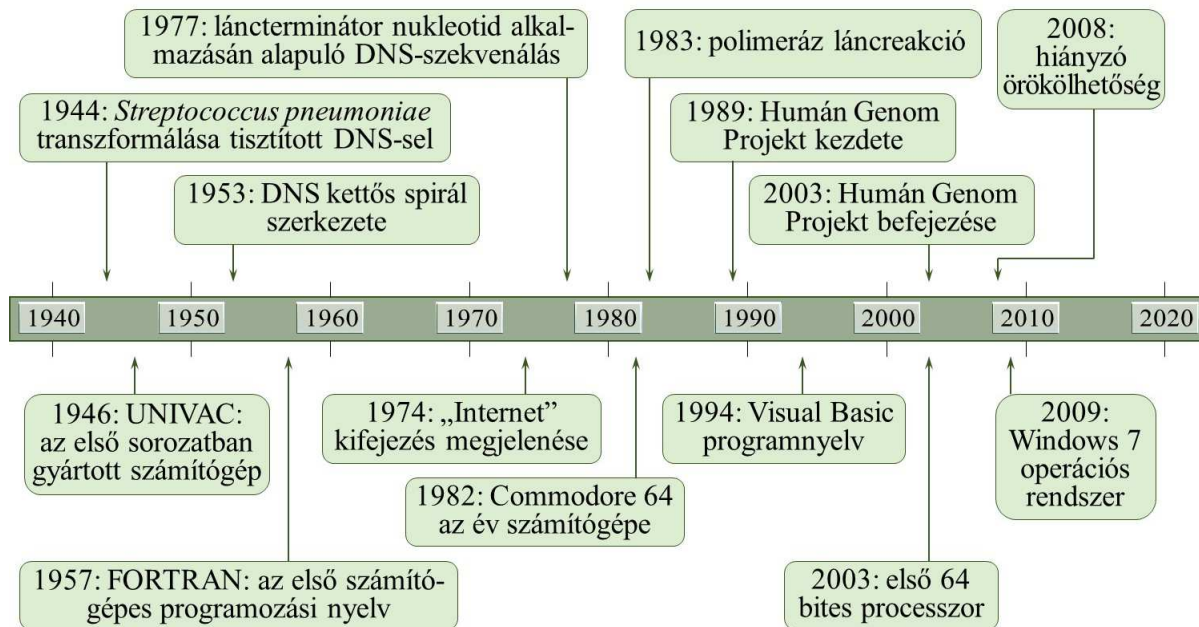
az elválasztó mátrix cseréje, a minták betöltése, az elválasztás, az adatgyűjtés és -elemzés (szekvencia leolvasása) teljes folyamata a felhasználó beavatkozása nélkül, automatikusan történik [17].

A kutatócsoportok – részben kompetitív – munkája, valamint az elméleti, gyakorlati és technikai vívmányok eredményeképp 2001-re elkészült az emberi genom szekvenciájának első verziója. Ez a szekvencia a teljes genetikai információ kb. 90%-át tartalmazta, és a hiányzó szekvenciák mellett a már meglévők pontos fizikai elhelyezkedése sem volt még teljesen ismert [14, 20]. Ennek ellenére a rendelkezésre álló – hatalmas mennyiségű – adat számos jelentős megfigyelés alapjául szolgált. Kiderült, hogy a fehérje-kódoló gének száma a korábban feltételezett kb. 100 000-nél jóval alacsonyabb. Számos információ vált ismertté az ismétlődő régiókkal kapcsolatban, ezek kiterjedését a teljes genom több mint felére becsülték, ami az eddig vizsgált genomokhoz képest kiemelkedően magas hányad volt. Az ismétlődő DNS-szakaszokat elhelyezkedésük és kiterjedésük alapján öt csoportba – transzpozon eredetű ismétlődések, működő gének részlegesen áthelyeződött inaktív szakaszai (pseudogének), egyszerű szekvencia ismétlődések, szegmentális duplikáció, tandem ismétlődések – sorolták, és leginkább biológiai funkcióval nem rendelkező genomi régióknak tartották. Felfigyeltek emellett a CG-dinukleotidok számos sajátosságára is: a szekvencia megismerésével szembeötlő volt, hogy gyakran a gének 5' régiójában helyezkednek el, és számuk a statisztikailag vártnak csupán kb. 1/5-e, aminek hátterében az áll, hogy a CG-dinukleotidok citozinje gyakran metilálódik, és dezaminálódás révén timinné alakul. Fény derült emellett a genom további variációira is: már ebben a kezdeti fázisban közel 1,5 millió SNP-t azonosítottak [14].

Az elkövetkező néhány év fő feladata a hiányzó szekvenciák meghatározása és a hibás régiók bázissorrendjének pontosítása volt. Ezen munka eredményeként 2003 áprilisában – pontosan ötven évvel Watson és Crick DNS kettős hélix szerkezetéről szóló közleményének megjelenése után – bejelentették a Humán Genom Projekt sikeres befejezését. Ez a genomi szekvencia („Build 35”) már csupán 341 (többségében ismétlődő szakaszok környezetében elhelyezkedő) hiányzó régiót tartalmazott, 99,99% pontosságúnak tartották. A szekvencia elemzése alapján a fehérjekódoló gének számát már csak 20 000–25 000-re becsülték [21].

Bár az informatika robbanásszerű fejlődése az élet csaknem minden területén meghatározó változásokat hozott, a genetikai információ megismerésében játszott szerepe mégis kiemelkedőnek mondható (*1. ábra*). Az emberi genomot alkotó 3 milliárd nukleotidnyi DNS-szekvencia

kb. 750 MB adatnak felel meg (mivel a nukleotid csak négyféle, egy byte pedig 256-féle lehet, így egy byte-tal 4 nukleotid kódolható), ekkora mennyiségű adat tárolása az 1990-es évek elején még viszonylag komoly feladatnak mondható. Napjainkban – azaz csupán 3 évtized elteltével – egy átlagos mobiltelefonon 50–100 személy teljes genomját tárolhatjuk.



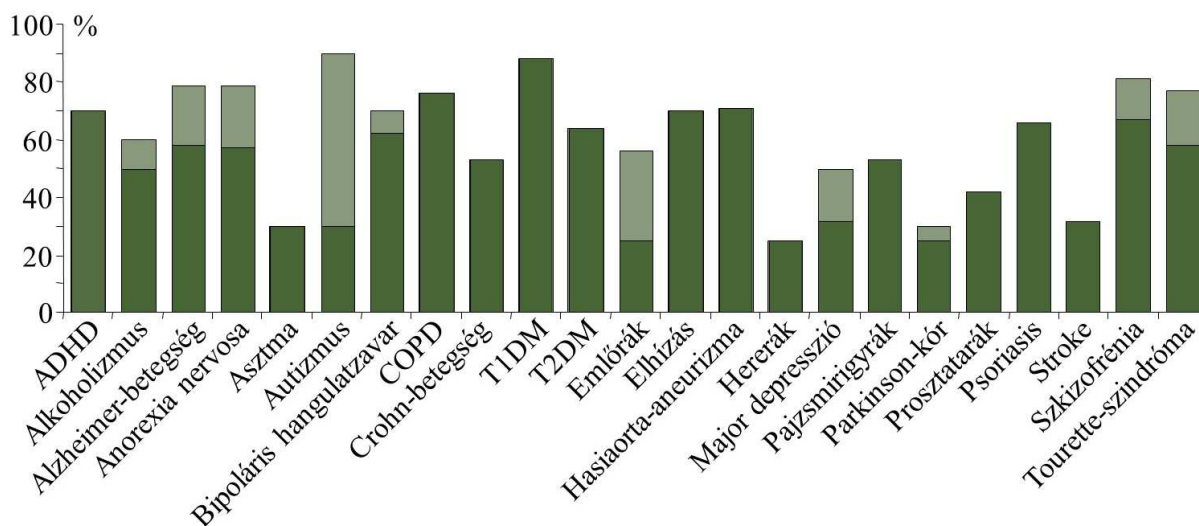
1. ábra. A genetika és az informatika fejlődésének néhány kiemelkedő mérföldköve

1.1.3 A hiányzó örökölhetőség

A Humán Genom Projekt befejező fázisában, illetve a lezárását követő években a munka vezető kutatói igen optimista várakozásokat fogalmaztak meg az eredmények rövid távú, orvosi, gyakorlati alkalmazását illetően. Az elkészülő emberi genom szekvenciájára „A Humán Biológia kézikönyve”-ként tekintettek, és abban reménykedtek, hogy ebből a könyvből a betegségek hátterében álló gének és genetikai variációk, a betegségekre való egyéni örökletes hajlam, a gyógyszerek iránti különböző érzékenység rövid időn belül kiolvasható lesz [22, 23]. Bár az emberi genom szekvenciájának ismerete a humán genetikai kutatások kétségtelen mérföldköve, néhány év alatt világossá vált, hogy a fenti célok meghatározása hosszabb – még napjainkig is tartó – munkát jelent [24].

Régóta ismert, hogy a komplex kórképek kialakulásában a környezeti tényezők mellett az örökletes faktorok is döntő szerepet játszanak, más szóval – sok esetben – meglehetősen magas

heritabilitás értékkel rendelkeznek (2. ábra) [XV]. A humán genom szekvenciájának megismerését követő 5–10 évben azonban a feltárt örökletes komponensek az egyes betegségek esetében a heritabilitás csupán 6–50%-át magyarázták [25, 26]: a fennmaradó, még ismeretlen örökletes tényezők összességét hiányzó örökölhetőségnek („missing heritability”) nevezték el [27].



2. ábra. Néhány betegség heritabilitás értéke. A két eltérő színű oszloprész az egyes forrásokból származó legkisebb (sötétzöld) és legnagyobb (világoszöld) értéket ábrázolja. ADHD: figyelemhiányos hiperaktivitás, COPD: krónikus obstruktív tüdőbetegség, T1DM, T2DM: 1-es illetve 2-es típusú diabetes mellitus. Forrás: <https://snpedia.com/index.php/Heritability>, <http://tga.nig.ac.jp/h2db/>, [28-31].

Kiderült tehát, hogy noha „a” genom ismerete nélkülözhetetlen alap, jelentésének megértéséhez azonban szükséges, hogy az egyes emberek genetikai információja, illetve legalább a köztük lévő különbségek feltérképezésre kerüljenek. Ezek az eltérések – a genetikai variációk: mutációk és polimorfizmusok – felelnek ugyanis a fenotípusban megjelenő különbözőségegek egy részéért, és így a betegségek örökletes hátteréért. A genetikai információ megismeréséhez így az informatikai, a genetikai és a molekuláris biológiai eljárások párhuzamos fejlődése (1. ábra) továbbra is nélkülözhetetlen maradt, melynek eredményeként mára már egy személy teljes genomjának meghatározása csakúgy, mint a kapott szekvencia elemzése, feldolgozása elérhető közelségbe került.

1.1.4 A genetikai variációk feltérképezése

A 2002-ben alapított HapMap Projekt fő célul a gyakori ($MAF \geq 0,05$) SNP-k azonosítását, illetve a polimorfizmusok közötti kapcsolatot: az SNP-k haplotípusának feltérképezését tűzte ki. Már ekkor ismert volt, hogy a gyakori variációk a komplex jellegek és betegségek genetikai rizikófaktorai. Nagy jelentőséget tulajdonítottak a polimorfizmusok közötti LD (linkage disequilibrium – „kapcsoltsági egyensúly hiánya”) érték meghatározásának. Ez az információ nem csak a kromoszómákon lévő rekombinációs helyek („hotspot”), valamint az SNP-k korának meghatározása szempontjából tűnt érdekesnek, hanem azért is, mert ebben az időben a szekvencia analízise, a polimorfizmusok genotípusának meghatározása nagy kihívásnak, idő- és költségigényes munkának bizonyult. Úgy vélték, hogy az egyes régiókra jellemző LD-értékek ismeretében 200 000–1 000 000 megfelelően kiválasztott ún. tag-SNP elemzése elegendő a haplotípusok, és így az azokat alkotó polimorfizmusok genotípusának meghatározásához [32].

A projekt első és második fázisában 270, különböző populációhoz tartozó személy DNS-mintájának elemzését tűzték ki célul, mivel ismert volt, hogy az egyes polimorfizmusok allél- és genotípusgyakorisága jelentős földrajzi változatosságot mutat. A minták részben nem-rokon személyektől, részben pedig családi trióktól (két szülő és felnőtt gyermekük) származtak. A családi triók genotípus elemzése a Mendel-szabályok alkalmazásával nagy segítséget jelentett a haplotípusok megalkotása során. A vizsgálatba 30 Ibadanból, Nigéria fővárosából származó joruba etnikai csoportba tartozó triót (YRI), 30 utahi, észak-, illetve nyugat-európai származású családot (CEU), valamint 45 nem-rokon pekingi, han kínai (CHB) és 45 nem-rokon tokiói japán személyt (JPT) vontak be. Elsődleges célként a genomban egymástól kb. 5 000 bázispár távolságban elhelyezkedő 600 000 SNP elemzését tűzték ki, amit további kb. 400 000 lókuszt vizsgálataival terveztek kiegészíteni azokon a régiókon, ahol a haplotípus térkép megalkotásához nagyobb felbontás szükséges. A genotípus elemzést Japán, az Egyesült Királyság, Kína, Kanada, Nigéria és az USA 18 kutatóintézetében végezték el, az egyes centrumok az összes minta meghatározott kromoszómáinak vizsgálatáért feleltek. A munka során 5 különböző genotipizáló eljárást alkalmaztak. A megbízhatóság ellenőrzése céljából a projekt kezdetén a 90 európai származású minta 1 500 kiválasztott SNP-jének genotípusát valamennyi kutatóhelyen meghatározták. Emellett a mérések során minden 96-os tálcán öt DNS-minta duplikátumként szerepelt, valamint a családi triók vizsgálata a Mendel szabályok ismeretében szintén lehetőséget kínált a genotipizálás megbízhatóságának kontrolljára. Végül a minőség-ellenőrzés további

lépését jelentette, hogy a független minták esetén az eltérő kromoszómákon elhelyezkedő SNP-ekre a Hardy–Weinberg egyensúlynak teljesülnie kell [32].

A HapMap projekt harmadik fázisa során bővült mind a bevont polimorfizmusok, valamint az elemzett minták száma és eloszlása is. Az elemzés kiterjedt kb. 1,4 millió, részben a korábban alacsonyabb MAF értékkel rendelkező SNP genotipizálására. A vizsgálat során alkalmazott Affymetrix és Illumina hibridizációs „chip” eljárás emellett lehetővé tette több mint 1 600 nagy kiterjedésű CNV (ismétlődési szám variáció – copy number variation) azonosítását és meghatározását is. Az elemzett minták száma 1 184-re emelkedett, a résztvevők az eddigi négy mellett további 7 populációt – afrikai származású délnyugat-amerikai (ASW), Denverből származó kínai (CHD), gujatarai népcsoporthoz tartozó texasi (GIH), Kenya nyugati részén élő Luhya etnikai csoporthoz (LWK), valamint Kenya északi, középső és déli részén élő Maasi csoporthoz (MKK) tartozó; mexikói származású Los Angeles-ben élő (MXL) és olaszországi, toszkán (TSI) személyek – képviseltek [33]. Ezen adatok megismerése és szabad hozzáférhetősége (<https://www.sanger.ac.uk/resources/downloads/human/hapmap3.html>) a komplex kórképek genetikai hátterét kutató asszociáció elemzések fontos bázisává vált.

A DNS-szekvencia elemzésének gyors fejlődése hamarosan lehetővé tette a HapMap Projekt során létrehozott haplotípus térkép felbontásának további növelését. Az 1000 Genom Projekt 14 etnikai csoport 1 092 személyének genomját elemezte új generációs genom-, illetve exom-szekvenálási eljárással. Ebben a vizsgálatban 38 millió SNP-t, 1,4 millió kis, és több mint 14 000 nagy kiterjedésű inzerció–deléció határoztak meg [34]. Bebizonyosodott, hogy különösen az alacsony MAF értékű polimorfizmusok meglehetősen nagy földrajzi eltérést mutatnak, és a genom-szintű elemzés további betekintést engedett a genom nem fehérje-kódoló régióiban elhelyezkedő, biológiai funkcióval rendelkező polimorfizmusok csoportjába is. Az adatok a jelenleg aktuális humán genomba (GRCh38) integráltak mind az Ensembl (http://www.ensembl.org/Homo_sapiens/Info/Index) genom böngésző alkalmazásaival, mind az NCBI oldalain (<https://www.ncbi.nlm.nih.gov/projects/fasftp/1000genomes/>) mindenki számára elérhetők, és számos klinikai genetikai kutatás alapjául szolgálnak napjainkban is [35–37].

A populációk genetikai elemzését célzó orvosi kutatások az Ensembl és az NCBI forrásai mellett ma már számos további Internetes adatbázisra támaszkodhatnak. A gnomAD (Genome Aggregation Database, <https://gnomad.broadinstitute.org/>) adatbázis 2014-ben ExAc (Exome

Aggregation Consortium) néven 60 000 személy exomszekvenálási eredményét tartalmazta, a jelenlegi verzióban több mint kétszerannyi exom-, és közel 16 000 teljes genom-szekvencia érhető el. A feldolgozott minták különböző etnikai csoportokból származnak [38]. A Kaviar (<http://db.systemsbiology.net/kaviar/>) 35 projekt eredményeit egyesítve 77 781 nem-rokon személy genomja alapján több mint 200 millió polimorfizmus (162 millió SNP és 50 millió rövid inzerció / delécio) adatait tartalmazza [39]. A DGV (Database of Genomic Variants, <http://dgv.tcag.ca/dgv/app/about?ref=GRCh37/hg19>) az egészséges személyekben előforduló, legalább 50 bp hosszúságú genomi szekvenciát érintő polimorfizmusok katalógusa [40]. A Mitomap a mitokondriális DNS adatbázisa [41], az OMIM (Online Mendelian Inheritance in Man[®] <https://omim.org/>) és az NCBI ClinVar [42] adatbázisa pedig a gyakori és ritka betegségekhez, illetve jellegekhez köthető genetikai variációk tárháza.

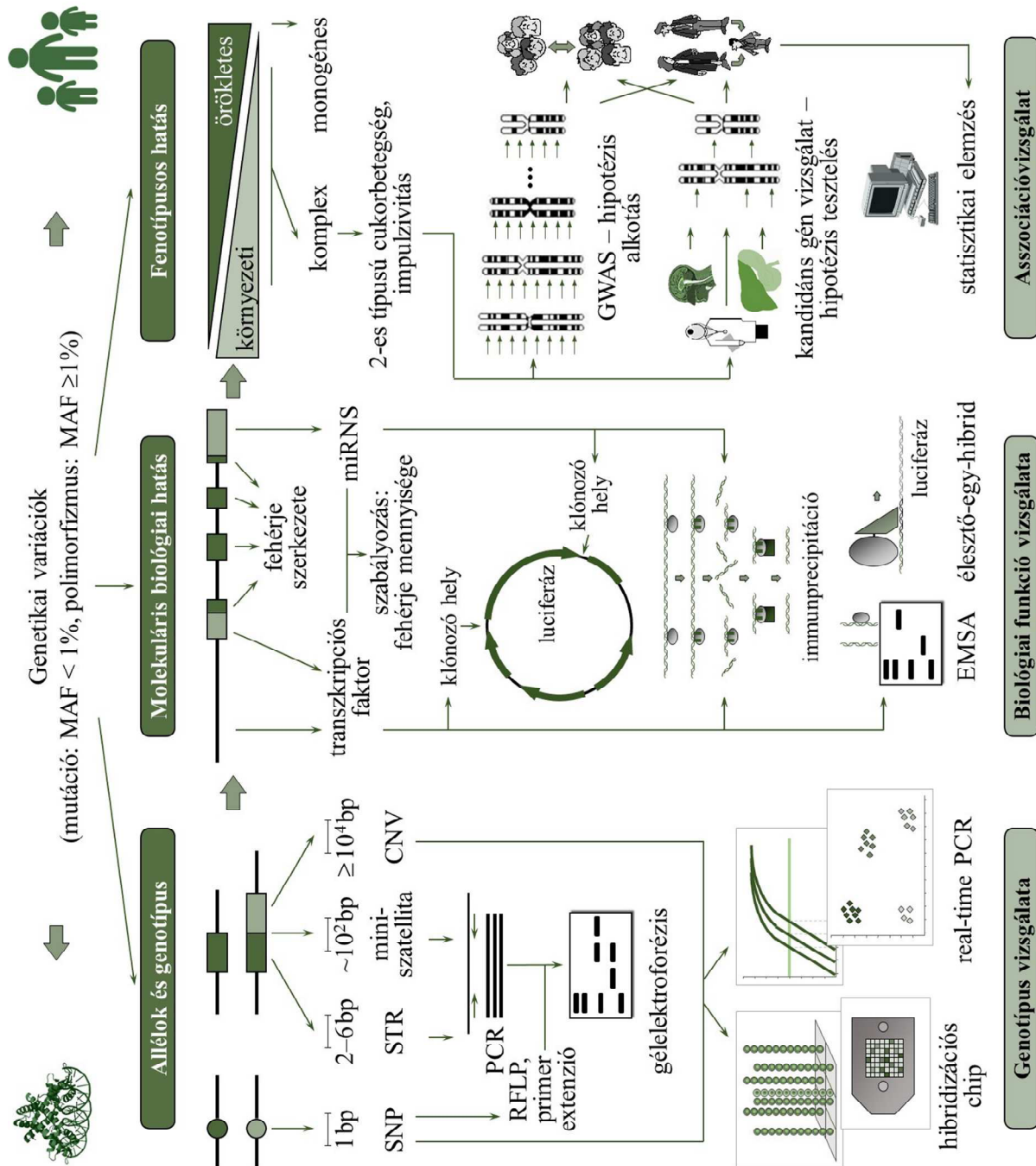
Látható, hogy az elmúlt valamivel több mint fél évszázad a genetikai információ megismerésében igen nagymértékű előrelépést hozott, és – ahogy a fenti összefoglaló is mutatja – rendkívüli mennyiségű adat és tudás érhető már el. Ennek eredményeként számos ritka, monogénes betegség oka ismert, és egyre többet tudunk a gyakori, komplex módon öröklődő betegségek hátterében meghúzódó genetikai tényezőkről is. Bizonyos – ma még nem túl nagyszámú – kórkép (pl. 1-es és 2-es típusú cukorbetegség, Alzheimer-kór, emlő- és prosztaták, gyulladáshoz vezető bélbetegségek) esetén a genetikai faktorok ismerete alapján a betegségek kialakulására való kockázat is megbecsülhető [43].

2 IRODALMI ÁTTEKINTÉS

Két nem-rokon személy genomja között megközelítőleg 0,5%-nyi eltérés mutatkozik. Ezek a különbségek a genetikai variációk, melyeket előfordulási gyakoriságuk (MAF: ritka allél gyakorisága) alapján hagyományosan mutációknak (MAF < 1%) vagy polimorfizmusoknak (MAF \geq 1%) nevezünk. Ismert azonban, hogy a ritka allél frekvenciája jelentős földrajzi, etnikai különbségeket mutat, így esetenként célravezetőbb összefoglaló néven genetikai variációkról beszélni.

2.1 A genetikai variációk osztályozása

A genetikai variációkat – szerkezeti szempontból – több csoportba sorolhatjuk (3. ábra). Legnagyobb számban az egypontos nukleotid polimorfizmusok, illetve a pontmutációk fordulnak elő, amelyek rendszerint a genom egyetlen nukleotidjának kicserélődését jelentik, de ide tartoznak az egyetlen nukleotidot érintő inzerciók és deléciók is. Az NCBI dbSNP-adatbázisában (<https://www.ncbi.nlm.nih.gov/snp/>) 2022 januárjában több mint 957 millió humán polimorfizmus található, figyelemre méltó, hogy ez a szám négy évvel ezelőtt csupán 672 millió volt [XV]. Genomunkban az SNP-k és pontmutációk mellett számos ismétlődési variáció is előfordul. Noha ezek száma alacsonyabb, kiterjedésük és a variációs lehetőségek ugyanakkor nagyobbak, így az SNP-knél jelentősebb mértékben járulnak hozzá az emberek közötti genetikai különbözőséghez. A viszonylag rövid szakaszokból álló ismétlődések a VNTR-ek (variable number of tandem repeats), ezek egyik alcsoportját képezik a legrövidebb szakaszok (2–6 bp) ismétlődéséből álló mikroszatelliták (más néven STR-ek, azaz „short tandem repeat”-ek), valamint a valamivel hosszabb (10–100 bp nagyságrend) miniszatelliták [44]. Az ismétlődési variációk különleges csoportját alkotják a CNV-k (copy number variation), melyek igen hosszú (10^4 – 10^5 bp) DNS-szakaszok ismétlődését jelentik. Kimutatták, hogy ezek a polimorfizmusok – a kisebb kiterjedésű genetikai variációkhoz hasonlóan – egészséges személyek genomjában is előfordulnak [45, 46].



3. ábra. Irodalmi áttekintés. A genetikai variációk – mutációk és polimorfizmusok – gyakran egyetlen nukleotid (SNP) máskor akár 10^4 – 10^5 bp hosszúságú DNS-szakaszok variabilitását jelentik. Genotipizálásuk rendszerint PCR-alapú eljárásokkal történik. Molekuláris funkciójuk kiterjedésüktől, illetve lokalizációjuktól függ: a nem kódoló szakasz polimorfizmusai a génkifejeződés szabályozásában játszhatnak szerepet. A polimorfizmusok biológiai funkciója különböző molekuláris biológiai eljárásokkal, a komplex jellegek genetikai rizikófaktorai pedig GWAS vagy kandidáns gén elemzésekkel azonosíthatók.

2.2 A genetikai variációk biológiai hatása

2.2.1 A genetikai variációk hatása molekuláris szinten

Mind az egyetlen bázispárt érintő SNP-k és pontmutációk, mind a hosszúság variációk molekuláris biológiai funkciója – a lokalizációval összefüggésben – igen eltérő lehet (3. ábra). Kezdetben a legtöbb tanulmány a fehérjét kódoló régiókban elhelyezkedő SNP-eket és pontmutációkat kutatta, mivel ezek esetében közvetlen összefüggés vezethető le a polimorfizmus és a keletkező fehérje aminosavsorrendje között. Legjelentősebb funkcióvesztést a stop kodont létrehozó nonsense variációk, illetve a hibás intronkivágódást eredményező báziscserék okoznak, ilyenkor ugyanis az eredeti fehérje jelentős része hiányzik. A kódoló régióban elhelyezkedő rövid ismétlődési variációk (VNTR-ek) hatását az ismétlődő modul hossza alapvetően befolyásolja. A transzláció alapelvéből adódóan (három nukleotid kódol egy aminosavat) ha ez az érték hárommal osztható, ismétlődő polipeptid szakasz keletkezésével számolhatunk. Ellenkező esetben viszont a leolvasási keret eltolódik, így a polimorfizmust követő génszakasz jelentése is megváltozik, ami gyakran a nonsense és a splice zavart okozó mutációkhoz hasonlóan súlyos funkcióvesztéssel jár. A missense variációk a polipeptidlánc egyetlen aminosavát változtatják meg, ennek funkcionális következménye – az érintett aminosav lokalizációjának és a csere jellegének, az új és eredeti aminosav kémiai hasonlóságának függvényében – igen széles skálán mozoghat. Ismert példa arra is, hogy egy báziscsere közvetett hatással rendelkezik. Vonatkozhat ez a missense variációk mellett a szinonim vagy más szóval „same sense” SNP-kre is, melyek bár a fehérje primer szerkezetét nem módosítják, mégsem tekinthetők biológiai szempontból minden esetben hatástalannak. Egy missense vagy szinonim polimorfizmus az mRNS élettartamát is megváltoztathatja, máskor a módosult primer szerkezetű fehérje el sem jut a rendeltetési helyére, hanem lebomlik [47, 48].

Biológiai, orvosi szempontból – ez utóbbiakhoz valamelyest hasonlóan – távolról sem hagyhatók figyelmen kívül azok a polimorfizmusok, melyek fehérje kódoló szakaszok nem transzlálódó régióiban, illetve nem fehérje kódoló génekben helyezkednek el. Így van ez önmagában amiatt is, mivel maga a genom legnagyobb része ebbe a kategóriába sorolható. Míg DNS-ünk csupán 1,2–1,5%-a felel ténylegesen a fehérjék aminosavsorrendjének leírásáért [44], legalább egy szövetben adott körülmények között a genom akár több mint 90%-a átíródik RNS-re [49, 50]. Az itt elhelyezkedő genetikai variációk biológiai hatása azonban legtöbb esetben a

DNS-szekvencia elemzése alapján csak alig jósolható meg, így sok esetben még ma is izgalmas kihívás ezen polimorfizmusok szerepének feltárása. Molekuláris biológiai jelentőségük, valamint a bemutatásra kerülő kutatási eredmények szempontjából kiemelhetők a gének promotereiben, valamint 3' UTR-ében elhelyezkedő polimorfizmusok.

A fehérjék túlnyomó többségének kifejeződése és működése az eukarióta sejtben pontos szabályozás alatt áll, ami több szinten (transzkripció, poszttranszkripció, transláció, poszttranszláció) valósul meg, és a genetikai polimorfizmusok hatása nélkül is igen összetett folyamat. A szabályozási lépésekben szerepet játszó komplexek kialakulása önmagában összetett, ráadásul nem ritka az egyes szintek közötti kapcsolat, ami akár klinikai jelentőségű is lehet: egy miRNS befolyásolhatja egy transzkripciós faktor kifejeződését [51], máskor egy hosszú, nem kódoló RNS (lncRNS) köthet meg egy miRNS-t, így az lncRNS transzkripció szintű szabályozása hatással van az adott miRNS és célgénjeinek interakciójára [52]. A genetikai polimorfizmusok ezeket az önmagukban is igen bonyolult folyamatokat tovább modulálják. Akár a transzkripciós faktor génjének kódoló szekvenciájában lévő variáció, akár a célgén promotereiben, az adott transzkripciós faktor kötőhelyében elhelyezkedő polimorfizmus befolyásolhatja az mRNS-szintézis aktivitását, és ismertek az ún. CpG-SNP-k is, melyek az epigenetikai, metilációs mintázat módosítása révén hatnak a transzkripcióra [53].

A kb. 22 nukleotid hosszú érett miRNS-ek leggyakrabban az mRNS-ek 3' UTR-hez, ritkábban 5' UTR-éhez vagy kódoló szakaszához kötődve fejtik ki szabályozó hatásukat. A miR-TarBase adatbázis több, mint 360 000 kísérletesen validált, a mirWalk pedig ennél is lényegesen nagyobb számú, többségében csupán jósolt mRNS–miRNS interakciót tartalmaz, ami jól mutatja ezen szabályozás rendkívül összetett, hálózatos jellegét. A kölcsönhatások jelentős része azonban csupán szekvencia illesztés alapján jósolt. Bár tekintettel arra, hogy RNS–RNS kapcsolatról van szó, feltételezhetnénk, hogy a molekulák bázissorrendje alapján a szabályozás hatékonysága könnyen előre jelezhető, a folyamat azonban mégis jóval bonyolultabb. Ennek oka részben az, hogy az mRNS és a miRNS kapcsolódását döntően, a miRNS 5' végétől számított 2–8. nukleotid közötti szakasz, az ún. „seed” régió határozza meg, amit azonban a miRNS 3' végi régiója modulálhat. Ehhez adódik hozzá a polimorfizmusok hatása: a két RNS közötti kapcsolódás hatékonysága ugyanis értelemszerűen megváltozik, ha azok szekvenciája módosul. Érdekes azonban, hogy míg a célgén 3' UTR-ében, a miRNS-ek kötőhelyében viszonylag nagy számú SNP (ún. miR-SNP) található, addig miRNS-ek génjében, különösen a

„seed” a szekvenciát kódoló szakaszokban kevesebb polimorfizmus van [54], ami talán a miRNS-eken keresztül megvalósuló szabályozás fontosságára utalhat.

2.2.2 A genetikai variációk hatása a fenotípusra

A jelentős biológiai funkcióvesztést okozó mutációk gyakran monogénes módon különböző súlyosságú öröklődő betegséghez vezetnek. Mivel ezekben az esetekben lineáris és direkt összefüggés figyelhető meg a genetikai variáció (rendszerint mutáció) és a fenotípus között, így a legtöbb ilyen kórkép, illetve a háttérben meghúzódó genetikai ok mára már ismertté vált. A század eleje, a humán genom szekvenciájának feltárása óta ugyanakkor ugrásszerűen megnőtt azoknak a kutatásoknak a száma, melyek a multifaktoriális, komplex módon öröklődő betegségek genetikai rizikófaktorainak azonosítását célozzák (3. ábra). Hamar kiderült, hogy a kihívás nagyobb [27], mint azt a Humán Genom Projekt lezárásakor remélték. A komplex jellegek kialakításában szerepet játszó környezeti és genetikai tényezők ugyanis bonyolult hálózatként egymást is befolyásolva rendkívül összetett rendszert alkotnak, melyben egy-egy alkotóelem (akár környezeti, akár genetikai faktor) hatása kicsi, és önmagában rendszerint sem nem elégséges, sem nem szükséges feltétele az adott fenotípus megjelenésének.

2.2.2.1 A D4-es dopamin receptor

A pszichogenetikai asszociáció vizsgálatok úttörő munkája a D4-es dopamin receptor (*DRD4*) gén 3. exonjában lévő 48 bp-os VNTR és az „újdomságkeresés” személyiségjegy közötti kapcsolat leírása volt [55].

A dopamin a katekolaminok közé tartozó ingerület-átvivő, szerepet játszik a mozgásszabályozásban, a neuroendokrin működésben, és befolyásolja az érzelmek, indulatok kifejeződését. Ennek megfelelően idegrendszerünkben három fő dopaminerg rendszer különíthető el. A substantia nigrából a nucleus caudatus és a putamen felé irányuló pálya a mozgások kivitelezésében nélkülözhetetlen. A mezolimbikus pálya a prefrontális kéreghez, az amygdalához és a hippocampushoz vezet, a limbikus rendszer részeként fontos az érzelmek, indulatok szabályozásában. Az endokrin funkciók szabályozója a hypothalamus nucleus arcuatusából a hypophysishez vezető pálya.

A dopamin receptorok hét transzmembrán régióval rendelkeznek. A DRD4 agonista kötődésekor a G_i fehérjét aktiválja, ennek következtében a sejten belül az adenil-cikláz gátlódik, így a cAMP-szint csökken.

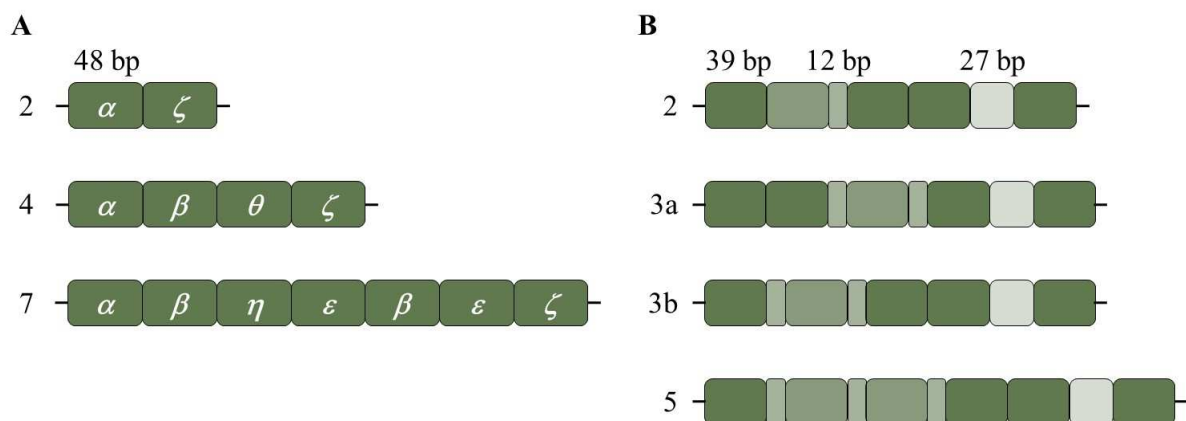
A dopamin rendszer és a DRD4 – ennek megfelelően – számos jelentős klinikai vonatkozással rendelkezik. A szkizofrénia hátterében központi idegrendszeri dopamin túlsúly áll, a betegek agyában post mortem elemzések tanúbizonysága szerint a DRD4 mennyisége jelentősen emelkedett [56], és a kezelés során használt atípusos antipszichotikum, a clozapin is ehhez a receptorhoz kötődik legmagasabb affinitással [57].

A gén 3. exonjában elhelyezkedő 48 bp-os ismétlődési polimorfizmust napjainkig nagy érdeklődés övezi. Ez a génrégió a harmadik citoplazmatikus hurok kódolásában játszik szerepet, melynek funkciója a G-fehérjével való kapcsolat. Az ismétlődési szám 2 és 10 között változhat, ami a polipeptidláncban egy 16 aminosav hosszúságú szakasz ismétlődését eredményezi, ahol ez egyes modulok egymástól kis mértékben különbözhetnek (4. ábra) [58]. Meglepő, hogy a polimorfizmus a receptorfehérje funkció szempontjából egyik alapvető részét érinti, az egyes formák – a legrövidebbtől a leghosszabbig – mégis mind működőképesek, ugyanakkor a gyakori „hosszú” változat, a 7× allél által kódolt fehérje hatékonysága kb. fele akkora, mint a leggyakoribb 4× vagy a 2× változatú receptoré [59]. Érdekes, hogy a polimorfizmussal asszociációt mutató fenotípusok és betegségek elemzését mindmáig nagy érdeklődés övezi, ugyanakkor további molekuláris szintű vizsgálat a különböző receptorfehérjék működésére vonatkozóan alig található a szakirodalomban [60].

A DRD4 3. exonjában lévő VNTR-rel kapcsolatos korai eredmény a hosszú (7×) változat és a Cloninger féle pszicho-biológiai személyiség-modell újdonságkeresés dimenziója között írt le összefüggést [55, 61]. Az eredményt a későbbiekben nem sikerült minden esetben kimutatni [62, 63], az asszociációt ugyanakkor több tanulmány – köztük egy nemrég megjelent metaanalízis [64] – megerősítette [65, 66]. Ezek az ellentmondó eredmények önmagukban felhívják a figyelmet a genetikai asszociáció vizsgálatok jellemző technikai nehézségére: a komplex jellegek esetén az egyes, önmagukban kis hatású genetikai faktorok kimutatása nem könnyű feladat, ahol mind az álpozitív, mind az álnegatív eredményekre fontos figyelmet fordítani.

A VNTR-t számos további jellel – szexuális érdeklődés [67], különböző személyiségjegyek [68], dohányzás [69] – hozták összefüggésbe, valamint asszociációt mutattak ki a polimorfizmus és egyes pszichiátriai rendellenességek, illetve betegségek, mint a szorongás [70], a figyelemhiányos hiperaktivitás [71] és a szerfüggőség, illetve abúzus [72, 73] között.

Említésre méltó, hogy míg a gyakran modellként alkalmazott rágcsálók hasonló polimorf régióval nem rendelkeznek, a kutya *DRD4* génjének 3. exonjában található egy, a humán VNTR-nek valamelyest megfelelő ismétlődési polimorfizmus. Ez a VNTR azonban nem 48, hanem 27, 39 és 12 bp-os modulokból áll, és a régió variabilitását a 39 és a 12 bp nagyságú elemek változó számú és sorrendű előfordulása adja (4. ábra) [74]. A kutya – emberrel együtt „szocializált” életmódja miatt – a viselkedésgenetikai vizsgálatok során a patkánynál vagy az egéرنél jobb modellnek tekinthető: számos kutyafajtában figyeltek meg asszociációt az egyes allélváltozatok, illetve az állatokra jellemző különböző viselkedésjegyek között [75].



4. ábra. A *DRD4* gén 3. exonjában lévő VNTR leggyakoribb allélvariánsai emberben és kutyában. *A* Az emberben lévő polimorfizmus 48 bp-os modulokból áll, ezek bázissorrendje azonban kis mértékben eltér egymástól [76]. *B* A kutyák ismétlődési variációja 12, 27 és 39 bp nagyságú blokkokból áll, melyek száma és sorrendje is különbözhet.

A *DRD4* génben természetesen számtalan további polimorfizmus található. A gén promóterében a mindmáig „–521C/T SNP”-ként említett rs1800955 polimorfizmus T allélja *in vitro* mérések alapján csökkent mennyiségű mRNS képződéséhez vezet [77]. Az SNP több, az újdonságkereséssel kapcsolatba hozható fenotípus mellett [78, 79] a szkizofrénia genetikai rizikófaktoraként is felmerül [80]. A transzkripció startponttól kissé távolabb található egy 120

bp-os duplikáció, mely sejtes rendszerben hatással van a transzkripció aktivitására [81], befolyásolja a clozapin-kezelés hatékonyságát [82], és – más polimorfizmusokkal együtt – összefüggésben állhat a figyelemhiányos hiperaktivitással [83].

2.2.2.2 A wolframin

A wolframin fehérje, illetve genetikai variációk okozta működésváltozása szintén összefüggésben áll központi idegrendszeri, neuropszichiátriai funkciókkal.

A wolframin az endoplazmás retikulum membránjában elhelyezkedő glikoprotein, 9 transzmembrán szakasszal rendelkezik, N-terminális régiója a citoplazma, C-terminális pedig az endoplazmás retikulum lumene felé helyezkedik el. Csaknem minden szövetben kifejeződik, legmagasabb expressziót a szívben, a központi idegrendszerben [84], valamint a hasnyálmirigy β -sejtjeiben és inzulinoma sejtvonalakon mutattak ki [85]. Számos megfigyelés utal arra, hogy a wolframin fontos szerepet játszik a hasnyálmirigy szigetsejtjeinek inzulintermelésében. Inzulinszekréció és endoplazmás retikulum stressz során a wolframin expressziójának aktivitása megemelkedik: az IRE1 α , valamint a PERK útvonalakon keresztül a fehérje hozzájárul az endoplazmás retikulum egyensúlyának fenntartásához [86]. Mesterségesen létrehozott wolframin hiányos egerekben glükóz hatására az inzulinszekréció nem emelkedik, progresszív inzulinhiány és diabetes mellitus figyelhető meg. Az állatokból izolált szigetsejtjeiben immunhisztokémiai elemzések csökkent β -sejtszámot mutattak ki, magas glükózkoncentráció és endoplazmás retikulum stressz következtében kialakuló apoptózis volt megfigyelhető [87]. Kimutatták, hogy a wolframin nem csak a β -sejtek endoplazmás retikulumában van jelen, hanem a szekretált granulumok membránjában is megtalálható, és a kémhatás szabályozása révén hozzájárul a proinzulin–inzulin átalakuláshoz és az exocitózishoz [88]. Fény derült emellett arra is, hogy a wolframin kalmodulint köt, így számos celluláris fehérjével és Ca^{2+} -függő folyamattal áll kapcsolatban, valamint szerepet játszik a sejten belüli Ca^{2+} -egyensúly fenntartásában [89].

A wolframin fehérje funkciókiesése ismert jelenség emberben: a Wolfram- vagy DIDMOAD-szindrómaként számon tartott ritka, autoszomális recesszív módon öröklődő neurodegeneratív kórkép (OMIM 222300) szerteágazó klinikai jegyeket mutathat: diaabetes insipidus, diaabetes mellitus, optikus atrófia és siketség (deafness) mellett más neuropszichiátriai

tünetek (ataxia, nisztagnus, perifériális neuropátia, mentális retardáció, pszichiátriai rendellenességek) is kialakulhatnak [90]. A betegség klinikai megjelenése igen változatos, leggyakrabban a nem autoimmun, inzulinfüggő diabetes mellitus alakul ki először 6 éves kor körül, s a génben máig azonosítanak különböző manifesztációt eredményező új mutációkat [91]. A prognózis meglehetősen kedvezőtlen, fiatal felnőtt korban a neurodegeneratív folyamatok rendszerint centrális légzési elégtelenséghez vezetnek [92].

A Wolfram-szindróma jellegzetes tüneti alapján vetődött fel annak gondolata, hogy a gén mérsékeltebb funkcióváltozást okozó polimorfizmusai a diabetes mellitus, valamint különböző neuropszichiátriai rendellenességek örökletes komponensei lehetnek. Egy genom szintű asszociációvizsgálat a T2DM (2-es típusú cukorbetegség) és a gén rs4689388 SNP-je között mutatott ki asszociációt [93], emellett az rs1046320 és az rs10010131 polimorfizmusokat szintén a betegség rizikófaktoraként azonosították [94, 95]. Asszociációt írtak le a T2DM és a wolframin kifejeződésének szabályozásában vélhetően szerepet játszó két SNP között: az rs1801214 polimorfizmus a metilációs mintázat módosítása révén, a gén 3' UTR-ében elhelyezkedő rs1046320 SNP pedig miRNS bekötődésének megváltoztatásával járulhat hozzá a sejtben jelen lévő fehérje mennyiségének meghatározásához [96]. Ezek az eredmények egy koncepcionális kérdésre világítanak rá. Egy génen belül egymástól nem túl nagy távolságban elhelyezkedő SNP-k között gyakran LD figyelhető meg. Ez azt is eredményezheti, hogy a genotípus–fenotípus asszociáció olyan polimorfizmusokkal is statisztikailag kimutatható, melyek tényleges biológiai hatással nem rendelkeznek. Egy adott jelleg háttérében álló kromoszóma szakaszok, illetve gének azonosításához ez nem is feltétlen szükséges, ugyanakkor a betegségek kialakulásában szereplő molekuláris folyamatok megértéséhez (és esetleges befolyásolásához) az SNP-k részletes LD-, haplotípus- és funkcionális elemzése elengedhetetlen.

A Wolfram-szindróma tüneteinek megfelelően a gén polimorfizmusai pszichiátriai rendellenességek genetikai rizikófaktoraként is felvetődnek. Hangulatzavarral kapcsolatban pozitív és negatív eredmények egyaránt fellelhetők a szakirodalomban [97, 98], a gén aminosavcserét okozó mutációi szkizofréniával és affektív zavarokkal nem mutattak szignifikáns összefüggést [99], a témában megjelent közlemények száma azonban lényegesen alacsonyabb, mint a WFS1 és a T2DM összefüggését kutató publikációké.

2.2.2.3 A SNAP-25

A szinaptoszóma-asszociált fehérje-25 szintén nélkülözhetetlen a központi idegrendszer megfelelő működéséhez, de érdekes módon – a wolframinhoz hasonlóan – szerepet játszik a hasnyálmirigy β -sejtjeinek inzulin szekréciójában is [100]. A fehérje a SNARE (soluble N-ethylmaleimide sensitive factor attachment protein receptor) komplex részeként részt vesz az intracelluláris vezikulumok exocitózisában. Ennek kulcslépése a vezikulumok sejtmembránnal történő összeolvadása, melyben alapvető a v-SNARE (vezikuláris) és t-SNARE (target) fehérjék közötti nagy affinitású kölcsönhatások létrejötte [101]. A SNAP-25 a szintaxin 1-gyel együtt a t-SNARE alkotója, nem rendelkezik transzmembrán régióval, csupán egy palmitát segítségével kötődik a membránhoz [102]

A SNAP-25 fehérje funkcióját állatkísérletek eredményei is magyarázzák. A *coloboma* egér a 2. kromoszómán egy 4,2 Mb hosszúságú deléciót hordoz, mely magában foglalja a SNAP-25 gént is. A deléció homozigóta formában az élettel nem bizonyult összeegyeztethetőnek: az állatok a 6. embrionális nap előtt elpusztultak. A delécióra nézve heterozigóta egerek életben maradtak, azonban lokomotoros aktivitásuk lényegesen magasabb volt, mint a kontroll állatoké [103].

A SNAP-25 karakterizálása során ismertté vált, hogy a génről két különböző transzkripciós variáns keletkezhet, melyek különböző 5. exont tartalmaznak, ami fehérjeszinten 9 aminosav eltérésében mutatkozik meg. Megfigyelték, hogy az egerek fejlődése során a két izoforma aránya megváltozik: míg a prenatális korban a SNAP-25 *a* izoforma mennyisége magasabb, a születés utáni hetedik napon a *b* változat válik dominánssá [104]. A fehérje *b* izoformájának selektív hiánya (a SNAP-25 azonos össz mennyisége mellett) egerekben fejlődésben való elmaradást és spontán görcsöket okozott, a hippocampusz területén morfológiai eltérések jelentek meg, és a tanulás, térbeli tájékozódás terén az egészséges állatokhoz képest jelentős lemaradás volt megfigyelhető [105]. Egerekből származó minták vizsgálata alapján a két izoforma makroszkópos és szubcelluláris lokalizációja is eltérő: a *b* izoforma kifejeződése az agy egész területére, a szinapszisokban gazdag régiókra, szubcellulárisan pedig a membrán frakcióra jellemző, a SNAP-25 *a* variáns ezzel szemben főképp a rosztrális területeken, a citoszólban fordul elő [106].

A SNAP-25 3' UTR-ében elhelyezkedő 2 SNP-t (rs3746544 és rs1051312) több vizsgálat a figyelemhiányos hiperaktivitás rizikófaktoraként azonosította, és megfigyelték, hogy az

ADHD-hoz kapcsolódó major depresszió jelenléte esetén az asszociáció még kifejezettebb [107, 108]. A SNAP-25 fehérje mennyiségét bipoláris hangulatzavarban szenvedő betegek agyának frontális lebenyében, a 9. Brodmann-féle kérgi mező területéről származó post mortem mintákban vizsgálták, és megállapították, hogy a SNAP-25 mennyisége – mind mRNS-, mind fehérjeszinten – szignifikánsan alacsonyabb, mint a pszichiátriai betegségben nem szenvedő személyekben. Bár hasonló eltérés szkizofréniában szenvedő betegek esetén nem volt megfigyelhető [109], a gén polimorfizmusai és a betegség előfordulása között gyenge összefüggés kimutatható [110].

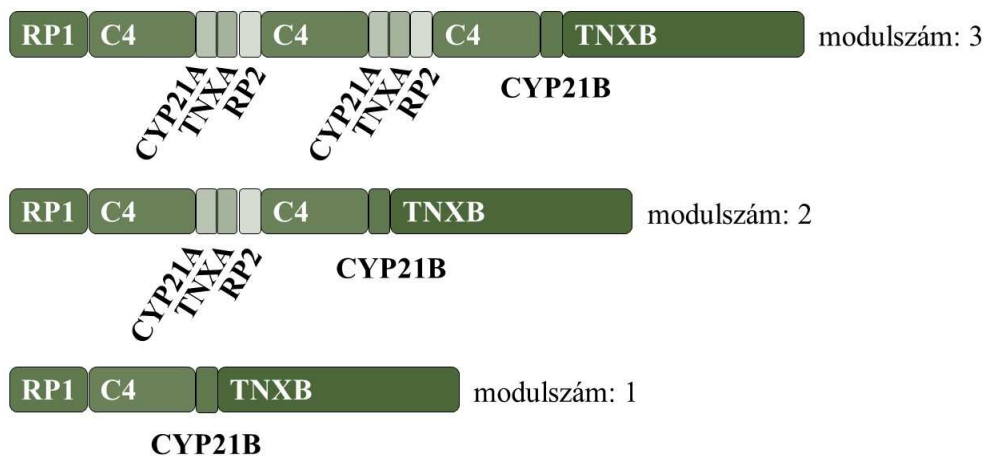
2.2.2.4 A 4-es komplement komponens

Az MHC régióban elhelyezkedő 4-es komplement komponens fehérje génje variabilitását tekintve jelentősen eltér az eddig példaként bemutatott génektől.

A humán DNS szekvenciájának feltárása során, illetve azt követően vált ismertté, hogy genomunkban nagy számban fordulnak elő akár 10^5 bázispár hosszúságú szakaszokat érintő duplikációk, ismétlődések, illetve deléciók, melyeket összefoglaló néven CNV-nek nevezünk. Ezek a polimorfizmusok egészséges személyek genomjában is igen nagy számban előfordulnak [45, 46], nem okoznak jelentős funkcióváltozást, illetve -károsodást, egyenként feltételezhetően kis hatásúak, és – a rövid szakaszokat érintő genetikai variációkhoz hasonlóan – hozzájárulhatnak a komplex jellegek genetikai hátterének meghatározásához. Ilyen CNV található a 6. kromoszóma rövid karján elhelyezkedő MHC-régióban, mely genomterület abból a szempontból is különleges, hogy ezen a szakaszon igen nagy kiterjedésű, rendkívül konzervált ősi haplotípus struktúrák vannak jelen. Az ősi haplotípusok konzervált, folyamatos szekvenciák, amelyek több ezer bázispár hosszúságú „befagyott blokkokból” állnak. Ezekben a régiókban nagyszámú, kapcsolatosan öröklődő polimorfizmus található, és érdekes módon rekombináció csak a blokkok között történik, a rögzült szekvenciák viszont állandósult módon örökítődnek tovább generációról generációra [111].

Az MHC régió génjei három osztályba sorolhatók: a telomerhez közelebb található az MHC I, középen helyezkedik el az MHC III, a 6. kromoszóma centromer szakasza felé pedig az MHC II lokalizálódik. Ez utóbbi régió géneiben – az egész genom viszonylatában is – rendkívül gazdag. Itt helyezkedik el az ún. RCCX-modul, melyet szerkezeti egységet képezve egy nukleáris protein kináz (*RP*), a 4-es komplement komponens (*C4*), a 21-hidroxiláz (*CYP21*) és

a tenaszcin-X (*TNX*) génje alkot. A négy génből álló modul sajátos tulajdonságú ismétlődési variáció jellemzi létrehozva az egy, kettő vagy három kópiát tartalmazó komplexet (5. ábra) [112]. A magi szerin / treonin kináz *RP* génből az ismétlődési számtól függetlenül mindig csak egy aktív (*RPI*), a további modulokban elhelyezkedő *RP2* részleges duplikáció révén jött létre, és nem kódol működő fehérjét. A 21-hidroxiláz enzim (*CYP21*) a glükokortikoid és mineralokortikoid hormonok szintézisének egyik lépését katalizálja. Az előzőhöz hasonlóan minden esetben csak egy blokk kódol működő enzimet (*CYP21B*), az „A”-val jelölt pszeudogénről egy 8 bázispáros deléció következtében kialakuló kereteltolódás, illetve korai stop kodon miatt nem íródik át aktív fehérje. Ugyanígy csak a modulok utolsó tagjában elhelyezkedő *TNXB* génről keletkezik működőképes extracelluláris mátrix fehérje, a *TNXA* kereteltolódást tartalmaz. Ezzel szemben – érdekes módon – a 4-es komplement fehérje (*C4*) génje a modulokat felépítő mindegyik építőelemben aktív, így genomunkban akár 6 működő *C4* gén is jelen lehet, és megfigyelték, hogy a képződő fehérje mennyisége a génszámmal összefüggést mutat [113].



5. ábra. Az „RCCX”-modul. Az ismétlődő elemeket a magi szerin / treonin kináz (RP), a C4-es komplement fehérje (C4), a 21-hidroxiláz (CYP21) és a tenaszcin-X aktív (sötétebb) és inaktív (világosabb) génjei alkotják.

A már önmagában sem szokványos genomi konstellációt tovább tarkítja, hogy a változó számban jelenlévő *C4* géneknek két különböző típusa, a *C4A* és a *C4B* ismert. Így nem csak az abszolút génszám, hanem a két izoforma aránya is variabilitást mutat [112]. Bár a két változat

csupán a 26. exon 5 nukleotidjában tér el egymástól, ami fehérje szinten 4 aminosav különbséget eredményezi, a két fehérje funkciója mégis valamelyest eltérő.

A *C4* gén leginkább a májban fejeződik ki, és a jelen lévő fehérje mennyisége alapvetően a génszámtól függ [114]. A komplementrendszer az immunválasz alapvető résztvevője. A kórokozók opszonizálásában, illetve lízisében, valamint a saját sejtek károsodásának gátlásában egy kaszkád módjára aktiválódó enzimrendszer vesz részt. Ennek során a *C4* fehérje limitált proteolízissel aktívvá válik, és a felszínre kerülő tioészter kötés segítségével kapcsolódni tud az antigénhez, valamint hozzájárul a *C2*-proenzim aktiválásához. A *C4A* és *C4B* változatban az 1101–1106. pozícióban 4 aminosav tér el egymástól, melyek közül a célmolekulával kapcsolatos eltérő kölcsönhatásért elsősorban az 1106. aminosav felel, ami a *C4A* fehérjében aszpartát, a *C4B* izoformában hisztidin. A *C4A* elsősorban aminocsoporttal rendelkező antigénhez kötődik, a *C4B*-ben a hisztidin egy köztitermék kialakulásán keresztül segíti elő a tioészter-kötés tranzacetilációját, de az intermedier hamar lebomlik, ennek az izoformának hidroxil-csoport-hoz magasabb az affinitása [112].

A populáció csaknem háromnegyedében két egységből álló *RCCX* modul található mindkét kromoszómán, melyek egy-egy *C4A* és *C4B* fehérjét kódolnak. Csökkent *C4* mennyiség következtében romolhat a fertőzésekkel szembeni ellenállóképesség, az emelkedett génszám, illetve fehérjemennyiség pedig autoimmun kórképek kialakulását eredményezheti. Nem ritka (25–33%) a két izoforma arányának eltolódása, illetve valamelyik változat teljes hiánya sem. A *C4A* mennyiségének csökkenése a fertőzésekre való fokozott hajlam mellett szisztémás lupus erythematosus (SLE) [115, 116], valamint érkárosodás talaján fiatal korban kialakuló miokardiális infarktus rizikófaktora lehet [117]. A *C4B*-hiány következtében erythema nodosum [118] és IgA-nefropátia [119] kialakulásáról számoltak be. A *C4*-es komplementrendszer rendkívüli genetikai változatossága miatt napjainkban is sok kutatás érdeklődésének középpontjában áll [120]. Összefüggést mutattak ki az alacsony *C4A* géndózis és a pajzsmirigy autoimmun túlműködése, a Graves–Basedow-kór között [121], a megemelkedett *C4A* génszám pedig az időskori makuladegeneráció genetikai rizikófaktora lehet [122], és felvetődött a *C4*-aktiválás és a koronavírusfertőzés talaján kialakuló diffúz mikrovaszkuláris trombózis közötti kapcsolat lehetősége is [120].

2.3 A genetikai variációk elemzése

2.3.1 A genotípus meghatározása

A genetikai mutációk és polimorfizmusok vizsgálata mind a kutató-, mind a klinikai munka feladata lehet. Bármelyikről is legyen szó, nagyszámú minta egy vagy több polimorfizmusának gyors, megbízható és reprodukálható elemzése a cél. Lényeges tehát a megfelelő technika kiválasztása (3. ábra), melynek során szem előtt kell tartani a vizsgálandó polimorfizmusok típusát (SNP, VNTR, CNV) és számát is. Kiszámú lókuszt meghatározása célzott eljárásokkal gyorsan, pontosan, hatékonyan és gazdaságosan megvalósítható, ugyanakkor ma már a klinikai diagnosztika során is helyet kapnak azok a nagy átírási képességű eljárások (pl. NGS), melyek genomszintű elemzést is lehetővé tesznek [123].

2.3.1.1 A genotipizáló módszerek alapelve, VNTR-ek elemzése

A célzott genotipizáló technikák többsége a vizsgált genomszakasz polimeráz láncreakcióval (PCR) történő felszorzásán, és a kapott DNS-termékek elektroforetikus illetve fluoreszcens analízisén alapul. A PCR során az adott genetikai variációt tartalmazó régióról ciklusos folyamatban nagyszámú azonos másolat keletkezik. Mivel a reakciót katalizáló DNS-függő DNS-polimeráz egy új lánc felépítését elkezdni nem tudja, csupán már jelen lévő molekulát képes meghosszabbítani, így mesterséges, egyszálú DNS-molekulákkal, a primerekkel meghatározható az elemezni, illetve felszorzani kívánt régió. Az emberi genom szekvenciájának feltárása ennek megfelelően jelentős előrelépés volt ezen a területen is, mivel a vizsgált genomi régiók bázissorrendjének ismerete elengedhetetlen a primertervezés során. Egy PCR-hez legalább két primerre van szükség: ezek úgy kötődnek a minta-DNS egyik illetve másik láncához, hogy a 3' végük egymás felé tekint, így a ciklusos folyamat során a primerek által közrefogott régióról exponenciálisan növekvő számú másolat jön létre. A képződő DNS-fragmentumok gélelektroforetikus technikákkal vagy real-time PCR esetén a folyamat során képződő fluoreszcens jel mérésével azonosíthatók.

Az egyszerű, horizontális, alámerülő agaróz gélelektroforézist mindmáig széleskörűen alkalmazzák DNS-fragmentum analízisre. A technika közel egy évszázaddal ezelőtti felfedezése óta azonban igen sokat fejlődött, számos hatékony, érzékeny, gyors és automatizált változatát kidolgozták.

A kapilláris elektroforézis több eltérő alkalmazási lehetőséget kínál, melyekhez csak a vizsgálat során használt kapillárist, sőt esetenként csupán a futtató puffert és egyéb anyagokat kell kicserélni. A kapilláris kis belső átmérője (20–180 μm) miatt az elektromos ellenállás igen nagy, ami magas (100–1000 V/cm-es) elektromos térerő alkalmazását teszi lehetővé, mivel az áramerősség és a hőfejlődés még így is alacsony marad. A nagy térerő rövid mérési időt és nagy felbontást eredményez, a valós idejű UV vagy lézer indukált fluoreszcencia elvén megvalósuló detektálás pedig magas érzékenységet biztosít [124]. Egyes berendezések 96 minta egyidejű analizisét is lehetővé teszik, az ilyen nagy áteresztőképességű eszközök jelentősen hozzájárultak a Humán Genom Projekt sikeres megvalósításához.

A technika legegyszerűbb változata a kapilláris zóna-elektroforézis, mely ionok tömeg-töltés arány szerinti elválasztását teszi lehetővé. Ebben a rendszerben számolni kell az elektroozmotikus áramlás létrejöttével, ami a kapilláris falán lévő szilanol csoportok (SiO^-) miatt alakul ki. A negatív kapilláris az oldatban lévő pozitív töltésű részecskéket magához vonzza, az így keletkező „pozitív falú cső” elmozdul a negatív pólus irányába, ami lényegében a kapillárisban lévő folyadék tömegtranszportját jelenti. Ez hozzáadódik az ionok saját mobilitásához, ennek következtében a negatív ionok és a töltéssel nem rendelkező molekulák is a katód felé vándorolnak [125]. DNS analízishez a kapillárist valamilyen géllal kell megtölteni (kapilláris gélelektroforézis), mivel az eltérő méretű molekulák tömeg-töltés aránya azonos, így ezek kapilláris zóna-elektroforézis segítségével nem választhatók el egymástól. A jól reprodukálható elválasztás ebben az esetben viszont az elektroozmotikus áramlás minimalizálását teszi szükségessé: a kapillárisfal kémiai módosítására például metil-cellulóz, poliakrilamid vagy polietilén-glikol [126] alkalmas.

A technika és a molekuláris genetica általános fejlődési tendenciáinak megfelelően a miniatürizálás az elválasztástechnikák körében is tért hódított. Sok minta gyors vizsgálata esetén ez számos előnnyel jár. Lerövidül az elektroforetikus analízisek időtartama (perc–másodperc nagyságrend), kis területen párhuzamosan több minta elemezhető, több lehetőség kínálkozik az automatizálásra, ami számottevő munka-, idő-, anyag- és költségmegtakarítást jelent [127].

A mikrofluidikai eszközök miniatürizált kapilláris elektroforézis berendezéseknek tekinthetők: a minták analízise egy kisméretű üveg- vagy műanyaglap csatornáiban történik. Ez a rendszer a méretcsökkenésből származó előnyökön túl még egy ígéretes lehetőséget kínál: a hagyományos kapilláris elektroforézissel szemben csatornák összetett, két dimenziós hálózata

alkotható meg, ami az elektroforézist megelőző lépések integrálását is lehetővé teszi. Ezen elrendezés segítségével hagyományosan több lépésből álló analízisek egyszerűen, hatékonyan, akár a mintavétel helyszínén megvalósíthatók („lab on chip”) [128]. A lapkában lévő kapillárisok rendszerint 10–40 μm szélesek és 60–200 μm magasak. A csatornák a chip felszínén lévő nyílásokban végződnek, itt történik az oldatok (minta, puffer, elválasztó mátrix) betöltése. Az ide csatlakozó elektródokkal a csatornában különböző nagyságú és polaritású elektromos mezők hozhatók létre, ami nem csak az elektroforetikus elemzést, hanem az anyagok egyszerű mozgását is lehetővé teszi. A legegyszerűbb mikrofluidikai eszközök két egymást keresztező csatornát tartalmaznak. A rövidebben történik a minta bejuttatása, a hosszabban az elválasztás. Természetesen ennél lényegesen bonyolultabb felépítésű rendszerek is léteznek, az alkalmazási lehetőségek száma pedig csaknem végtelen: a „klasszikus” elektroforetikus analíziseken túl vírusok, baktériumok kimutatása, de akár tumor sejtek izolálását követő transzkriptom analízis is megvalósítható [128].

A polimorfizmust tartalmazó genomi régió PCR-rel történő amplifikálásával, majd azt követően a – megfelelő platformon megvalósított – elektroforetikus analízissel a kis kiterjedésű ismétlődési variációk (mikroszatelliták és VNTR-ek) genotipizálása elvégezhető, mivel ebben az esetben éppen az adott DNS-szakasz hosszát kell meghatározni. Ennek megfelelően a vizsgált régió felszorzozható egy primerpárral, mely közrefogja az érintett szekvenciát, majd a PCR-termékek gélelektroforetikus analízisével az ismétlődési szám leolvasható. Technikai nehézséget jelenthetnek azok a polimorfizmusok (mikroszatelliták és rövid inzerciók, deléciók), melyeknél az ismétlődő egység hossza nagyon kicsi, mivel ilyenkor a hagyományos horizontális agaróz gélelektroforézis felbontóképessége nem minden esetben megfelelő az egyes allélok megbízható elkülönítésére. Ilyen genetikai variációk elemzése során a kapilláris elektroforézis [I] vagy a mikrofluidikai eszközök [X] alkalmazása jelentheti a megoldást [129].

2.3.1.2 SNP-k genotipizálása

SNP-k tanulmányozása során ugyanakkor további, illetve eltérő eljárásokra van szükség a genotípus elemzése során. Ennek oka az, hogy a genetikai variáció az érintett DNS-szakasz hosszát nem befolyásolja, így a PCR-rel történő amplifikáció, majd a keletkezett termékek elektroforetikus méret-meghatározása önmagában nem alkalmas a genotípus meghatározására. Számos technikát dolgoztak ki az elmúlt évtizedek alatt, melyekkel a DNS-szekvencia eltérése,

azaz az SNP gélelektroforézissel kimutatható hosszúságkülönbséggé vagy valamilyen fluoreszcens szignállá konvertálható.

Az RFLP (restrikciós fragmentum hosszúság polimorfizmus) lényege a II-es típusú restrikciós endonukleázok alkalmazása. Ezek az enzimek igen specifikusan, egy adott szekvenciánál hasítják a dupla szálú DNS mindkét láncát. Ha egy SNP vagy pontmutáció egy endonukleáz felismerőhelyén található, akkor az adott szekvenciát az egyik allél létrehozza, a másik viszont elrontja, így a különböző genotípusoknak megfelelően eltérő hasítási mintázat jön létre, ami gélelektroforézissel megjeleníthető.

A „Surveyor” nukleáz egy különleges, DNS-t hasító enzim, ami szintén alkalmas SNP-k tanulmányozására. Jellemzője, hogy a DNS-t azokon a pontokon hasítja el, ahol hibás bázispár van jelen, bár említésre méltó, hogy az enzim affinitása az egyes szabálytalan bázispárok iránt valamelyest eltérő. Az eljárás lényege, hogy a PCR után a mintákat ismert bázissorrendű fragmentummal hibridizáltatják, így ott, ahol a vizsgált minta szekvenciája eltér a referenciától, kialakulhatnak a hibás bázispárok. Gélelektroforézis segítségével így – a hasítási mintázat alapján – megállapítható, hogy volt-e eltérés a két minta szekvenciája között. A módszer ebből adódóan – sok más technikával ellentétben – a vizsgált szakaszon belül bármilyen, akár eddig ismeretlen mutáció vagy polimorfizmus azonosítására alkalmazható. Az újonnan kimutatott genetikai variáció pontos helye és a megjelenő szekvencia viszont csak további módszerekkel (leggyakrabban DNS-szekvenálással) határozható meg [130].

Az allélspecifikus amplifikáció (ARMS – Amplification-refractory mutation system) alapelve olyan primer használata a PCR során, melynek 3' vége éppen a vizsgált SNP pozíciójára esik [131]. Emellett olyan DNS-polimerázt szükséges alkalmazni, mely nem rendelkezik 3' exonukleáz aktivitással, ellenkező esetben épp a technika lényege vész el, ha az enzim a primer utolsó, nem komplementer nukleotidját lehasítja. Ebben a rendszerben csak akkor keletkezik PCR-termék, ha a reakcióelegyben lévő allélspecifikus primer 3' végével komplementer allél van jelen a vizsgált lókuszon. A primer hossza a genotipizálás megbízhatóságát nagyban befolyásolja, túl hosszú (több mint 20–22 nukleotidból álló) primer alkalmazása álpozitív eredmény jelentkezését okozhatja. Egy minta elemzése a két allélnak megfelelően rendszerint két külön reakcióelegyben történik, ugyanakkor amennyiben az SNP körüli szekvencia ezt lehetővé teszi, a két allélspecifikus primer eltérő irányultsággal egy reakcióban is alkalmazható. Két reakcióelegyben történő genotipizálás esetén előnyös, ha a reakciók egy az allélspecifikussal azonos

irányú külső primert is tartalmaznak. Az így mindig keletkező kontroll fragmentum mutatja, hogy az allél-specifikus termék hiánya nem technikai hiba következménye.

A DNS-ligáz (a DNS-polimerázhoz hasonlóan) csak akkor tud egy hiányzó foszfodiészter kötést létrehozni (vagyis egy „nick”-et megszüntetni), ha a DNS-szálak tökéletesen illeszkednek egymáshoz. Az enzim az allélspecifikus próbát így a közvetlenül mellé hibridizáló DNS-molekulával csak akkor kapcsolja össze, ha megfelelő allél van jelen a vizsgált DNS-szálon („oligonucleotide ligation assay” – allélspecifikus oligonukleotid ligálás). Az így létrejött termék (gyakran univerzális) végeire hibridizáló primerekkel a termék felsokszorozható, amennyiben a két próba ligálása a vizsgált szekvenciában lévő polimorfizmus miatt nem történik meg, értelemszerűen nem képződik PCR-termék. Az eljárás kapilláris elektroforézis és eltérő hosszúságú próbák alkalmazásával hatékonyan multiplexálható: ez a technika az MLPA („multiplex ligation-dependent probe amplification” – multiplex ligálás-függő próba amplifikáció), mely az SNP-k és pontmutációk genotipizálása mellett kisebb és nagyobb kiterjedésű inzerciók, deléciók és kópia szám polimorfizmusok (CNV-k) vizsgálatára is alkalmazható [132-135]. Említésre méltó ugyanakkor, hogy az MLPA esetén előfordulhatnak álpozitív eredmények, így a módszer alkalmazása és az eredmények értékelése során kellő körültekintés szükséges [136].

DNS-láncok hibridizációján alapul a nagy áteresztő képességű „DNS-chip” technológia is. Számos új platform főképp az SNP-k, de akár a CNV-k egyidejű elemzését kínálja, egy lemezen 250 000–4 000 000 polimorfizmus vizsgálható [137], s ma már a technika érzékenysége is kitűnő, a genotipizáláshoz rendkívül kis mennyiségű minta-DNS elegendő [138]. A módszer lényege röviden, hogy a fragmentált, fluoreszcens festékkel jelölt, egyszálú DNS-molekulák bá-zissorrendjüknek megfelelően a chip felszínére rögzített próbákhoz kötődnek. Ezt követően a berendezés leolvassa az egyes pozíciókban mérhető fluoreszcens jelet: ennek alapján nem csak az adott szekvencia jelenléte vagy hiánya, hanem annak relatív mennyisége is megállapítható [139]. Bár a technika számos – elméleti és klinikai – területen hatékony megoldást jelenthet, mégis szembetűnő, hogy a módszerrel kapcsolatos tudományos közlemények száma 2013 óta alig több mint egyharmadára csökkent. Ennek oka vélhetően az, hogy az NGS eljárások mára a közepes, illetve viszonylag nagy áteresztőképességű genotipizáló technikák reális alternatívájává váltak: a módszerek költsége hasonló nagyságrendű, ugyanakkor az új generációs szekvenálás rendszerint lényegesen több szekvencia adatot generál.

A primerextenzió – kémiai szempontból – a Sanger-szekvenálással rokon, egyes források emiatt a technikát miniszekvenálásnak is nevezik. Ennél az eljárásnál – szemben az allélspecifikus amplifikációval – a primer 3' vége éppen a vizsgált lókuszt melletti nukleotidhoz hibridizál, és a reakcióelegyben jelen van mind a négy láncterminátor (leggyakrabban didezoxi- vagy aciklo-) nukleotid más-más színű fluoreszcens festékekkel megjelölve. A folyamat során a primer egyetlen nukleotiddal hosszabbodik meg, a DNS-szintézis természetes szubsztrátjait (dATP, dCTP, dGTP és dTTP) ugyanis az oldat nem tartalmazza. A létrejövő termék színe a genotípustól függ, mivel a minta-DNS-ben jelenlévő allél határozza meg, hogy a primer melyik jelölt láncterminátorral hosszabbodik meg. A képződő DNS-fragmentumok olyan – a Sanger-szekvenálás során is alkalmazott – kapilláris elektroforézissel tehetővé láthatóvá, ami képes a négy fluoreszcens festék egyidejű detektálására. Az eljárás hatékonyabb, multiplex formában is kivitelezhető. A genotípus leolvasása ugyanis a szín detektálásán alapul, a kapilláris elektroforézis viszont természetesen hatékonyan alkalmazható a termékek méret szerinti elválasztására is. Így különböző hosszú extenziós primerek segítségével párhuzamosan 6–8 SNP genotipizálása is megvalósítható [140, 141]. Kidolgozták a módszer olyan, módosított változatát, ahol az alapelv hasonló, de a detektálás elektrokémiai úton történik, így szükségtelenné téve a viszonylag drága és helyigényes CCD kamera alkalmazását [142].

A valós idejű (real-time) PCR alapelve, hogy a képződő termékeket nem csak a reakció után, hanem az amplifikáció minden egyes ciklusában detektálják. Ez két különböző technikai eljárás segítségével valósítható meg: alkalmazhatunk a DNS-hez aspecifikusan kötődő interkalátor festékeket (pl. Eva Green, SYBR Green, SYTO-család stb.), vagy a szekvencia specifikus fluoreszcens próbákat [143]. A folyamatos detektálás nem csak a munkafolyamat gyorsítását, így a hatékonyság növelését eredményezi, hanem lehetővé teszi a minta-DNS kiindulási mennyiségének pontos meghatározását is. Ha ugyanis a fluoreszcens jelet a ciklusszám függvényében ábrázoljuk, leolvasható, hogy a reakció meddig megy hatékonyan végbe: meddig tud a PCR-termékek mennyisége ciklusról ciklusra (jó közelítéssel) megkétszereződni. A kiválasztott küszöb fluoreszcencia eléréséhez szükséges ciklusszám (C_T vagy C_q) alapján pontos abszolút vagy relatív kvantifikálás végezhető el, és a módszer előnye hogy egy kísérletben a 2–3×-os és a 10^5 – 10^6 ×-os koncentrációkülönbségek egyidejűleg megbízhatóan kimutathatók.

A valós idejű PCR technikák az SNP-genotipizálás során eredményesen alkalmazhatók. Bár itt mennyiségi elemzés nem történik, de a technika hatékonysága, gyorsasága és a genotípus

meghatározását is magában foglaló automatizálhatóság nagy előnyt jelent. A módszer elve, hogy a reakcióelegy – a PCR szokásos összetevőin kívül – két próbát is tartalmaz. Ezek az egyik illetve a másik alléllal komplementerek, és különböző színű fluoreszcens festékekkel jelöltek. Többféle technika létezik: az egyik leggyakrabban használt eljárás a „molecular beacon” (molekuláris jelzőfény) nevű próbák alkalmazásán alapul, szintén széleskörűen elterjedtek a TaqMan próbák is [144]. Mindkét megközelítés lényege, hogy az intakt próbák nem bocsátanak ki fluoreszcens jelet, mert egy csillapító („quencher”) molekula a „riporter” fluoreszcens festék által kibocsátott fényt elnyeli. A „molecular beacon” esetén ezt a molekula hajtűszerű szerkezete eredményezi. A próba a megfelelő allél jelenléte esetén ugyanakkor a minta DNS-sel stabilabb duplaszálú struktúrát képez, így a molekula „kinyílik”, a csillapító hatása nem érvényesül, és a próba jelet bocsát ki [145]. A TaqMan-próbák esetén a csillapító molekula hatása úgy szűnik meg, hogy – a megfelelő allél jelenlétében – a minta-DNS-hez hibridizáló próbát a DNS-polimeráz a PCR során 5’ exonukleáz aktivitásával nukleotidokra lebontja, így a quencher és a „riporter” festék el tud diffundálni egymástól [146]. Az elemzés során a fluoreszcens jelek valós idejű mérése nem is feltétlenül szükséges: a PCR után történő „végpontdetektálás” alapján a két különböző színű fluoreszcens szignál aránya meghatározható, és így a genotípus leolvasható. A TaqMan OpenArray® rendszer a technika nagy áteresztőképességű változata. Egy mikroszkópos tárgylemezhez hasonló nagyságú acéllemezen, 33 nl térfogatú reakcióelegyekben történik az amplifikáció, egy lemezen 3072 párhuzamos PCR valósítható meg [147, 148]. Ez a miniatürizált eljárás SNP-k genotipizálása mellett mennyiségi elemzések kivitelezésére is alkalmas [149].

Az SNP-k genotipizálása, valamint azonosítása az olvadáspont mérésével is megvalósítható. Az elemzés szekvencia specifikus próbák helyett a DNS-interkalátor festékek alkalmazásán alapul. A módszer lényege az, hogy ezek a festékek csak a kettős láncú DNS-sel képeznek komplexet, egy szálú DNS jelenlétében nem adnak jelet. A DNS olvadáspontja – vagyis az a hőmérséklet (T_m), ahol a molekula két lánc elválik egymástól – ezt kihasználva meghatározható. A berendezés a vizsgált mintát felmelegíti, eközben folyamatosan méri a fluoreszcens jelet, ami az olvadáspontnál hirtelen lecsökken. Megfelelő érzékenységű detektor segítségével akár egyetlen nukleotidnyi szekvencia különbség, azaz egy SNP is kimutatható [150]. Az olvadáspont megváltozását természetesen a vizsgált molekulában jelen lévő bármilyen szekvencia

eltérés okozhatja, így az eljárás önmagában nem nyújt biztos információt egy adott SNP genotípusáról és még lokalizációjáról sem, ugyanakkor éppen emiatt alkalmas új, eddig még ismeretlen polimorfizmusok azonosítására [151].

A denaturáló gradiens gél elektroforézis (DGGE) és az egyszálú konformáció polimorfizmus (SSCP – single-strand conformation polymorphism) bizonyos szempontból hasonlít az olvadáspontra. Az alapelv ennél a két technikánál az, hogy az egyszálú DNS térszerkezetét, valamint (a duplaszálú DNS-ben, illetve az egyszálú molekula egyes komplementer szakaszai között) a hidrogén-hidak létrejöttét a szekvencia befolyásolja, így a vizsgált régióban jelen lévő bármely – korábban még nem azonosított – polimorfizmus feltárható [152]. A denaturáló gradiens gélelektroforézis során olyan gélt alkalmaznak, ami a molekulák vándorlási irányának megfelelően egyre növekvő koncentrációjú denaturáló ágenst (ureát vagy formamidot) tartalmaz. A különböző molekulák így (szekvenciájuktól, azaz lényegében a H-kötések számától függően) a gél más-más pontján denaturálódnak, vándorlási sebességük itt jelentősen csökken, így a kapott mintázat alapján következtetni lehet a minták szekvenciájának eltérésére. A PCR amplifikáció során az egyik primer 5' végén egy GC-gazdag régió van („GC-horgony”), ez megakadályozza, hogy a két DNS-lánc teljesen elváljon egymástól az elválasztás során [153]. Az eljárás módosított változata a hőmérséklet gradiens gél elektroforézis (TGGE – temperature gradient gelelectrophoresis), ahol a real-time PCR alapú technikához hasonlóan a vizsgált DNS-ek olvadáspontja közötti eltérés mutatható ki [154]. Az egyszálú konformáció polimorfizmus során az egyszálúvá tett DNS-t denaturáló körülmények között analizálják gélelektroforézis segítségével. Optimális esetben egyetlen nukleotidnyi különbség is megváltoztatja a nukleinsav lánc térszerkezetét, aminek eredményeként módosul a molekula vándorlási sebessége az elválasztás során [155].

2.3.1.3 CNV-k genotipizálása

Bár a CNV-k ismétlődési variációknak is tekinthetők, meghatározásuk mégsem könnyű feladat. Ez éppen méretüknek köszönhető: a variábilis DNS-szakasz hossza rendszerint lényegesen nagyobb, mint amekkora régió PCR segítségével amplifikálható, így vizsgálatuk PCR-rel és azt követő gélelektroforézissel nem valósítható meg. Elemzésük ezzel szemben kvantitatív analízisként értelmezhető: azt határozzák meg, hogy egy adott, az ismétlődő egységben jelenlévő

(rövid) szekvencia hány példányban található meg a genomban. Ennek megfelelően kimutatásuk több – mennyiségi mérésre alkalmas – technikával megvalósítható. Egyik legelterjedtebb, hatékony technika az aCGH (array based comparative genome hybridization). Az eljárás lényegében a DNS-chip technológiával egyező módon működik. A lapka felszínére rögzítetten helyezkednek el a teljes genomot reprezentáló próbák (egyszálú DNS-molekulák). Ezzel hibridizáltatják az elemzett genomot megfelelően fragmentált formában egy ún. referencia genommal együtt. A két minta (vizsgált és referencia) DNS-molekulái eltérő színű fluoreszcens festékekkel jelöltek. A különböző színű molekulák kompetitív módon kötődnek a chipen lévő próbákhoz, eltérő génszám (1:1-nél nagyobb vagy kisebb arány) esetén a megfelelő pozíciók színe megváltozik, ami alapján az ismétlődési szám leolvasható [156].

Célzott, egy vagy néhány régióra irányuló elemzés esetén a real-time PCR alapú technikák igen pontos és megbízható eredményt szolgáltatnak. A vizsgálat mind DNS-interkalátor, mind specifikus próbák használatával megvalósítható. A relatív kvantifikálás során belső kontrollként olyan genomi régió alkalmazása szükséges, melyet egyetlen CNV sem foglal magában, így az ismétlődési szám referenciaként 2-nek tekinthető [157, 158].

2.3.1.4 DNS-szekvenálás

A DNS bázissorrendjének elemzése, a genetikai mutációk és polimorfizmusok azonosítása természetesen DNS-szekvenálási technikákkal is megvalósítható. A ma már hagyományosnak mondható Sanger-szekvenálás lényege a DNS-szintézis természetes szubsztrátjai (dATP, dCTP, dGTP és dTTP) mellett a négy különböző színű fluoreszcens festékekkel jelölt láncterminátor (didezoxi-, aciklo-) nukleotidok alkalmazása. Amikor ezek a molekulák beépülnek a képződő láncba, szabad 3' OH-csoport híján a DNS-szintézis befejeződik. A keletkezett termék mérete és színe kapilláris gélelektroforézissel automatizált módon meghatározható, ezen szignálok „összeolvasásával” a szekvencia meghatározható. Ez a technika többek között kiemelkedő pontossága miatt a genotipizáló módszerek validálása, valamint számos egyéb molekuláris biológiai munka (pl. rekombináns DNS-ek szekvenciájának ellenőrzése) során még napjainkban is nélkülözhetetlen. Célzott genotipizálásra ugyanakkor kevésbé hatékony, mivel át-eresztőképessége kiterjedtebb régiók vizsgálatához nem elég nagy, rövid (< 1000 bp) szakaszo-

kon belül elhelyezkedő egy vagy néhány polimorfizmus és mutáció meghatározása pedig a bemutatott technikák valamelyikével rendszerint egyszerűbben, olcsóbban és gyorsabban megvalósítható.

Az új generációs szekvenálási (NGS) eljárások az elmúlt másfél évtizedben ugyanakkor robbanásszerű fejlődésen mentek keresztül, kapacitásuk 2–3 nagyságrenddel emelkedett, költségük pedig szignifikánsan csökkent. Ezekkel a technikákkal így ma már óriási mennyiségű szekvencia adat nyerhető, melynek elemzése ugyanakkor kihívást jelent amiatt is, mert a legtöbb módszer egy lépésben lényegesen rövidebb szekvencia (35–700 bp) leolvasására képes, a hibásan meghatározott nukleotidok aránya (0,1–15%) pedig jelentősen magasabb, mint a Sanger-szekvenálás esetén. Az NGS technikák kiindulópontja az adott platformhoz illeszkedő módon előkészített DNS-könyvtárak létrehozása. Első lépés a minta-DNS fragmentálása, ezt követi a rövid DNS-ek (gyöngyök felszínén, víz–olaj emulzióban lévő vízcseppekben történő) felsokszorozása. Így válik lehetővé, hogy a szekvenálás több millió reakcióterben egyidejűleg „masszívan párhuzamos” módon valósul meg, ami nélkül a 10^8 – 10^{10} bp nagyságrendbe eső adat meghatározása nem volna lehetséges. Másik lényeges különbség a hagyományos, Sanger-szekvenáláshoz képest, hogy a detektálás nem elektroforézissel, hanem a szintézis során történik, ennek pontos módja az egyes eljárások esetén különböző [159].

A szekvencia leolvasása alapulhat jelölt próbák hibridizálása és ligálása során keletkező fluoreszcens szignálok mérésével (SOLiD – „sequencing by oligo ligation detection). A próbák 3' végén adott dinukleotid áll, amit további degenerált vagy univerzális bázisok követnek. A 16 különböző próbának megfelelően 16 jól elkülöníthető fluoreszcens festék nem áll rendelkezésre, így a rendszer csupán 4 eltérő jelet detektál, azaz 4–4 dinukleotid azonos színnel jelölt. A megfelelő próba bekötődését követően megtörténik a szignál leolvasása, majd a próba 5' végét a fluoreszcens festékekkel együtt lehasítják, és így – néhány nukleotid kihagyását követően – megtörténhet a következő dinukleotid analízise. Rendszerint tíz hibridizálás / ligálás / detektálás / hasítás ciklus után a szekvencia analízise újratekintődik olyan illesztő oligonukleotid alkalmazásával, mellyel a leolvasási pozíció egy bázispárral eltolódik. Több ilyen lépés eredményeinek együttes feldolgozásával mind az azonos színnel jelölt dinukleotidok, mind az eddig kimaradt szakaszok szekvenciája meghatározható [160]. Az egyes nukleotidokat így több próba

is lefedi, ennek következtében ez a technika meglehetősen nagy pontosságú, az egyben leolvasott szekvencia hossza ugyanakkor meglehetősen alacsony (kevesebb mint 100 bp), emiatt a hosszabb kiterjedésű ismétlődő szakaszok azonosítása nehézségekbe ütközhet.

Másik lehetőség a DNS-szintézis során keletkező fluoreszcens jelek detektálása a reakció során. Ennek egyik régóta ismert módja a piroszekvenálás, mely egyedi reakcióként is kivitelezhető rövid szakaszok bázissorrendjének elemzésére, több mint 1,5 millió, 10^{-12} liter nagyságrendű párhuzamos reakcióteret tartalmazó változata pedig az NGS egyik hatékony technikája. A piroszekvenálás lényege, hogy egy lépésben mindig csak egy nukleotidot (dATP-t, dCTP-t, dGTP-t vagy dTTP-t) adnak a reakcióhoz, melyet a polimeráz beépít a képződő láncba, amennyiben az komplementer a vizsgált DNS-sel. A nukleotid beépülését egy kapcsolt, többlépéses enzimreakció során fény keletkezése kíséri. Ennek lényege, hogy a DNS-polimeráz által katalizált reakcióban a beépülő dezoxinukleotidból egy pirofoszfát lehasad. Ezt a csoportot az ATP-szulfuriláz adenzin-5-foszfoszulfátra helyezi, így ATP jön létre. A luciferáz enzim a luciferint ATP jelenlétében oxiluciferinné alakítja, miközben fény keletkezik, ezt a 36 cm²-es lemez felszínén CCD kamera detektálja [161]. A feleslegben maradó nukleotidokat (a szubsztrátot és az ATP-t) az ATP-difoszfátáz enzim lebontja. A rendszer előnye, hogy kivitelezéséhez nem szükséges módosított (jelölt, illetve láncterminátor) nukleotidok alkalmazása, mivel a lánckukleotidoként történő meghosszabbítását biztosítja, hogy minden lépésben csupán egyféle nukleotid van jelen. Épp emiatt a nukleotidokat egyesével kell a reakcióelegyhez adni, és még ennek ellenére is megtörténhet, hogy a lánck több nukleotiddal hosszabbodik meg olyan szakaszokon, ahol azonos bázisok állnak egymás mellett. Ilyenkor a fény intenzitása arányos a beépült szubsztrátok számával, megjegyzendő ugyanakkor, hogy 7–8 nukleotidból álló, vagy ennél hosszabb ismétlődés esetén a lineáris összefüggés már nem teljesen megbízható, így ilyen szakaszok szekvenálása során a módszer pontatlan lehet.

Hasonló elven működik az Ion Torrent platform, a detektálás azonban eltérő. A kapcsolt, fényfelvillanást eredményező enzimreakció helyett a módszer a dezoxiribonukleotidok beépülése során felszabaduló H⁺-ionok okozta pH-változás mérésén alapul [162]. Az eljárás az adott alkalmazáshoz illeszthető, különböző áteresztőképességű változatban elérhető, és gyorsabb a legtöbb alternatív NGS módszernél, emiatt a technikát napjainkban számos klinikai kutatás során is alkalmazzák [163-166].

A DNS-szintézis valós idejű detektálásán alapuló eljárások másik válfaja némileg hasonlít a hagyományos Sanger-eljáráshoz. A DNS-polimeráz kiindulópontjául szolgáló primer egy adapter szekvenciához kötődik, így nem szükséges a bázissorrend részleges ismerete a folyamat elindításához. A folyamat lényege olyan nukleotidok alkalmazása, melyek fluoreszcens festékekkel jelöltek, és a 3' C-atomon lévő módosítás következtében a további lánchosszabbítást megakadályozzák. Hosszabb régió bázissorrendjének leolvasása úgy lehetséges, ha a 3' OH-t elfedő csoport csakúgy, mint a fluoreszcens festék a detektálást követően lehasítható, és így a lánccsintézise folytatható. Ilyen kémiai módosításra alkalmas az azidometil [167] vagy az allil csoport, ez utóbbi a fluoreszcens festék kötésére is használható, így egyetlen közös lépésben lehetővé válik a 3' OH csoport regenerálása és a detektálást követően a jelölő molekula eltávolítása [168]. A folyamat ciklusosan megismételhető, nagy előnye, hogy a 4 jelölt nukleotid minden egyes lépés során egyszerre adható a meghosszabbodó lánchoz, és minden esetben pontosan egy nukleotid épül csak be a vizsgált DNS-szál szekvenciájának megfelelően, mivel a 3' OH-t fedő csoport eltávolításáig a további extenzió nem lehetséges [168], ennek megfelelően a technika megbízhatósága nem csökken azonos nukleotidok hosszabb ismétlődését tartalmazó szekvenciák esetén sem.

Ezen technikák közös jellemzője, hogy az egy lépésben leolvasott DNS hossza legfeljebb néhány száz bp lehet. Ismert ugyanakkor, hogy az emberi genomban számos hosszú ismétlődési variáció (CNV) található, ezek pontos azonosítása ilyen rövid szekvenciák összeillesztése alapján nem végezhető el minden esetben megbízhatóan. Kidolgoztak ezért olyan módszereket, melyek alkalmazásával $\sim 10^4$ bp hosszúságú régiók szekvenciája egyben leolvasható. Az egyik eljárás (SMRT – single molecule real-time szekvenálás) lényege immobilizált DNS-polimeráz alkalmazása. Ennek eredményeként a DNS-lánc meghosszabbítása – molekuláris szinten – adott helyen történik. A reakcióelegy négy különböző színnel jelölt dezoxiribonukleotidot tartalmaz, a beépülés folyamata – melynek során a megfelelő nukleotid az enzimhez kötődik – egy rendkívüli érzékenységgű szenzorral detektálható [169]. A lánccsintézise során a fluoreszcens molekula is lehasad a szubsztrátról, és eldiffundál a kamera látóteréből. A módszer lehetővé tette a humán genomban lévő számos – különböző kiterjedésű – ismétlődési variáció pontos meghatározását [170]. Hosszabb DNS-fragmentumok bázissorrendje egy lényegében egyszerű köztes lépés alkalmazásával a rövid szakaszok leolvasására alkalmas NGS technikák

segítségével is egyértelműen és pontosan meghatározható. Ezen eljárások lényege, hogy a genomi DNS-t első lépésben 10^4 bp nagyságú szakaszokra darabolják fel, és ezeket – pl. egy mikrotiter tálca egyes pozícióba – elkülönítik. A molekulákat tovább darabolják, és egy mesterséges DNS-címkével („vonalkód”) látják el. Ezt követően az így előkészített minta-DNS-t tartalmazó oldatok egyesíthetők, és valamely NGS módszerrel szekvenálhatók, mivel a „vonalkód” alapján egyértelműen ismert, hogy mely fragmentumok építenek fel egy adott hosszú DNS-szakaszt [171].

2.3.2 A genotípus–fenotípus összefüggés kutatása

Az újgenerációs szekvenálási eljárások (NGS) robbanásszerű fejlődésének eredményeként a polimorfizmusok elemzése nem jelent limitáló tényezőt az örökletes jellegek genetikai hátterének kutatása során, ennek köszönhetően tért hódítottak a genom szintű asszociáció vizsgálatok. Ennek ellenére megfelelően megválasztott körülmények között a kandidáns gén vizsgálatok napjainkban is helyet kaphatnak ezen a területen, amennyiben az elemzés valamilyen elméleti vagy korábbi kutatásból származó hipotézisen alapul, és ritka variánsok elemzését célozza. Leegyszerűsítve a genom szintű asszociáció elemzések (GWAS) hipotézis alkotó, a kandidáns gén vizsgálatok pedig hipotézis elemző eljárásnak tekinthetők (3. ábra) [172].

2.3.2.1 Genom szintű asszociáció vizsgálat (GWAS)

A komplex jellegek és betegségek genetikai hátterét képező polimorfizmusokra vonatkozó ismeretünk zöme a genom szintű asszociációvizsgálatokból (GWAS) származik. A GWAS-ok előzetes hipotézis nélkül a teljes genomban keresik azokat a polimorfizmusokat, melyek valamilyen jelleggel asszociációt mutatnak. Ez kétségtelen előnyt jelent a célzott, kandidáns gén vizsgálatokkal szemben, a teljes genomra vonatkozó hatalmas mennyiségű – és esetenként részben nem releváns – adat feldolgozása során ugyanakkor számos elméleti és technikai nehézséggel is számolni kell.

A legtöbb GWAS SNP-k elemzésére fókuszál, esetenként CNV-eket is vizsgálnak, ugyanakkor említésre méltó, hogy VNTR-ek genotipizálására szinte soha nem kerül sor annak ellenére, hogy az ismétlődési variációk jelentős mértékben hozzájárulnak az emberek közötti ge-

netikai különbözőségekhez, és az SNP-khez hasonlóan örökletes jellegek és betegségek genetikai rizikófaktorai lehetnek. A GWAS-ok során gyakran SNP blokkokat azonosítanak, melyek statisztikailag korrelálnak az adott jelleggel, ezeket „rizikó lókusznak” nevezzük. Egyes esetekben ezekben a blokkokban a betegség kialakulásáért felelős gén azonosítható: az IL-12 / IL-23 út Crohn-betegségben gyógyszer célponttá vált [173]. A GWAS eredmények alapján ma már egyes betegségek (pl. koszorúér-betegség, T2DM, emlőrák) esetén meghatározható egy poligénes rizikó pontszám (PRS: polygenic risk score), ami alapján a kórkép kialakulásának esélye megjósolható. Máig több mint 5700 GWAS-t végeztek el több mint 3300 jelleggel kapcsolódóan vizsgálatonként akár több mint egymillió minta analízisével [174, 175], az NCBI PubMed adatbázisában 2010 óta évente 150–200 közlemény címében szerepel a „GWAS” kifejezés. Az eredmények azt mutatták, hogy a legtöbb jelleg SNP-k ezreivel asszociál, melyek jelentős része számos más jelleggel is összefüggést mutat, és a polimorfizmusok közötti LD miatt ezek jelentős része nem rendelkezik tényleges biológiai hatással, így a funkcionális SNP-k kiválasztása és azonosítása nagy kihívás.

Egy genom szintű asszociáció elemzés alapja a megfelelő mintapopuláció kiválasztása: reprodukálható, szignifikáns asszociáció kimutatása igen nagy mintacsoport bevonását teszi szükségessé, melynek ideális létszáma a Genetic Power Calculator segítségével becsülhető meg [176]. Ez már önmagában rávilágít arra, hogy egy GWAS idő- és költségigénye csaknem bizonyosan meghaladja a legtöbb önálló kutatólaboratórium lehetőségeit. Ugyanakkor napjainkban már elérhetőek olyan nyilvános források, melyek nagy populációk részletes fenotípus és genotípus adatait tartalmazzák. Érdeemes azonban szem előtt tartani, hogy a vizsgálatba bevont résztvevők megfelelően gondos kiválasztása még ilyen adatforrások esetén is alapvető: sok kutatás veszi például alapul a UK Biobankból elérhető mintákat, ahova a résztvevőket önkéntes alapon toborozzák, akik ennek megfelelően egészségesebbek, jómódúbbak és iskolázottabbak az átlagpopulációnál [177]. Hasonló módon a kórházakban gyűjtött minták szintén – értelemszerűen – eltérnek a globális átlagtól. Jól ismert, hogy a polimorfizmusok allélfrekvencia értéke igen nagy földrajzi eltérést mutat, így bár eltérő etnikai eredetű személyek bevonhatók egy elemzésbe, de lényeges, hogy a vizsgált populációk összetétele azonos legyen.

A GWAS másik alappillére a megbízható, robosztus genotípusmeghatározás jelenti. A genotípus adatok napjainkban már új generációs (NGS) genom (WGS) vagy exom szekvenálásból (WES) is származhatnak. Míg a súlyos, monogénesen öröklődő betegségekért felelős

genetikai variációk több mint 90%-ának azonosításához az exom szekvenálásból származó adatok is elegendők, a komplex jellegek genetikai hátteréhez hozzájáruló gyakori variánsok sok esetben az exonokon kívül helyezkednek el. Noha a genom szekvenálás értelemszerűen költségesebb, nem csak több polimorfizmus azonosítására alkalmas, de SNP-k esetén lényegesen alacsonyabb az álpozitív találatok aránya, és csupán ez a megközelítés képes a CNV-k kimutatására [178]. A szekvenálás mellett a microarray alapú genotipizáló eljárások szintén a genotípusmeghatározás eszközeként szolgálhatnak. Ezzel az eljárással napjainkban még a legalacsonyabb költséggel a legtöbb viszonylag magas ritka allél frekvenciájú SNP nagy pontossággal meghatározható. Feltételezhető ugyanakkor, hogy a WGS költségének további csökkenése eredményeként a közeljövőben a genom szekvenálás lesz a GWAS során elsődlegesen alkalmazott genotipizálási technika.

Bár a genotipizálás során kapott rendkívüli mennyiségű adat felbecsülhetetlen információt hordoz, ennek elemzése komoly kihívást jelent. A tényleges asszociációvizsgálat gondos előzetes elemzéseket követően végezhető csak el. Ezek közé tartozik a nem polimorf lókuszok, és a Hardy–Weinberg-egyensúlytól szignifikánsan eltérő, valamint a sikertelenül genotipizált SNP-k eltávolítása. A megfelelő adatfeldolgozás és -elemzés jó indikátora a résztvevők által megadott nem és az X, illetve Y kromoszómák jelenléte alapján megállapított nem egyezése. Mindezen minőségellenőrző lépések kivitelezésének gyakori eszköze a PLINK alkalmazás [179]. Lényeges a genotípus eredmények validálása különösen abban az esetben, ha az elemzés során összehasonlításra kerülő csoportok genotipizálása nem párhuzamosan, esetleg eltérő eljárással történt. További igen fontos szempont – nem csupán a genom szintű, hanem a célzott, kandidáns gén vizsgálatok esetén is – a vizsgált személyek közötti rokoni kapcsolatok, illetve a származás, etnikai hovatartozás figyelembe vétele [180]. A populáció rétegződés (stratifikáció), azaz a vizsgálati csoportokon belül elkülöníthető eltérő származású, fenotípusú és különböző létszámú alpopulációk jelenléte álpozitív eredmények, valótlan genotípus–fenotípus összefüggések kimutatásához vezethet [181].

A genotípus–fenotípus kapcsolat kimutatása legtöbbször eset–kontroll elemzés keretében valósul meg. Ebben az esetben a jelleg rendszerint diszkrét változó: a két populációba a vizsgálati személyek az adott tulajdonság jelenléte vagy hiánya alapján kerülnek besorolásra. Fontos szempont a kontroll személyek megfelelő beválogatása. Valamilyen betegség genetikai hátterének elemzése során egy „átlagpopulációban” értelemszerűen a betegség prevalenciájának

megfelelő számban beteg személyek is jelen vannak. A nem, kor, eredet szerint illesztett, illetve az adott jelleg hiánya alapján „aktívan” beválogatott kontroll csoport alkalmazása ennek megfelelően a GWAS statisztikai erejét nagyban növelheti. A genotípus–fenotípus összefüggés meghatározása alapvetően lineáris vagy logisztikus regresszió alkalmazásával történik attól függően, hogy a vizsgált fenotípus folytonos vagy diszkrét változó.

Az elsőfokú rokonok, családi triók alkalmazásán alapuló elemzések napjainkban is reális alternatívát jelentenek [182]. Ez a megközelítés nagyobb mintaszámot igényel, és a megfelelő családok beválasztása sokszor nehezebb, mint független személyek toborzása. Ugyanakkor a közvetlen rokonok (érintett gyermekek és szülei) közötti allélátadódás gyakoriságának elemzése – „transmission disequilibrium test” – a genotípus–fenotípus összefüggés hatékony jelzője lehet [183], és nagy előnyt jelent, hogy mivel nem külön populációkat hasonlítunk össze, populációrétegződésből adódó álpozitív eredmények nem jelentkezhetnek. Mind az eset–kontroll, mind a családvizsgálat esetén elengedhetetlen a többszörös tesztelésből származó, véletlenszerűen felbukkanó eredmények kiszűrése. Ennek gyakori módja a Bonferroni-korrekciónak, melynek lényege, hogy a szokásosan alkalmazott szignifikancia szintet ($p < 0,05$) az elvégzett független tesztek számával osztjuk el. Ennek következtében az új p -érték rendszerint 4–6 nagyságrenddel alacsonyabb tartományba kerül, ami esetenként a kis hatású polimorfizmusok, illetve gének azonosítását megnehezíti. Erre a problémára megoldást jelenthet a Benjamini–Hochberg-féle korrekció („false discovery rate”) alkalmazása, mellyel a többszörös tesztelés miatti korrekció a Bonferroni-módszernél megengedőbb módon valósítható meg [184].

A GWAS végeredménye egy a statisztikai szignifikanciát jelentő p -értékeket, hatásnagyságokat, illetve az eltérések irányát tartalmazó lista, ami on-line eszközök – FUMA [185], Locus Zoom [186] – segítségével is megjeleníthető. A további lépések közül kiemelhető a tényleges biológiai funkcióval rendelkező variánsok azonosítása, mely *in silico* elemzések (pl. valamilyen biológiai rendszerhez, útvonalhoz tartozó gének és polimorfizmusainak meghatározása), valamint *in vitro* molekuláris biológiai elemzések segítségével valósítható meg.

2.3.2.2 Kandidáns gén vizsgálat

Bár a GWAS-okból származó információ jelentősége kétségtelen, a kandidáns gén vizsgálatok helye és szerepe ezzel együtt nem elhanyagolható a komplex jellegek és betegségek genetikai

hátterének megismerése során. Ugyan az NGS technikák egyre szélesebb körű alkalmazása lehetőséget teremt a genom összes polimorfizmusának tanulmányozására, mégis – részben a statisztikai elemzés lehetőségeiből adódóan – a GWAS-ok főképp a gyakori (MAF $\geq 0,05$) variánsokra fókuszálnak elsősorban. Noha ezek – az egészséges személyekben is előforduló – polimorfizmusok a *gyakori betegség–gyakori variáns* hipotézis szerint kicsi, mégis szignifikáns hatású genetikai rizikófaktorok lehetnek [187], a komplex jellegek genetikai hátterének kialakításában a ritka polimorfizmusok szintén jelentős szerepet játszanak [188]. A genom szintű elemzések esetén további korlátozó tényező, hogy ezek a kutatások leggyakrabban figyelmen kívül hagyják az ismétlődési polimorfizmusokat (VNTR-eket), melyek száma az SNP-kénél bár lényegesen alacsonyabb, de kiterjedésük és nagyfokú variabilitásuk miatt mégis jelentősen hozzájárulnak az emberek közötti genetikai különbséghez.

A kandidáns gén vizsgálatok sikerének és megbízhatóságának alapja a kutatási protokoll gondos megtervezése, valamint a megfelelő módszerek kiválasztása és alkalmazása. Döntő fontosságú a fenotípus lehető legpontosabb definiálása. Nehézséget jelent, hogy sok esetben a napjainkban alkalmazott klinikai diagnosztikus kategóriák meglehetősen általánosak, illetve gyakran több olyan állapotot foglalnak magukban, melyek különböző molekuláris mechanizmus révén alakulhatnak ki. Az ilyen heterogén kategóriák kialakításában a számos genetikai faktor mellett a környezeti tényezők is különböző súllyal játszanak szerepet. Sikeres megközelítést jelenthet azon személyek vizsgálata, akik az adott betegség súlyos vagy korán manifesztálódó formájában szenvednek, mivel ezekben az esetekben a genetikai tényezők kifejezettebb hatására lehet számítani. Gyakori eljárás a szub-, illetve endofenotípusok alkalmazása. Ezek olyan szűkebb, jól definiálható és (sok esetben kvantitatív módon) mérhető jellegek, melyek genetikai, biológiai szinten eredményesebben értelmezhetők, illetve magyarázhatók.

A kandidáns gén elemzések során – értelemszerűen – alapvető a vizsgálni kívánt gének és polimorfizmusaik kiválasztása. Az adott kórkép hátterében álló élettani, illetve molekuláris szintű folyamatok logikus kiindulópontot jelentenek, számos adatbázis és on-line eszköz segíti a releváns targetek objektív, megbízható azonosítását. Ennek speciális válfaja olyan gének elemzése, melyek a vizsgált fenotípushoz köthető súlyos, monogénes betegségek kialakulásáért felelősek. Logikus kiindulópontot jelenthetnek ezen túlmenően korábbi genom-szintű elemzések eredményei, állatmodellek megfigyeléséből vagy funkcionális elemzésekből (pl. átfogó, genom szintű génexpresszió vizsgálatok) származó adatok. A kiválasztott géneken belül az

elemzésbe bevonni kívánt lókuszek azonosítása sem egyszerű feladat, mivel számuk átlagosan a 10^3 – 10^4 -es nagyságrendben mozog, s a választás az SNP-ken túl a különböző nagyságú ismétlődési variációkra is kiterjedhet. Jó támpontot jelent a haplotípus, valamint az LD vizsgálata, ezek ismeretében rendszerint meghatározhatók olyan ún. „tag-SNP-k”, melyek segítségével nagyobb kiterjedésű haplotípusok megbízhatóan azonosíthatók. Ugyanakkor éppen emiatt – a GWAS-okhoz hasonlóan itt is – szem előtt kell tartani, hogy az esetlegesen statisztikai asszociációt mutató polimorfizmus jó eséllyel genetikai marker csupán, mely a vizsgált kórképpel kapcsolatos tényleges biológiai funkcióval nem rendelkezik.

A genotípus–fenotípus összefüggés tanulmányozása – csakúgy, mint GWAS-ok esetében – eset–kontroll vizsgálat mellett családi triók elemzésével is történhet. Mindkét esetben lényeges a kutatásban résztvevő populáció optimális létszámának előzetes becslése. A genotipizálás pontosságának ellenőrzésére hatékony eljárás az alkalmazott technika validálása független módszerrel a vizsgálati populáció egy részében, illetve a mért és a Hardy–Weinberg egyensúly alapján számított várható genotípus gyakoriságok összevetése. A genotípus–fenotípus asszociáció meghatározásának egyik alap statisztikai eszköze a χ^2 -próba [172], alkalmazható eljárás a Cochran–Armitage-teszt, valamint a logisztikus regresszió alapuló Wald-féle módszer is [189]. A célzott vizsgálatok során természetesen éppúgy alapvető a többszörös tesztelés miatt várható álpozitív eredmények elkerülése, mely a Bonferroni- vagy a Benjamini–Hochberg-féle korrekcióval lehetséges.

2.3.3 A genetikai variációk molekuláris hatásának elemzése

A genetikai asszociáció elemzések során – különösen GWAS esetében – a vizsgált fenotípussal statisztikailag szignifikáns összefüggést mutató polimorfizmusok gyakran markerek csupán, melyek nincsenek biológiai kapcsolatban az adott jelleggel, betegséggel. Ennek ellenére rávilágíthatnak a molekuláris mechanizmusban szerepet játszó kromoszómaszakaszra, illetve génre, emellett általában alkalmasak lehetnek a klinikumban biomarkerként. Mégis a betegség hátterében álló molekuláris szintű folyamatok feltárása, azaz a biológiai funkcióval rendelkező gének és polimorfizmusok felkutatása mind elméleti szempontból, mind esetleges új gyógyszer-célpontok azonosítása céljából lényeges. Ennek megfelelően az asszociáció vizsgálatokat

rendszerint funkcionális analízis követi, melynek során *in vivo*, *in vitro* vagy akár *in silico* eszközökkel elemezhető az eltérő szekvenciák molekuláris működése (3. ábra). Bár a kódoló szakaszokban elhelyezkedő mutációk és polimorfizmusok biológiai hatása sem jósolható meg egyértelműen a fehérje primer szerkezetének megváltozása alapján, a nem kódoló régiók variációinak funkciója még összetettebb kérdés, mivel ezek szabályozó hatása igen bonyolult és szerteágazó módon a transzkripció és a transláció szintjén is megvalósulhat.

2.3.3.1 A DNS–fehérje kölcsönhatás vizsgálata

A transzkripció szabályozásában a DNS-hez kötődő, valamint a DNS RNS-re történő átíródásában résztvevő fehérjék alapvető szerepet játszanak. Döntő a transzkripciós faktorok működése, de emellett a nukleoszóma kialakulásában szereplő fehérjék, a kromatinstruktúra szabályozását végző enzimek és a transzkripciós faktor komplex létrejöttében résztvevő fehérjék (pl. kofaktorok) is hozzájárulnak a génátíródás szabályozásához.

A DNS-hez *in vivo* kötődő fehérjék azonosításának egyik megbízható eszköze a kromatin immunprecipitáció. Bár az eljárás igen sokoldalúan használható, mégis fontos az adott kérdés megválaszolására leginkább megfelelő módozatának kiválasztása. A módszer elve, hogy az élő sejtben lévő DNS–fehérje komplexeket stabilizálják, majd a DNS fragmentálását követően specifikus antitestekkel elválasztják, és valamilyen módszerrel azonosítják azokat a DNS-szakaszokat, melyek kötődtek a célfehérjéhez.

A kromatin immun-precipitáció két fő típusa az N-ChIP és az X-ChIP. Az N-ChIP lényege, hogy a kromatin natív formában van jelen, ami lehetővé teszi a DNS nukleázzal történő fragmentálását, és az immunprecipitáció során specifikusabb és hatékonyabb antitestkötődés figyelhető meg. A DNS kinyerése annyival hatékonyabb, hogy az azonosítás során PCR-rel történő amplifikálás nem szükséges. Ezzel az eljárással ugyanakkor – a DNS–fehérje komplexek mesterséges stabilizálása híján – elsősorban a DNS-hez igen erősen kötődő fehérjék, leginkább csupán a hisztonok elemezhetők megbízhatóan. Az X-ChIP első lépése a DNS és a fehérjék közötti mesterséges keresztkötések kialakítása. Ennek előnye, hogy az analízis során az *in vivo* gyengébben kötődő fehérjék is stabil komplexet alkotnak, így bármilyen fehérjék (pl. transzkripciós faktorok) hatékonyan tanulmányozhatók. A DNS enzimatikusan ugyanakkor nem emészthető, sőt a mesterséges keresztkötések esetenként a szonikálással történő fragmentálás hatékonyságát is rontják, és a keresztkötő reagens, valamint maga a szonikálás a fehérjék antitestkötését is

befolyásolhatja. Az X-ChIP ugyanakkor igen érzékeny, így kevesebb sejtre illetve antitestre van szükség a vizsgálat során [190].

A DNS–fehérje komplexek stabilizálása, a mesterséges keresztkötések létrehozása leggyakrabban formaldehiddel történik. Ennek előnye, hogy a kötődés reverzibilis, a formaldehid a médiumból a membránon keresztül bejut a sejtekbe, és nem csak a DNS–fehérje, de az RNS–fehérje és a fehérje–fehérje komplexeket is stabilizálja Schiff-bázist kialakítva a fehérjében lévő lizinek, illetve a nukleinsavakban található adeninek, guaninok és citozinok között.

A DNS feldarabolását követően az eljárás lényegi része, az immunprecipitáció, előzetes hipotézist igényel. Specifikus antitest segítségével történik ugyanis azon DNS-molekulák elkülönítése, melyek ténylegesen kötődnek az adott fehérjéhez [191]. Ismert ugyanakkor a technika antitest nélkül működő változata, ahol a DNS-kötésben vélhetően szerepet játszó fehérjékre egy mesterséges címkét fuzionálnak a „Halo Tag” vektor alkalmazásával, az immunprecipitáció pedig ezen peptidszakasz segítségével történik [192]. A kigyűjtött DNS-eket ezután a fehérjétől elválasztják, majd valamilyen eljárással azonosítják. Ez történhet egyszerű PCR-rel vagy kvantitatív PCR-rel, ami azonban a kapott szekvenciák előzetes ismeretét tételezi fel. A DNS-molekulák klónozást követő szekvenálása vagy hibridizációs chipen történő elemzése nagy áteresztőképességű, hipotézis mentes megközelítést jelent [190].

A DNS-adenin-metiltranszferáz (Dam) alkalmazásán alapuló technika (DAMID) a kromatin immunprecipitáció hatékony alternatívája lehet. A módszer lényege, hogy a vizsgálandó fehérjét a bakteriális Dam enzimmel fuzionáltatják, és a fúziós fehérjét eukarióta sejtben kifejeztetik. Azokban a régiókban, ahova a fehérje a DNS-hez *in vivo* kötődik, a Dam alegység a baktériumokra jellemző metilációs mintázatot hoz létre a GATC-szekvenciákban lévő adeninek metilálása révén. Ezekben a helyeken a DNS a *Dpn* I enzim segítségével elhasítható, és a kapott fragmentumok a PCR-rel, real-time PCR-rel vagy hibridizációs chip alkalmazásával azonosíthatók. A módszer előnye a kromatin immunprecipitációval szemben, hogy nincs szükség a DNS–fehérje komplexek keresztkötésére és antitestek alkalmazására, ugyanakkor nehézséget jelenthet a fúziós fehérje létrehozása, illetve az, hogy a Dam a célfehérje funkcióját némileg befolyásolhatja [193].

Az elektroforetikus mobilitás megváltozásán alapuló eljárás (EMSA – „electrophoretic mobility shift assay”) az *in vitro* DNS–fehérje kötődés kimutatásának hatékony és viszonylag

egyszerű eszköze. A technika alapelve, hogy a DNS–fehérje komplex elektroforetikus mobilitása lényegesen eltér a szabad DNS vándorlási sebességétől. A módszer előnye, hogy kivitelezése viszonylag egyszerű, nem igényel különleges eszközháttérrel, és flexibilis mind a DNS, mind a fehérje tekintetében. A vizsgált DNS rövid oligonukleotid, hosszú, akár több ezer bázispár nagyságú fragmentum, sőt, cirkuláris molekula is lehet. Rövid nukleinsavak alkalmazásának előnye – amellet, hogy előállításuk olcsóbb és egyszerűbb –, hogy ezeken kisebb eséllyel fordulnak elő aspecifikus kötőhelyek, és a fehérje kötődése nagyobb mobilitásváltozást eredményez. A fehérje azonban – értelemszerűen – a DNS-lánc végéhez közel kötődik, ami a természetes kölcsönhatástól eltérő interakciók létrejöttéhez vezethet. Hasonlóan nincs megkötés az alkalmazott fehérjére és annak méretére vonatkozóan sem, megfelelő körülmények között tisztított fehérjék éppúgy használhatók a mérés során, mint nyers sejtkivonatok. Radioaktívan jelölt nukleinsav alkalmazásával kiemelkedő érzékenység érhető el, ugyanakkor a fluoreszcens és kemilumineszcens detektálási módok is hatékony kimutatást tesznek lehetővé, és fluoreszcens jelölés esetén a kölcsönhatás kvantitatív elemzése is megvalósítható. Jelöletlen, kompetitor nukleinsavak alkalmazhatók a kötődés erősségének jellemzése során, és hozzájárulhatnak az aspecifikus kötődések létrejöttének csökkentéséhez. Érdeemes ugyanakkor szem előtt tartani, hogy maga a gélelektroforézis a DNS és a fehérje disszociációját eredményezheti, de megeshet az is, hogy az elválasztás során a gélben a komplex stabilitása emelkedik.

Az elektroforézis során többféle puffer (Tris–acetát–EDTA, Tris–borát–EDTA, Tris–glicin) közül érdemes azt választani, melyben a DNS–fehérje kölcsönhatás legkevésbé romlik az elválasztás alatt. Ehhez különböző reagensek (glicerin, etilén-glikol) is hozzájárulhatnak. Az elválasztás poliakrilamid vagy agaróz gélben is történhet, a poliakrilamid rendszerint jobb felbontást eredményez, és kevésbé rontja a DNS–fehérje komplex stabilitását [194]. Kapilláris elektroforézis és lézeres detektálás alkalmazásával hatékony, gyors és érzékeny elemzés valósítható meg kapilláris zóna-elektroforézis technikával is.

A DNáz „footprinting” („lábnyom”) vizsgálat alapja, hogy a DNS-hez kötődő fehérjék (és egyéb molekulák) a kölcsönhatás helyén meggátolják a nukleinsav lánc mesterséges lebontását [195]. A DNS hasítását leggyakrabban a DNáz I enzimmel végzik el, ami a kettős hélix kis árkába kötődve hasítja a foszfodiészter kötéseket 5’ foszfát és 3’ hidroxil csoportokat létrehozva. A vizsgált DNS-szakasz rendszerint 50–200 bp nagyságú, és az elemzés során tisztított fehérje mellett sejtkivonat is alkalmazható [196]. Utóbbi esetben a vizsgálat során jelöletlen,

kompetitor oligonukleotidok használata szükséges, amivel csökkenthető az aspecifikus DNS–fehérje kölcsönhatásokból adódó álpozitív eredmények jelentkezése. A méréshez 5' vagy 3' végén radioaktív izotóppal vagy fluoreszcens festékkel jelölt oligonukleotid szükséges [197]. Ezt a nukleinsavat a vizsgált fehérje jelenlétében, illetve anélkül emésztik különböző koncentrációjú DNáz I-gyel eltérő ideig annak érdekében, hogy az enzim véletlenszerűen minden pozícióban elhasítsa a kontroll próbát. A fehérjét tartalmazó minta esetén a hiányzó hasítóhelyek alapján kirajzolódik a vizsgált fehérje kötőhelye, amit rendszerint denaturáló poliakrilamid gélelektroforézissel lehet megbízhatóan láthatóvá tenni. A technika a fehérjék mellett a DNS-hez kötődő más molekulák kimutatására szintén alkalmazható [198], PCR-rel kombinálva az *in vivo* DNS–fehérje kölcsönhatás kimutatásának eszköze lehet [199], és kifejlesztettek olyan vízoldékony arany nanorészecskékhez kötött DNS-ek alkalmazásán alapuló eljárást is, mellyel a hagyományos gélelektroforézishez hasonló felbontású, mennyiségi elemzés valósítható meg [200].

Lényegében azonos logikán, valamint a Förster-féle fluoreszcencia rezonancia energia átadás (FRET) elvén alapulnak azok a technikák, melyek elve, hogy alapállapotban a duplaszálú próba két láncán lévő fluorofórok (pl. FAM és TAMRA) kölcsönhatása létrejön, és így szignál nem mérhető. A két fluoreszcens festék közelségét valamilyen DNS-bontó enzim (*Exo* III exonukleáz [201], restriktív endonukleáz [202]) megszünteti, így a jel létrejön, ha azonban a DNS-hez fehérje kötődik, a láncok degradációja, és ennek következtében a szignál megemelkedése elmarad. A FRET jelenségét kihasználva két molekula közötti kölcsönhatás közvetlenül, a köztes DNS-hasító lépés elhagyása nélkül is kimutatható [203]. Hasonló elven alapul az Alpha Screen (Amplified Luminescent Proximity Homogeneous Assay) eljárás, melynek során lézerrel történő gerjesztés hatására a donor gyöngyön magasabb energiájú szingulett oxigén jön létre, ami – a vizsgált kölcsönhatás létrejötte esetén – az akceptor gyöngyön lévő fluorofórt aktiválja. A szingulett oxigén élettartama rövid, kb. 200 nm távolságra tud diffundálni. A donor gyöngyön másodpercenként kb. 60 000 szingulett oxigén keletkezik, ami a jel nagy mértékű felerősödését eredményezi, így a módszert még miniatürizált formában (8 μ l) is kimagasló érzékenység, nagy dinamikus mérési tartomány és automatizálhatóság jellemzi [204].

Az élesztő-egy-hibrid rendszer a fehérje–fehérje kölcsönhatások vizsgálatára alkalmas élesztő-két-hibrid eljárás módosított változata. A rendszer két fő összetevőből áll: egy riporter konstrukció a riporter gén és az azt megelőző indukálható promóter előtt tartalmazza a vizsgálni

kívánt DNS-szakaszt, egy expressziós konstrukció pedig egy fúziós fehérjét kódol, melynek egyik része az elemzett – feltételezetten – DNS-kötő fehérje, másik fele pedig a riporter gén indukálható promoterét aktiváló molekula. Amennyiben létrejön a kölcsönhatás a vizsgált DNS és fehérje között, a riporter gén promotere aktiválódik, ami a riporter fehérje megjelenésével detektálható. Ennek egyik gyakran alkalmazott módja olyan élesztő sejtek alkalmazása, melyek valamely az anyagcserében alapvető molekula (rendszerint hisztidin, máskor triptofán, leucin vagy uracil) előállítására képtelenek, a riporter pedig éppen az az enzim, ami a megfelelő vegyület szintézisét katalizálja. Ennek megfelelően csak azok a sejtek élnek túl, illetve növekednek, ahol a riporter gén kifejeződik, azaz létrejött a kölcsönhatás a vizsgált (mesterségesen bejuttatott) DNS és fehérje között. Más rendszerek a bakteriális *lac*-operonhoz, azaz a β -galaktozidázhoz kapcsolt színreakció kimutatásán alapulnak. Megjegyzendő, hogy a rendszer – értelemszerűen – a DNS és a fehérje fizikai kapcsolatának kimutatására alkalmas, a riporter gén promoterét a fúziós fehérje másik alegysége, az aktiváló domén indukálja. Így függetlenül a vizsgált fehérje *in vivo* gátló vagy aktiváló hatásától a kölcsönhatás kialakulása esetén mindenképp megemelkedő riporterszint mérhető [205]. Ismert a rendszer „fordított” változata is, melynél a riporter konstrukció valamilyen toxikus anyagot fejez ki, így az elemzett DNS és fehérje kapcsolódása a sejtek túlélését gátolja [206]. Ezzel a rendszerrel hatékonyan azonosíthatók olyan mutációk, melyek a DNS–fehérje kölcsönhatás kialakulását megakadályozzák, mivel így a riporter konstrukció aktiválódásának elmaradása miatt a sejtek életben maradnak.

Bár a „kimenet” teljesen eltérő, valamelyest hasonló fúziós molekulák alkalmazásán alapul a nemzetközi szakirodalomban „proximity ligation assay” néven ismert technika. A módszer lényege, hogy mind a vizsgált kettőszálú DNS-hez, mind a hozzá feltételezetten kapcsolódó fehérjét specifikusan felismerő antitestre egy-egy rövid, egyszálú DNS-láncot helyeznek. Ha létrejön a DNS–fehérje komplex (amihez kötődik az antitest is), akkor a két egyszálú DNS egymás közelségébe kerül, és egy (végükkel komplementer szekvenciájú) összekötő oligonukleotid segítségével összeligálható. A keletkezett DNS-lánc real-time PCR-rel kimutatható, ennek következtében a módszer kimagaslóan érzékeny és a kötődés erősségét jellemző mennyiségi mérések elvégzésére is alkalmas [207].

A szabályozó régiók vizsgálatának régóta ismert hatékony és viszonylag egyszerű eszköze a riporter konstrukciók létrehozása és elemzése. Ennek lényege, hogy a riporter vektor egy olyan fehérjét kódol, melynek mennyisége pontosan és érzékenyen mérhető: igen elterjedt a

biolumineszcens fényt kibocsátó luciferáz enzim alkalmazása. Számos különböző luciferáz riporter vektor (pl. pGL3-, pGL-4 család, pMIR-Report, pmirGLO Dual-Luciferase stb.) létezik, melyek különböző tulajdonságú saját promoterral és eltérő lokalizációjú klónozó hellyel rendelkeznek, így az inzert, azaz a vizsgálni kívánt szabályozó DNS-régió a kísérleti kérdésnek megfelelő helyre (riportert kódoló gén elé vagy mögé) és megfelelő környezetbe (pl. promoterként vagy promoter mellé) klónozható. Az így létrehozott mesterséges gén élő sejtben a sejt saját transzkripciós és transzlációs apparátusával, illetve esetlegesen mesterségesen kotranszfektált egyéb faktorok jelenlétében kifejeződik: a keletkező riporter fehérje mennyisége alapján következtetni lehet a vizsgált DNS-szakasz szabályozó hatására [208]. Irányított mutagenézissel a konstrukcióban egyetlen vagy több bázispárra kiterjedő szekvenciaváltozatok könnyen létrehozhatók, így ez a rendszer a szabályozó régiókban (pl. promoter, 5' és 3' UTR) elhelyezkedő SNP-k, pontmutációk és rövid ismétlődési variációk biológiai funkcióinak elemzésére eredményesen alkalmazható.

2.3.3.2 A miRNS-ek hatásának vizsgálata

Az elmúlt évtizedekben egyre több figyelem irányult a kb. 22 nukleotid hosszúságú miRNS-ek révén megvalósuló szabályozó mechanizmusokra, ami részben annak köszönhető, hogy szerepük számos kóros folyamat, illetve betegség kialakulásához kapcsolódóan ismertté vált. A miR-124a expresszióját befolyásoló rs531564 SNP összefüggést mutat a 2-es típusú cukorbetegséggel [209], a miR-34a megemelkedett szintje a hasnyálmirigy β -sejtjeinek apoptózisához köthető [210], a miR-16 család [211], a miR-196, a miR-132 és a miR-134 pedig különböző daganatok kialakulásával áll kapcsolatban [212]. A miRNS-ek hatásának feltérképezése nyilvánvaló jelentőségű tehát, vizsgálatuk és megismerésük a diagnosztika és a kezelés eszköze lehet. A miR-34 szolid tumorok kezelésének fázis I. vizsgálati szintjéig jutott, ahol azonban ismeretlen okból súlyos immunreakció kialakulását eredményezte, ami jól szemlélteti a miRNS-ekben rejlő ígéretes lehetőségeket csakúgy, mint további elemzésük és megismerésük fontosságát [213].

A miRNS-ek rendszerint az mRNS-ek 3' UTR-éhez kötődve fejtik ki legtöbbször gátló hatásukat a transzláció vagy az mRNS élettartamának befolyásolása révén. Funkciójuk pontos, átfogó megismerését az is nehezíti, hogy egyetlen miRNS számos célgén kifejeződésére lehet

hatással, és egyetlen mRNS-t nagy számú miRNS szabályozhat, az így kialakuló bonyolult hálózat feltérképezése hatékony informatikai megközelítést igényel. Tovább bonyolítja a képet, hogy az egyszerű szekvencia analízis az interakciók megítéléséhez nem elegendő, így a molekuláris genetikai elemzések szintén nélkülözhetetlenek ezen rendszer megismerésében.

A miRNS-ek funkciójának vizsgálata során az *in silico*, on-line eszközök hasznos kiindulópontot jelentenek. A pri-miRNS-ek hajtú szerkezete, az érett miRNS-ek szekvenciája, valamint a vele kapcsolatos tudományos publikációk fellelhetők a miRBase-ben. A Diana–TarBase és a miRTarBase adatbázisok a laboratóriumi eljárásokkal igazolt mRNS–miRNS kölcsönhatásokat gyűjti össze, a miRWalk jósló és kísérletesen validált kapcsolatokat egyaránt tartalmaz [214]. A miRGator 3.0 adattárban a miRNS-ek expressziós mintázatáról is tájékozódhatunk [215], ezek különböző humán tumorokra jellemző megváltozása olvasható a miRCancer adatbázisban, a PolymiRTS pedig a miRNS-eket kódoló szakaszokon, illetve a kötőhelyekben megtalálható SNP-eket foglalja össze. A Firefly Discovery Engine a miRNS-ekkel kapcsolatos tudományos közlemények tárháza. A miRNS-ek célgénjeinek jóslására számos on-line eszköz (miRDB, TargetScan, miRSearch) elérhető.

Az mRNS–miRNS kölcsönhatás biológiai szerepe és működése természetesen alapvetően eltér a DNS–fehérje kölcsönhatásokétól, mégis ebben az esetben is két molekula között kialakuló kapcsolat elemzése a feladat, így a korábban bemutatott technikák egy része – megfelelő módosításokat követően – alkalmazható ezen szabályozási rendszerek kutatására is.

Logikus lehetőség a luciferáz riportter rendszerek alkalmazása, melynek során a vizsgálni kívánt 3' UTR-t természetes elhelyezkedésének megfelelően a luciferázt kódoló szakasz mögé kell klónozni, erre alkalmas pl. a pMIR-Report vagy a pmirGLO Dual-Luciferase vektor. A rendszer alkalmas nemcsak a „seed” szekvencia, de akár egyetlen SNP hatásának pontos tanulmányozására, és említésre méltó, hogy sok esetben a sejtes rendszerben az *in silico* jósló kölcsönhatások nem megerősíthetők, ugyanakkor olyan miRNS–mRNS kapcsolatok is kimutathatók, melyeket a predikciós eljárások nem jeleznek [216].

A célgének miRNS hatására bekövetkező expresszió változása szintén információt szolgáltat arról, hogy egy miRNS mely szabályozási folyamatokban játszik szerepet. Egy miRNS túltermelését vagy gátlását követően a vele kapcsolatban álló gének kifejeződése megváltozik. Annak ellenére, hogy a miRNS-ek az mRNS-hez kötődve a transzkripciót követően fejtik ki hatásukat, mégis sok esetben hatásuk – pl. microarray alapú mérés segítségével –

mRNS-szinten is nyomon követhető [217], ráadásul szövet specifikus módon. Egyetlen miRNS akár több mint száz mRNS működését, illetve mennyiségét befolyásolhatja, ami a rendszer hálózatos jellegén túlmenően központi szabályozók szerepét is felvetheti [218]. A fehérjeexpresszió mérése nagy áteresztőképességű tömegspektrometriás eljárással megvalósítható: a kimutatható változások annál nagyobb összhangban állnak az mRNS szintű mintázattal, minél hosszabb (7 vagy 8 nukleotid) szakaszon hibridizál a miRNS az mRNS-sel [219].

Immunprecipitáció segítségével azonosíthatók azok a célgének, melyek egy adott miRNS-sel közvetlen kölcsönhatásban állnak, mivel az mRNS–miRNS kapcsolat a RISC komplex kialakulásán keresztül jön létre. Az eljárás leggyakrabban az argonauta fehérjét specifikusan felismerő antitestekkel végezhető el, a komplex stabilitása azonban viszonylag alacsony, emiatt disszociálhat a RISC és az mRNS az immunprecipitáció során [220]. Megoldást jelenthet erre, hogy UV-fény hatására a fehérje és az RNS fényérzékeny csoportjai között keresztkötések jönnek létre, aminek hatékonysága 4-tiouridin alkalmazásával tovább fokozható [221]. Lehetőség van az mRNS és a miRNS ligálására is az immunprecipitáció folyamat során, melynek révén a miRNS és célgénje együtt szekvenálható [222].

Az alapelv hasonló azon eljárások esetén, melyek – rendszerint biotinnal – jelölt miRNS-eket alkalmaznak a target mRNS-ek azonosítása céljából [223, 224]. A sejtben létrejövő biotinilált miRNS–mRNS komplexek streptavidin segítségével izolálhatók, a rendszer érzékenysége és hatékonysága fényérzékeny csoportok (pl. psoralennel konjugált uridin) alkalmazásával növelhető, mivel a nukleinsavak között UV fény segítségével stabil keresztkötések hozhatók létre [225]. Szem előtt kell azonban tartani, hogy a miRNS-ek működését a biotin jelentősen befolyásolja, és egyes esetekben megfigyelhető, hogy a biotinilált miRNS affinitása az argonauta fehérjék iránt lényegesen csökken [226]. Ennek megfelelően a miRNS-ek jelölésére alternatív megközelítéseket dolgoztak ki. Hatékonyak bizonyult egy tetrazolt, valamint két benzol gyűrűt tartalmazó csoporttal módosított miRNS-ek használata. Az így konjugált molekulák megőrizték eredeti funkciójukat, és hatékonyan beépültek a RISC-be. A miRNS–mRNS komplex létrejöttét követően a miRNS-en lévő mesterséges kötőhelyhez fumaráttal konjugált biotin UV fény hatására stabilan bekötődött („fotoklikk” reakció), a létrejött komplexeket pedig streptavidin segítségével lehetett hatékonyan izolálni [227].

3 CÉLKITŰZÉSEK

A bemutatott kutatómunka fő célkitűzése genetikai polimorfizmusok három – DNS-, molekuláris (génkifejeződésre gyakorolt) és fenotípusos – szintű elemzése volt.

- Hatékony genotipizáló eljárást kívántunk kidolgozni a humán *DRD4* és *WFSI* génekben lévő kiválasztott SNP-kre és hosszúságpolimorfizmusokra, valamint a kutya dopaminerg rendszeréhez tartozó gének (*DRD4*, *DBH*, *TH*) egyes hosszúságvariációinak elemzésére. Célunk volt ezen utóbbi génekben új, korábban nem ismert polimorfizmusok azonosítása.
- A genotipizáló és haplotipizáló eljárások részeként célunk volt olyan elektroforetikus rendszerek – ultravékony és kapilláris gélelektroforézis, valamint mikrofluidikai eszköz – beállítása és alkalmazása, melyekkel az elemzés szűk keresztmetszetét jelentő fragmentum analízis hatékonysága és felbontóképessége növelhető.
- Real-time PCR alapú eljárásokat kívántunk kidolgozni és alkalmazni a *C4A*, illetve *C4B*, valamint a *GSK3B* géneket érintő CNV-k elemzésére.
- Céljaink között szerepelt a genotipizáló rendszerek továbbfejlesztése oly módon, hogy alkalmassá váljanak szomszédos polimorfizmusok direkt, molekuláris haplotípuselemzésére.
- A beállított genotipizáló és haplotipizáló eljárásokat asszociáció vizsgálatok során kívántuk alkalmazni. Vizsgálni kívántuk a *WFSI* és a monoaminerg rendszer egyes génjeiben lévő kiválasztott polimorfizmusok és az agresszió közötti feltételezett összefüggést.
- Céljaink között szerepelt továbbá a *SNAP-25* SNP-i és az impulzivitás, valamint a *WFSI* polimorfizmusai és az 1-es, illetve 2-es típusú cukorbetegség közötti összefüggés elemzése.
- Kutatni terveztük, hogy bizonyos polimorfizmusok szerepet játszhatnak-e a génkifejeződés szabályozásában: elemezni kívántuk a *SNAP-25* promoterében, illetve 3' UTR-ében elhelyezkedő SNP-k hatását *in vitro* rendszerben.
- Célul tűztük ki a genetikai asszociáció elemzések során pozitív eredményt mutató *WFSI* polimorfizmusok transzkripció aktivitására, valamint miRNS-kötődésére gyakorolt moduláló hatásának tanulmányozását.

4 MÓDSZEREK

4.1 Polimorfizmusok genotipizálása

4.1.1 Vizsgálati személyek, mintavétel, DNS-izolálás

Az asszociáció vizsgálatokban résztvevő egészséges személyek önkéntes alapon vettek részt kutatásainkban. Az 1-es, illetve 2-es típusú diabetes mellitus-ban szenvedő betegek a Semmelweis Egyetem II., illetve III. Belgyógyászati Klinikáján, a hangulatzavarral diagnosztizált résztvevők a Semmelweis Egyetem Pszichiátriai és Pszichoterápiás Klinikáján álltak kezelés alatt. A cukorbetegség diagnózisának felállítása éhgyomri vércukorértékek, valamint az orális glükózterhelés teszt (OGTT) alapján történt a WHO előírásai szerint. A vizsgálatokat az Egészségügyi Tudományos Tanács Tudományos és Kutatásetikai Bizottsága (ETT-TUKEB) hagyta jóvá. A mintavétel, valamint a genetikai elemzések megkezdése előtt minden résztvevő részletes írásbeli tájékoztatót kapott a kutatás céljáról és menetéről, valamint írásos beleegyező nyilatkozatot töltött ki. Az egyedi kóddal ellátott mintákat anonim módon kezeltük. Az eset-kontroll elemzések során etnikailag homogén (kaukázusi eredetű) személyekből álló, nemben és korban illesztett csoportokat hasonlítottunk össze.

A DNS-mintavétel non-invazív módon történt: a vizsgálati személyek vattapálca segítségével az íny vagy a bucca területét dörzsölve szájnyálkahártya-sejteket gyűjtöttek. A mintákat további feldolgozásig $-20\text{ }^{\circ}\text{C}$ -on tároltuk. A DNS-tisztítás a sejtek lízisével kezdődött (0,1 M NaCl, 0,01 M Tris-HCl (pH = 8), 0,5% SDS, 0,2 mg/ml proteináz K). Az $56\text{ }^{\circ}\text{C}$ -os inkubációt követően a fehérjéket telített NaCl oldattal kicsaptuk. Centrifugálás után a DNS-t a felülúszóból izopropanollal és etanollal precipitáltuk, a csapadékot $0,5\times$ TE pufferben (5 mM Tris pH = 8, 0,5 mM EDTA) oldottuk fel.

4.1.2 Genotipizálás „hagyományos” PCR-rel

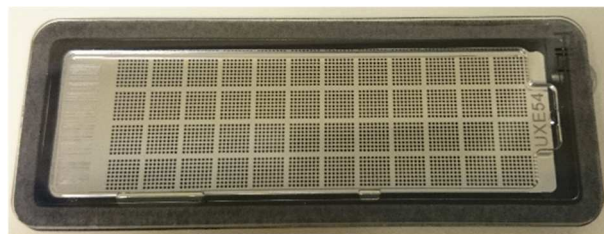
A VNTR-ek, valamint SNP-k PCR-RFLP-vel, illetve allél-specifikus PCR-rel történő genotipizálása során a Qiagen HotStarTaq[®] DNS-polimeráz kitet alkalmaztuk. A méréseket 6–10 μl végtérfogatban végeztük el, a reakcióelegyek 1–1 μM primereket (Függelék S1. táblázat), 200 μM dATP-t, dCTP-t, dGTP-t és dTTP-t; $1\times$ puffer oldatot, $1\times$ Q-oldatot, 0,25–0,5 U DNS

polimerázt és 4–10 ng genomi DNS-t tartalmaztak. A keletkező termékeket – várható méretük-től függően – különböző koncentrációjú hagyományos vagy metaphor agaróz gélelektroforézissel választottuk el egymástól. A *WFSI* SNP-k primer extenzióval történő elemzése során a régió amplifikálást követően a megmaradt primereket *Exo I* emésztéssel, a még jelenlévő dezoxiribonukleozid-trifoszfátokat pedig rekombináns alkalikus foszfáttal (rSAP) bontottuk le. Ezt követte az SBE reakció, melynek során a SNPStart® MasterMix oldatot alkalmaztuk. A feleslegben maradt, jelölt láncterminátor nukleotidok foszfát csoportjait rSAP-emésztéssel távolítottuk el, hogy az elektroforézis során ne adjanak zavaró szignált.

4.1.3 Genotipizálás real-time PCR-rel

Az SNP-k real-time PCR-rel történő elemzése során saját tervezésű, illetve kereskedelmi forgalomban beszerezhető primer- és próbapárokat alkalmaztunk. A 10 µl végtérfogatú reakcióelegy – a PCR alapvető komponensei mellett – ROX referencia festéket és AmpliTaq Gold® DNS-polimerázt tartalmazott. A laboratóriumunkban működő QuantStudio 12K Flex berendezés lehetővé tette a genotipizálás miniaturizált, nagy áteresztőképességű kivitelezését. Az OpenArray lemezek előre meghatározott elrendezésben tartalmazzák a megfelelő SNP-k genotipizálásához szükséges primereket és próbákat. A DNS-eket és a reakcióelegyet az AccuFill rendszer viszi fel: egy lemez (6. ábra) 48 db 64 mintahelyből álló régiója összesen 3072 párhuzamos elemzést tesz lehetővé, a reakciók 33 nl térfogatban zajlanak.

A *WFSI* gén promoterében lévő in / del polimorfizmus analízise során a régiót PCR-rel felsokszoroztuk, majd a reakcióelegyhez SYBR Green interkalálódó festéket adtunk, és a termék olvadáspontját real-time PCR berendezés segítségével határoztuk meg.



6. ábra. OpenArray lemez. Egy lemezen a 48 · 64 mintahelynek megfelelően párhuzamosan 3072 real-time PCR mérés végezhető el. A reakciók végtérfogata 33 nl.

A CNV-k elemzésekor a génszámot kvantitatív real-time PCR-rel állapítottuk meg. Minden esetben három technikai párhuzamost alkalmaztunk, a relatív kvantifikálás a ΔC_T -módszerrel történt. Belső kontrollként olyan genomi régiót (RNáz P gén) amplifikáltunk, melynek génszáma nem mutat polimorf változatosságot.

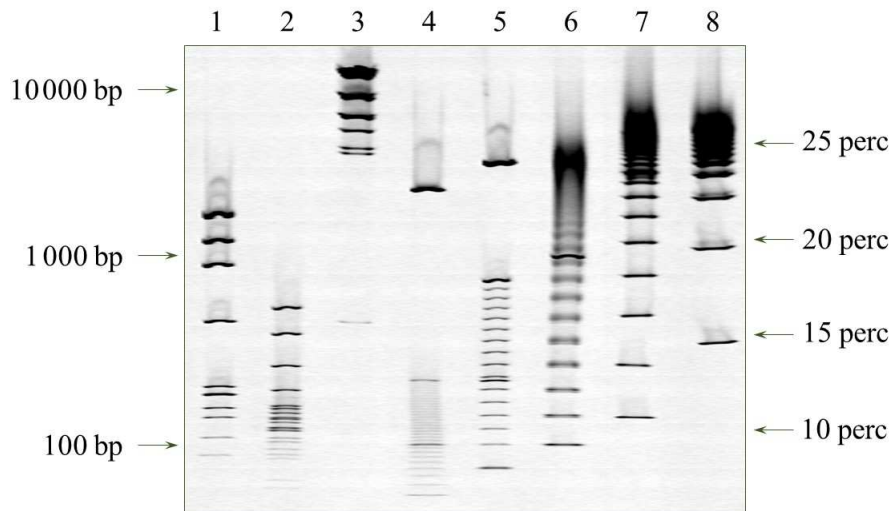
4.1.4 Automatizált elektroforetikus technikák

A genotipizáló munka során a nagyobb áteresztőképességű és jobb felbontású fragmentum analízis céljából a hagyományos, horizontális alámerülő agaróz gélelektroforézis mellett különböző, hatékony elválasztástechnikai eljárásokat is alkalmaztunk.

Az ultravékony gélelektroforézis során a DNS-fragmentumokat egy üvegekazettában lévő 190 μm vastagságú agaróz–lineáris poliakrilamid gélben választottuk el, a detektálás – a kapilláris elektroforézishez, illetve a mikrofluidikai eszközökhöz hasonlóan – lézer indukált fluoreszcencia segítségével az elválasztás alatt („időben”) történt. Az üvegekazetta két végén egy-egy 15 ml-es beépített puffer tartály található, ezeken keresztül történt a gél betöltése, illetve ide csatlakoztak az elektródok az elektroforézis során. A valós idejű detektálás miatt ugyanabban a gélben néhány órás időtartamon belül több elválasztás is elvégezhető. Kísérleteink során 2% agaróz–1% poliakrilamid–20 ng/ μl EtBr–1 \times TBE (89 mM tris, 89 mM borát, 2 mM Na₂EDTA puffer) gélmátrixot használtunk, mely igen széles (10–20 000 bp) mérettartományban alkalmasnak bizonyult a DNS-fragmentumok nagy felbontású elektroforetikus analízisére (7. ábra). A gélt folyékony formában (60 °C) töltöttük a kazettába. A minták injektálása egy fésű alakú membrán alkalmazásával történt, a „fésűfogakra” 0,2 μl mintát pipettáztunk, és a membránt a kazettába, a gélhez illesztettük. A DNS-molekulák az elektromos térerő hatására a membránról a gélbe jutottak, majd a gélben megkezdődött az elválasztás. A nagy felszín–térfogat arány hatékony hőleadást biztosított, így a hagyományos alámerülő technikához képest kb. egy nagyságrenddel magasabb (általában 75 V/cm) térerő volt alkalmazható. A detektor a gélt a kiindulási ponttól 2–6 cm távolságban folyamatosan pásztázta: a kazetta méretétől függően egy gélben párhuzamosan akár 96 minta elemzése is elvégezhető volt.

A kapilláris gélelektroforézis során a Beckman Coulter MDQ egy kapillárisos, illetve GeXP nyolc kapillárisos berendezésekkel dolgoztunk. Kísérleteink során 50–180 μm belső átmérőjű kapillárisokat alkalmaztunk, az elválasztás során a kapilláris effektív hossza (l) 20 és

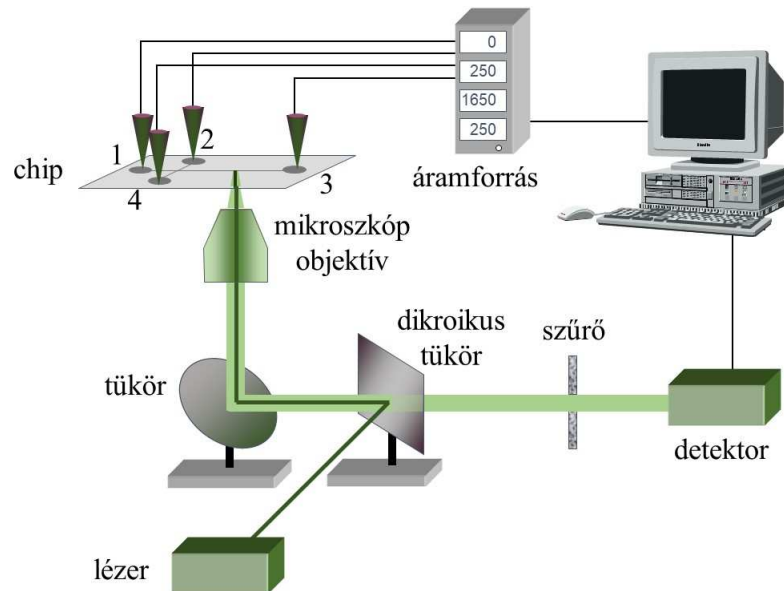
50 cm között változott. A mintákat elektrokinetikusán injektáltuk, az elválasztás 2–6 kV feszültséggel történt.



7. ábra. DNS-fragmentumok elválasztása ultravékony gélelektroforézissel. A 2% agaróz–1% poliakrilamid gél 10–20 000 bp tartományban hatékony elválasztást tesz lehetővé. 1: *Hae* III enzimmel emésztett Φ X 174 DNS, 2: *Msp* I endonukleázzal hasított pBR322 plazmid, 3: *Hind* III enzimmel emésztett λ DNS, 4–8: 10, 50, 100, 200 és 500 bp-os DNS-létra.

A mikrofluidikai eszköz elválasztási paramétereinek elemzése során 100 μ m-es „dupla-T” injektoros, üvegből készült elektroforetikus chipet alkalmaztunk. Az elválasztó csatorna teljes hossza 80 mm volt, az effektív hossz 10 és 75 mm között volt változtatható. Az elválasztások előtt a csatornákat 2% polivinilpirrolidon (PVP, $M_w = 1\,300\,000$ g/mol, $1\times$ TBE) oldattal mostuk át, ami a csatornák falán bevonatot képezve az elektroosmotikus áramlást hatékonyan csökkentette. Az elválasztásokat 0,5–2,5%-os polivinilpirrolidon használatával végeztük el. A mintákat a 4-es tartályból elektrokinetikusán injektáltuk ($U_1 = 240$ V, $U_2 = 480$ V, $U_3 = 240$ V, $U_4 = 0$ V), az elválasztás során 165 V/cm elektromos térerőt ($U_1 = 0$ V, $U_2 = 250$ V, $U_3 = 1650$ V, $U_4 = 250$ V) alkalmaztunk (8. ábra). A DNS molekulákat 0,5 μ M Sytox Orange interkalátor festék segítségével tettük láthatóvá. A DNS–Sytox Orange komplexet 532 nm-es zöld lézerrel gerjesztettük, a lézerfényt egy dikroikus tükrön keresztül mikroszkóp segítségével irányítottuk az elektroforetikus chip elválasztó csatornájának megfelelő pozíciójába. Ugyanez a mikroszkóp

objektív gyűjtötte össze a kibocsátott fluoreszcens jelet ($\lambda = 570 \text{ nm}$), ezt a fényt a dikroikus tükör átengedte, s így az a detektorba jutott (8. ábra).



8. ábra. A kísérleteink során alkalmazott mikrofluidikai eszköz felépítésének vázlata. Az 532 nm hullámhosszúságú lézert (vékony sötétzöld vonal) a dikroikus tükör visszaveri, így az a mikroszkópon keresztül megvilágítja az elektroforetikus chip elválasztó csatornáját. A fluoreszcens fény (szélesebb világoszöld sáv) a dikroikus szűrőn áthaladva a detektorba jut.

4.1.5 Haplotípus / LD vizsgálata

A polimorfizmusok „direkt”, PCR-alapú haplotípus meghatározása során a genotipizáláshoz hasonló technikákat alkalmaztunk. A HotStar Taq DNS-polimeráz kittel felszorzott PCR-termékeket hagyományos, alámerülő agaróz vagy kapilláris gélelektroforézis segítségével tettük láthatóvá. Nagyobb kiterjedésű génszakaszok, illetve több SNP esetén az egyes haplotípusok gyakoriságát, valamint a polimorfizmusok közötti LD-t a HaploView 4.2 segítségével számoltuk ki [228]. Az egyes személyek individuális haplotípusát a Phase 2.1 programcsomag alkalmazásával határoztuk meg [229].

4.2 Asszociáció vizsgálatok

4.2.1 Fenotípus-elemzés

Az 1-es, illetve a 2-es típusú diabetes mellitus-ban szenvedő betegek kezelése, illetve a kórképek diagnózisának felállítása a Semmelweis Egyetem II. és III. számú Belgyógyászati Klinikáján történt.

Az agresszió mérésére a Buss–Perry Agresszió Kérdőívet használtuk. Ez a kérdőív 29 kérdésből áll, 4 alskálájának megfelelően négy dimenzió: a fizikai és verbális agresszió, a hosztilitás (ellenségesség, rosszindulatúság – az agresszió kognitív összetevője) és a harag meghatározására alkalmazható.

A Barratt Impulzivitás Skála az egyik leggyakrabban alkalmazott önkitöltős kérdőív az impulzív viselkedés elemzésére. A teszt 30 tételt tartalmaz, a kérdőívet kitöltő személy reakcióiban megjelenő lobbanékonyt (motoros alskála), a jövőbeli események megtervezését (tervezés alskála), valamint a gondolkodásban megjelenő impulzivitást (kognitív funkciók alskála) határozza meg.

4.2.2 Statisztikai elemzés

A polimorfizmusok elemzése során alkalmazott genotipizáló technikák megbízhatóságát a Hardy–Weinberg-egyensúly tesztelésével ellenőriztük. A mért és az allélfrekvenciákból számított várható genotípus-gyakoriság értékeket, valamint eset–kontroll elemzések során a csoportok allél-, genotípus-, illetve haplotípusfrekvencia értékeit χ^2 -próbával hasonlítottuk össze. Azon kísérletek során, amikor a fenotípust folytonos változóval jellemeztük, az egyes genotípus- vagy haplotípus-csoportba tartozó személyek valamely tulajdonságot jellemző átlagértékei közötti különbséget egyszempontos variancia analízissel (ANOVA) vizsgáltuk. Mivel sok esetben számos polimorfizmust elemeztünk egyidejűleg, a többszörös tesztelés miatt a Bonferroni-korrektúrával módosított statisztikai szignifikancia (p) értéket vettük figyelembe az eredmények kiértékelése során.

4.3 Molekuláris biológiai vizsgálatok

4.3.1 Luciferáz riporter konstrukciók készítése

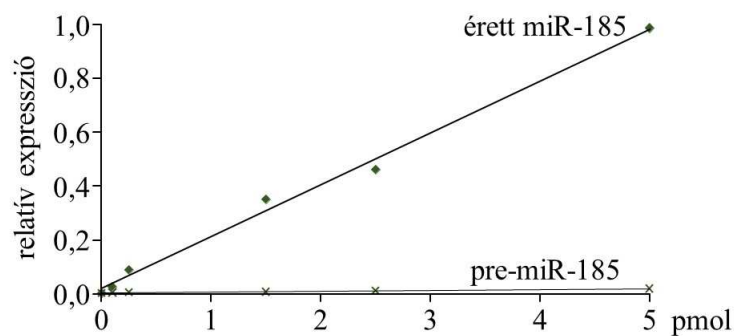
A szabályozó régiók polimorfizmusainak elemzéséhez pGL3-Basic, pGL3-Control és pMIR-Report vektorokat alkalmaztunk. Az inzerteket ismert genotípusú személyek genomi DNS-éről állítottuk elő PCR-rel, a primerek (*Függelék S1. táblázat*) 5' végén a klónozó helynek megfelelő restrikciós endonukleáz felismerőhely volt megtalálható. A *WFSI* promoter polimorfizmusait pGL3-Basic, ugyanezen gén, valamint a *SNAP-25* 3' UTR-ében lévő SNP-ket pedig pMIR-Report és pGL3-Control rendszer segítségével vizsgáltuk. A vektorokat és inzerteket a megfelelő restrikciós endonukleázzal emésztettük, majd a rekombináns DNS-t T4-ligáz segítségével készítettük el, és XL10-Gold® *E. coli* kompetens sejtekbe transzformáltuk. Az elkészült konstrukciókat PureYield™ Mini-, illetve Midiprep segítségével tisztítottuk, az inzert bázissorrendjét és megfelelő beépülését Sanger-szekvenálással ellenőriztük. A kísérletek során elemzett további variánsokat QuikChange® Lightning irányított mutagenézissel hoztuk létre.

4.3.2 Tranziens transzfekció, luciferáz aktivitás mérése

A promoterben és a 3' UTR-ben elhelyezkedő polimorfizmusok szabályozó hatásának luciferáz rendszerrel történő vizsgálata során HEK293T (humán embrionális vese) sejteket alkalmaztunk. A sejtvonalat termosztátban (37 °C, 5% CO₂), 10% FBS-t és 1% Penicillin–Streptomycin antibiotikumot tartalmazó DMEM oldatban tartottuk fenn. Az egyes mérésekhez a sejteket 24 lyukú tálcába, 500 µl médiumba tettük ki, majd 24 óra múlva Lipofectamine™ 2000 segítségével transzfektáltuk. A relatív luciferáz aktivitás meghatározása céljából a sejteket minden esetben β-galaktozidázt kódoló konstrukcióval kotranszfektáltuk, a mérések során a riporter fehérje aktivitását ezen enzim mennyiségével normalizáltuk.

A kísérletek megkezdése előtt vizsgáltuk, hogy milyen mennyiségű konstrukciók transzfektálása vezet a legoptimálisabb eredményhez. Azt a tartományt kerestük, melyben a relatív luciferáz aktivitás és annak változása megbízhatóan kimutatható. Ennek meghatározására a sejteket növekvő mennyiségű konstrukciókkal transzfektáltuk, és mértük a keletkező enzimek aktivitását. A 3' UTR-ben lévő SNP-k elemzése során a sejteket a luciferázt és a β-galaktozidázt kódoló konstrukció mellett pre-miRNS-sel is transzfektáltuk, ebben az esetben real-time PCR alapú eljárással azt is elemeztük, hogy a prekursor átalakul-e érett miRNS-sé (*9. ábra*).

A kalibrációs kísérletek eredményei alapján a mérések során 0,125 μg pGL3-, 0,05 μg pMIR-Report, valamint 0,2 μg β -galaktozidáz kódoló konstrukciót alkalmaztunk. Megállapítottuk, hogy a kimutatható miRNS szintje a transzfektált mennyiséggel megfelelő összefüggést mutat (9. ábra, érett miR: $R^2 = 0,994$, pre-miR: $R^2 = 0,977$), és ebben a tartományban a prekursor csaknem maradéktalanul átalakul érett formává. Az SNP-k hatásának elemzése során a sejteket 5 pmol miRNS-sel transzfektáltuk. A transzfekciós reakcióelegy a megfelelő konstrukciók mellett 2,5 μl 1 mg/ml LipfectamineTM 2000 oldatot tartalmazott 125 μl végtérfogatban, melyet a liposzómák kialakulását követően (20 perc inkubáció) adtunk az előkészített sejtekhez.



9. ábra. Prekursor miRNS átalakulása érett formává. HEK293T sejteket növekvő mennyiségű pre-miR-185-tel transzfektáltunk, majd meghatároztuk az érett miRNS és a prekursor mennyiségét. Az ábra egy reprezentatív kísérlet eredményét mutatja. A három technikai párhuzamos közötti különbségek szintje olyan alacsony, hogy az az ábrán (hibaszív formájában) nem jeleníthető meg.

A sejteket a transzfekció után 24 órával begyűjtöttük, három fagyasztás–olvasztás lépéssel feltártuk, majd centrifugálást (13000 rpm, 15 perc, 4 °C) követően az enzimeket tartalmazó felülúszót új Eppendorf csövekbe mértük. Az enzimaktivitások meghatározása Varioskan Flash berendezés segítségével történt, amely luminometriás és fotometriás mérések elvégzésére egyaránt alkalmazható. A luciferáz enzim mennyiségét 0,16 mg/ml luciferin kálium sótt tartalmazó reagenssel, a β -galaktozidáz aktivitását 40 mg/ml ONPG-oldattal (orto-nitrofenil- β -galaktozid) határoztuk meg. Minden kísérletben három párhuzamos mérést alkalmaztunk. A luciferáz enzim aktivitását a β -galaktozidáz mennyiségével normalizáltuk, az egyes kísérleti pontok eredményeit egyszempontos ANOVA elemzéssel hasonlítottuk össze.

5 EREDMÉNYEK ÉS MEGBESZÉLÉS

5.1 Polimorfizmusok genotipizálása

A polimorfizmusok és mutációk klinikai diagnosztikai, molekuláris, illetve asszociáció elemzése során egyaránt alapvető kiindulási pont a megbízható genotipizáló eljárás beállítása és használata. Bár a legtöbb módszer elve, működése első pillantásra egyértelműnek tűnhet, optimalizálásuk és alkalmazásuk szinte minden esetben gondos munkát és körültekintést tesz szükségessé. A genetikai polimorfizmusok elemzésére irányuló kutatásaink három fő pilléren nyugszanak (3. ábra). Első lépésben hatékony genotipizáló eljárásokat állítottunk be, melyeket ezt követően asszociáció vizsgálatok keretében egyes komplex jellegek örökletes hátterét kutatva alkalmaztunk. Végül a harmadik – talán legfontosabb – lépésként azt vizsgáltuk, hogy a pozitív asszociációt mutató polimorfizmusokhoz köthető-e molekuláris biológiai funkció.

Munkánk egyik kezdeti feladata ennek megfelelően olyan eljárások beállítása és optimalizálása volt, melyekkel a kiválasztott célgének (elsősorban a *DRD4*, a *WFS1* és a *SNAP-25*) egyes polimorfizmusai megbízhatóan és hatékonyan genotipizálhatók. Közeli variációk esetén külön figyelmet fordítottunk a haplotípus, azaz az allélok relatív kromoszomális lokalizációjának meghatározására, mivel ez a biológiai hatás szempontjából döntő jelentőségű lehet. Bár az SNP-k és az ismétlődési variációk vizsgálata valamelyest eltérő megközelítést igényel, közös bennük, hogy az elemzés szűk keresztmetszetét gyakran az elektroforetikus analízis jelenti. Munkánk során ezért hangsúlyt fektettünk olyan technikák – ultravékony és kapilláris gélelektroforézis, mikrofluidikai eszköz – alkalmazására, melyekkel a DNS-fragmentumok elválasztása automatizálható, az elektroforézis hatékonysága és érzékenysége fokozható.

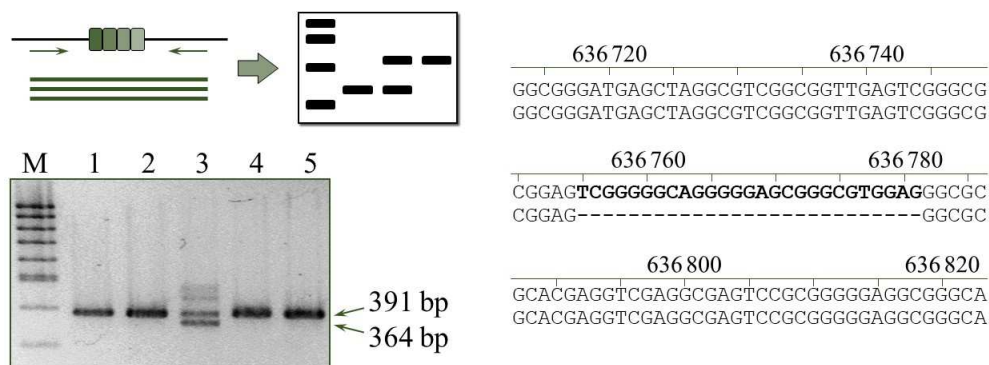
5.1.1 Ismétlődési variációk és SNP-k elemzése

Napjaink nagy áteresztőképességű genotipizáló technikái elsősorban az SNP-k elemzésére fókuszálnak, ugyanakkor az ismétlődési variációk mégis több szempontból különös érdeklődésre tarthatnak számot. Bár számuk az SNP-kénél kétségtelenül alacsonyabb, rendkívül heterogén csoportot képeznek, mivel mind az ismétlődő modul hossza, mind pedig az ismétlődési szám igen variábilis lehet, így jelentős biológiai funkcióval rendelkezhetnek, és nagymértékben hozzájárulnak az emberek közötti genetikai különbségek kialakításához.

5.1.1.1 A dopaminerg rendszer kiválasztott polimorfizmusainak elemzése

A *DRD4* gén 3. exonjában lévő 48 bp-os VNTR elemzése 25 éves múltra tekint vissza: a pszichogenetikai asszociáció elemzések hajnalán összefüggést figyeltek meg a 7 ismétlődést tartalmazó allél és az újdonságkeresés személyiségjegy között [55, 61]. Korábbi munkánk során megfigyeltük, hogy a guaninban és citozinban meglehetősen gazdag génrégióban a heterozigóta (pl. 4×/7×) genotípusok kimutatása a rövidebb allél preferenciális képződése („allél-kiesés”) miatt kihívást jelent, ami a hosszú allélok gyakoriságának alulbecsléséhez vezethet. Ennek elkerülése céljából a PCR-elegyben a dGTP-t részlegesen (50%-ban) dezoxi-inozintrifoszfáttal helyettesítettük, melyben a hipoxantin a guaninhoz hasonlóan a citozinnal képez bázispárt, de a molekulák között csak kettő H-kötés jön létre. A képződő eltérő nagyságú PCR-termékek olvadáspontja között így kisebb a különbség, ami hozzájárul a hosszabb fragmentum hatékonyabb képződéséhez [XVII].

A gén 5' régiójában elhelyezkedő SNP-ket (rs1800955 „-521 C/T”, rs747302 „-616 C/G”) korábban allél-specifikus PCR és RFLP technikákkal elemeztük [XVIII]. Felhívtuk a figyelmet arra, hogy az rs747302 SNP-re leírt *Ava* II restrikciós endonukleáz alkalmazásán alapuló genotipizáló eljárás [230] figyelmen kívül hagyja a lókuszt közvetlen szomszédságában elhelyezkedő rs796487306 „-615 A/G” polimorfizmust, és emiatt az rs747302 G allélt esetenként tévesen C variánsként azonosítja [XIX]. Az SNP-k elemzése során észrevettük, hogy néhány mintában a régió PCR-amplifikációja során a vártnál rövidebb termék képződik.



10. ábra. A humán *DRD4* gén 5' régiójában lévő 27 bp-os deléció. A 3. minta heterozigóta genotípusú. A szekvencia összehasonlítás mutatja a deléció GRCh38.p13 szerinti pontos elhelyezkedését.

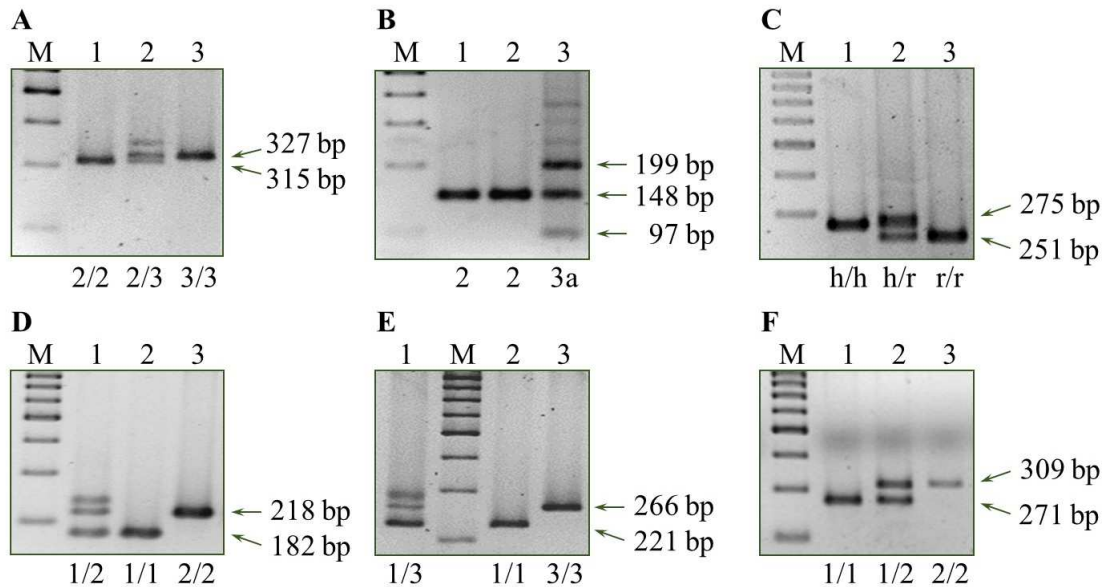
Az érintett szakasz két oldalára tervezett primerpár (*Függelék S1. táblázat*) alkalmazásával egy hosszúságvariáció rajzolódott ki, a szekvenálás eredménye alapján a deléció 27 bp hosszúságú szekvenciát foglal magában (*10. ábra*). Nem-rokon személyekből álló populáción ($N = 955$) vizsgáltuk az új allél előfordulási gyakoriságát, a deléció frekvenciája 0,16%-nak mutatkozott. Az egyik heterozigóta személy szüleitől származó DNS-minták is elérhetőek voltak, a 27 bp-os szekvenciakiesést egyikükben szintén sikerült kimutatni [II].

Ezek az eredmények mutatják, hogy a Humán Genom Projekt befejezését követően is szükséges, illetve lehetséges volt a DNS-szekvencia pontosabb megismerése és polimorfizmusainak feltárása. Más fajok esetén még kevesebb információ állt rendelkezésre. A humán *DRD4* gén polimorfizmusainak elemzéséből kiindulva vizsgálatainkat a kutya dopaminerg rendszerének variációira is kiterjesztettük, tekintve, hogy a kutya bizonyos (pszichológiai) fenotípusok szempontjából jobb modell a gyakran alkalmazott rágcsálóknál, és érdekes módon a *DRD4* génben lévő polimorfizmusai is több hasonlóságot mutatnak az emberi variánsokkal. *In silico* szekvencia analízissel (Tandem Repeats Finder) ismétlődő szekvenciákat kerestünk a dopamin- β -hidroxiláz (*DBH*), a dopamin transzporter (*DAT*), a tirozin-hidroxiláz (*TH*) és D4-es dopamin receptor (*DRD4*) génekben, a további vizsgálatokba azokat a régiókat vontuk be, ahol az ismétlődő modul hossza meghaladta a 6 bp-t, és a repetitív szakaszok bázissorrendje legalább 85%-os egyezést mutatott. Ezen kritériumok alapján öt korábban nem elemzett régió (*TH* 4. intron 36 bp duplikáció, *DBH* 4. intron 17 bp duplikáció, *DBH* 11. exon 24 bp duplikáció, *DAT* 5. intron 59 bp duplikáció és *DAT* 9. intron 38 bp duplikáció) elővizsgálatát kezdtük meg 60 (különböző fajtájú kutya és farkas) egyed bevonásával.

Az ismétlődő szekvenciákat közrefogó primerpárokat (*Függelék S1. táblázat*) terveztünk, az ezekkel kapott PCR-termékeket hagyományos alámerülő agaróz gélelektroforézissel elemeztük (*11. ábra*). A *DAT* 9. intronjában, valamint a *DBH* és a *TH* 4. intronjában lévő ismétlődő szakaszok polimorfnek bizonyultak a vizsgált populációkban, ezeket a variációkat és a *DRD4* gén 1. és 3. exonjában korábban leírt populációkat elemeztük tovább a teljes, 495 állatból (tervueren, groenandael és malinois belga juhászkutyák, németjuhászok és farkasok) álló mintában.

A *DRD4* gén 3. exonjában lévő polimorfizmus pontos elemzéséhez az eltérő modulokból álló szerkezet (*4. ábra*) miatt két reakció szükséges. Az első PCR során meghatározható az ismétlődő régió hossza, a „3a” és „3b” allélok elkülönítése külön mérésben történik (*Függelék*

SI. táblázat), mivel ezen két variáns esetén csak az ismétlődő modulok sorrendje különbözik (11. ábra). Ez a reakció azonban önmagában nem alkalmas a genotípus meghatározására, csak az első PCR eredményének ismeretében értelmezhető.



11. ábra. Kutya dopaminerg rendszer génjeiben lévő VNTR-ek genotipizálása. Az ismétlődési szám meghatározása a polimorf régió PCR-amplifikációját követő gélelektroforézissel történt. *A*: *DRD4* 3. exon VNTR hosszának elemzése, illetve *B*: a „3a” és „3b” allélok elkülönítése. *C*: A *DRD4* 1. exon 24 bp inzerció / delécio (h = hosszú, r = rövid allél), *D*: *TH* 4. intron, *E*: *DBH* 4. intron, *F*: *DAT* 9. intron VNTR vizsgálata. A heterozigóta genotípusú mintáknál látható legkisebb mobilitású fragmentum a heteroduplex.

A vizsgált populációban a „3b” allél egyetlen egyedben sem volt kimutatható. Németjuhászokban a „2” allél (64,6%), tervuerenekben és malinoisokban a „3a” allél (62,5%, illetve 64,0%) gyakorisága volt magasabb, míg groenandaelekben a két variáns frekvenciája közel azonos volt („2” allél: 51,4%). Kutyaiban más ismétlődés nem fordult elő, ezzel szemben farkasokban az 5 modulból álló allél volt legnagyobb gyakorisággal (43,3%) kimutatható, és a „2”, „3a” és „5” variánsok mellett egy korábban nem azonosított változat is jelen volt, melyet – hosszának megfelelően – „8” allélnak neveztünk el. Az allélgyakoriság értékek hasonló markáns különbséget mutattak az 1. exonban lévő 24 bp-os inzerció / delécio esetén (11. ábra).

Farkasokban kizárólag a hosszú változat volt jelen, és a rövid allél alig fordult elő (1,7%) németjuhászokban is. A tervuerenekben viszont a deléciót hordozó forma bizonyult gyakoribbnak (57,0%) [III].

1. táblázat. A dopaminerg rendszer génjeiben azonosított új polimorfizmusok allélfrekvencia értékei. A táblázat mindhárom polimorfizmus esetén a rövid („1”) változat gyakoriságát mutatja.

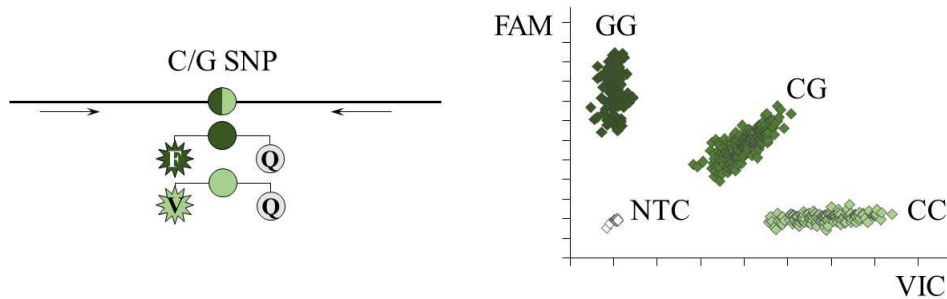
VNTR	Tervueren	Groenandael	Malinois	Németjuhász	Farkas
TH	40,0%	72,5%	9,0%	11,5%	27,3%
DBH	88,7%	77,9%	76,0%	96,9%	91,3%
DAT	5,9%	16,82%	56,25%	0%	17,4%

A TH, DBH és DAT génekben azonosított új polimorfizmusok esetén a kapott genotípus gyakoriság adatok nem mutattak szignifikáns eltérést a Hardy–Weinberg egyensúly alapján számított elméleti értékektől, ami az alkalmazott genotipizáló technikák megfelelő működését is mutatja. A TH génben a 36, a DAT-ban pedig a 38 bp hosszúságú ismétlődő szekvencia egy, illetve két példányban volt jelen, az allélgyakorisági értékek a DRD4 két polimorfizmusához hasonlóan az egyes fajtákban jelentős eltérést mutattak (1. táblázat). A DBH génben lévő 17 bp-os szakasz az eredményeink alapján – érdekes módon – vagy 1, vagy 3 példányban volt jelen. Ezen polimorfizmus esetén a ritka változat egységesen a rövid („1”) allél volt: itt figyeltük meg a legkisebb eltérést az egyes fajták allélgyakoriság értéke (3,1–24%) között [III]. Már ilyen kisszámú polimorfizmus vizsgálata esetén is szembeűnő, hogy a különböző kutyafajták genetikai sajátossága között jelentős eltérés mutatkozik, érdekes, hogy még a malinois fajta is lényegesen különbözik a belga juhászkutyák másik két csoportjától. Adataink alapján megfigyelhető az is, hogy esetenként a kutyafajták között nagyobb különbség tapasztalható, mint a kutyák és a farkasok között, ami részben azzal magyarázható, hogy a két faj genetikai állománya kb. 100 000 évvel ezelőtti szétválásukat követően is keveredhetett egymással [231].

5.1.1.2 A WFS1 gén polimorfizmusainak vizsgálata

A humán WFS1 génben 17 SNP-t real-time PCR módszerrel, TaqMan próbák alkalmazásával genotipizáltunk (12. ábra). Munkánk során az eljárás „hagyományos” technikája [IV] mellett a

miniatürizált, OpenArray[®] rendszert is alkalmaztuk [V], mellyel egyetlen lemezen 33 nl térfogatú reakciókban egyidejűleg 3072 párhuzamos mérés valósítható meg.

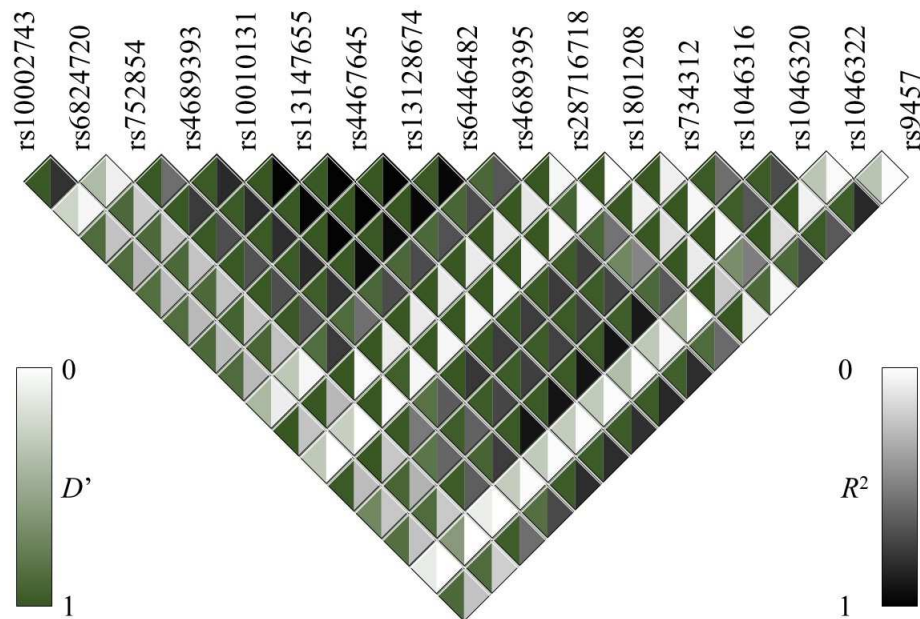


12. ábra. SNP-genotipizálás real-time PCR-rel, TaqMan próbák alkalmazásával. Az ábra a módszer elvét és a *WFS1* génben lévő rs9457 C/G SNP genotipizálásának eredményét mutatja ($N = 518$). F: FAM, V: VIC, Q: quencher, NTC: genomi DNS nélküli negatív kontroll.

A vizsgálatainkba bevinni kívánt polimorfizmusokat több szempont figyelembevételével választottuk ki. Olyan polimorfizmusokat kívántunk elemezni, melyek ritka allél gyakorisága eléri az 5%-ot. A gén minél átfogóbb elemzése céljából az 1000 Genom Projekt adatbázisából letöltöttük az elérhető genotípus adatokat, és kiszámoltuk a SNP-k közötti LD értékeket. Minden haplotípus blokkból legalább két SNP-t választottunk úgy, hogy az SNP-k fizikai távolsága is lehetőleg kiegyensúlyozott legyen (a kiválasztott lókuszok közötti távolság nem haladta meg a 6 kb-t). A 3' UTR-ben elhelyezkedő 3 SNP közül kettő (rs9457 és rs1046322) a Patrocles, illetve a PolymiRTS adatbázisok alapján miRNS-ek (miR-185 és miR-668) kötőhelyében található [IV, V]. Korábbi tanulmányok a harmadik (rs1046320) polimorfizmust a 2-es típusú cukorbetegség rizikófaktoraként azonosították, az SNP molekuláris funkcióját azonban *in vitro* elemzésekkel nem sikerült igazolni [94]. A 4. intronban elhelyezkedő rs10010131 SNP szerepe valószínűsíthető a 2-es típusú diabetes, illetve a betegséggel kapcsolatos fenotípusok örökletes hátterében [232], a 8. exonban lévő, aminosavcserét (p.H611R) okozó rs734312 polimorfizmust pedig GWAS során a kórkép genetikai rizikófaktoraként azonosították [233], és szerepe felmerült a hangulatzavarok örökletes komponenseként is [234].

Az SNP-k közötti LD mértékét 801 személy genotipizálását követően a Lewontin-féle D' , valamint az R^2 értékek (13. ábra) meghatározásával elemeztük [V]. A gén 3. intronjától a 7.

intronjaig kiterjedő több mint 13,5 kb hosszúságú szakaszon lévő SNP-k (rs4689393-től rs6446482-ig) egyetlen haplotípus blokkot alkotnak, ebben a régióban magas D' és R^2 értékeket kaptunk. Érdekes, hogy a csupán 292 bp távolságra, szintén a 7. intronban lévő rs4689395 polimorfizmus már nem tartozik bele ebbe a magas LD-vel jellemezhető szakaszba.

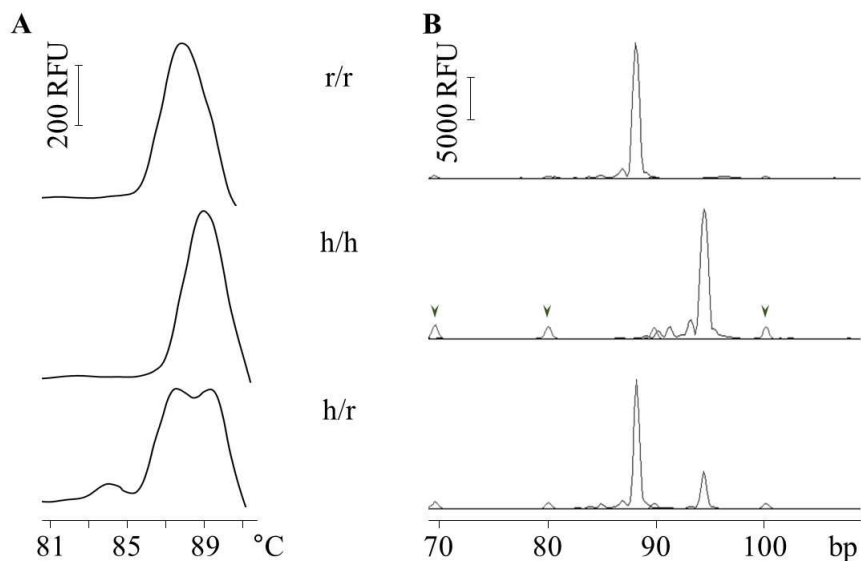


13. ábra. A *WFS1* gén 17 SNP-jének kapcsoltsági analízise. A négyzetek bal oldala (zöld skála) a Lewontin-féle D' , jobb fele (szürke skála) pedig az R^2 értékeket ábrázolja.

Szembetűnő, hogy az rs28716718 és az rs1801208 SNP-k esetén a D' és az R^2 között jelentős eltérés mutatkozik. A magas D' -höz társuló alacsony R^2 értékek abban az esetben jellemzőek, amikor a vizsgált populációban az SNP-k lehetséges négy kombinációjából csak három fordul elő, a negyedik haplotípus gyakorisága az allél-frekvenciák alapján számított értékénél lényegesen alacsonyabb. Figyelemre méltó, hogy a korábban a 2-es típusú cukorbetegség rizikófaktoraként felmerülő polimorfizmusok (rs1046320, rs10010131 és rs734312) és a miRNS-kötődést feltételezeten befolyásoló rs9457 SNP között magas LD értékeket kaptunk, ugyanakkor a másik vélhetően miRNS-kötőhelyen lévő SNP (rs1046322) független a gén csaknem valamennyi általunk elemzett polimorfizmusától [V].

Vizsgálatainkat a *WFS1* gén 5' nem kódoló régiójára is kiterjesztettük [VI]. Korábbi tanulmányok tanúsága szerint az rs4689388 és az rs4273545 SNP hatással van a génkifejeződés

szabályozására [235], és az előbbi polimorfizmust egy GWAS a 2-es típusú cukorbetegség genetikai rizikófaktoraként azonosította [93]. Kísérleteinkbe bevontuk a korábban nem vizsgált rs148797429 inzerció / deléció polimorfizmust is, mely a transzkripció starthelytől csupán 259 bp távolságra helyezkedik el, és *in silico* szekveneciaelemzés alapján hozzájárulhat a transzkripció aktivitásának szabályozásához SP1, valamint AP2 α transzkripció faktorok bekötődésének befolyásolása révén.



14. ábra. A *WFS1* rs148797429 inzerció / deléció polimorfizmus genotipizálása. A: olvadáspont analízis real-time PCR-rel, **B:** Cy5-jelölt PCR-termékek elemzése denaturáló kapilláris gélelektroforézissel. A nyílhegyek a molekulatömeg marker DNS-fragmentumait mutatják. Genotípusok: r/r: homozigóta deléció, h/h: homozigóta inzerció, h/r: heterozigóta.

A polimorfizmus egy 6 bp hosszúságú szakaszt érint (GGGGCG / –), melynek kimutatása hagyományos alámerülő gélelektroforézis segítségével technikai nehézséget jelent. A genotípus meghatározására ezért két független detektálási módszert állítottunk be (14. ábra). A polimorf régiót mindkét esetben PCR-rel amplifikáltuk (Függelék S1. táblázat), a 6 bp-os szakasz jelenlétekor egy 97, deléció esetén pedig egy 91 bp nagyságú fragmentum keletkezett. Az egyik eljárás során a PCR-termékhez SYBR Green DNS-interkalátor festéket adtunk, és real-time PCR segítségével meghatároztuk az ampikonok olvadáspontját. Bár ezzel a technikával a hosszú és a rövid változat hatékonyan elkülöníthetőnek bizonyult, megjegyzendő, hogy a PCR-

termékek olvadáspontját a felsokszorozott régióban lévő további polimorfizmusok is befolyásolhatják, ami az eljárás megbízhatóságát némileg csökkenti. A 6 bp-os szakasz jelenlétének kimutatására a denaturáló kapilláris gélelektroforézis hatékony és megbízható technikának bizonyult. Ezen módszer során Cy5-tel jelölt forward primert alkalmaztunk. A többcsatornás detektálás lehetővé tette, hogy a vizsgálandó mintához közvetlenül hozzáadjuk a Wellred D1 (Beckman) fluoreszcens festékkel jelölt molekulatömeg markert is, melynek szimultán elválasztása igen pontos fragmentumméret-meghatározást tett lehetővé [VI].

5.1.2 CNV-k genotipizálása

Bár a CNV-k az ismétlődési variációk közé sorolhatók, elemzésük a VNTR-ekhez hasonló módon a régió PCR-amplifikációját követő gélektroforézissel nagy kiterjedésük miatt nem valószínűsíthető meg. Az ismétlődési szám meghatározása felfogható viszont mennyiségi elemzésként, melynek a real-time PCR alapú eljárás hatékony eszköze lehet.

5.1.2.1 A *C4A* és *C4B* génszám elemzése

A 6. kromoszóma rövid karján elhelyezkedő CNV elméleti, klinikai és technikai szempontból is érdeklődésre tart számot. Itt helyezkedik el a 4 gént magában foglaló „RCCX”-modul (5. ábra), melynek egyik alkotója a C4 komplement komponens, ez a fehérje az adaptív immunválasz nélkülözhetetlen résztvevője. A változó (mindkét kromoszómán 1, 2 vagy 3, nagyon ritkán 4) génszám mellett érdekes, hogy a modul többi tagjával szemben a *C4* gén minden egységben aktív, ráadásul két, csupán 5 nukleotidban eltérő izoformája – a *C4A* és a *C4B* gén – lehet jelen, melyek valamelyest eltérő működésű fehérjét kódolnak. Annak ellenére, hogy a *C4A* és *C4B* génszám klinikai szempontból is jelentős, meghatározására csupán meglehetősen idő- és munkaigényes technikák álltak korábban rendelkezésre. A *C4A*, illetve *C4B* teljes hiánya az adott izoformára specifikus PCR-rel [236], a két variáns aránya, illetve ismétlődési száma *PshA* I és *Taq* I RFLP-vel [237], valamint Southern-blot technikával volt megállapítható [238]. Az általunk beállított real-time PCR-alapú eljárás segítségével a *C4A*- és *C4B*-génszám hatékonyabban és pontosabban meghatározható [VII].

Az eljárás a TaqMan próbákkal történő SNP genotipizálás és a kvantitatív analízis kombinációjának tekinthető. A *C4A* és *C4B* gének közötti összesen 5 nukleotidnyi szekvencia eltérés

egy 17 bp-os régióban helyezkedik el, a próbákat úgy terveztük meg, hogy ezen génspecifikus nukleotidokból négyet magukban foglaljanak (15. ábra). A relatív kvantifikálás során belső kontrollként az *RNáz P* génszámhoz viszonyítottunk, a hatékonyabb és megbízhatóbb mérés céljából a *C4A*, illetve a *C4B* és az *RNáz P* génszámot különböző fluoreszcens festékekkel jelölt próbák segítségével azonos reakcióelegyben határoztuk meg. Az 1. reakcióelegyben történt a *C4A* mennyiségi meghatározása VIC-kel jelölt génspecifikus és FAM-mal jelölt *RNáz P*-specifikus próbával, a 2. reakcióelegyben a *C4B*-hez kötődő próba volt FAM-mal, az *RNáz P*-hez kötődő pedig VIC-kel jelölve. A „fordított” jelölést azért vezettük be, mert így a *C4A*- és *C4B*-specifikus próbák egy reakcióelegyben együtt (*RNáz P* mérése nélkül) is alkalmazhatóak. Így egy minta elemzéséhez feleannyi reakció szükséges, ebben a rendszerben viszont a pontos génszám helyett csak a két izoforma aránya, illetve valamelyik variáns hiánya állapítható meg egyértelműen.

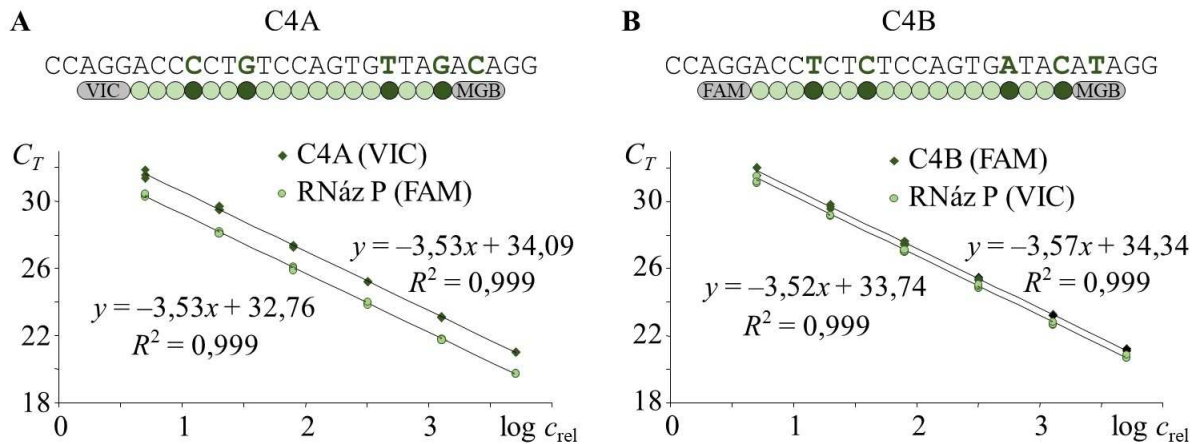
A génszám (n_{C4A} és n_{C4B}) meghatározása a ΔC_T -módszer alkalmazásán alapult ($C_{T\text{ RPF}}$ és $C_{T\text{ RPV}}$ az 1., illetve a 2. reakcióelegyben mért *RNáz P*-hez tartozó C_T értékek):

$$n_{C4A} = \frac{2^{C_{T\text{ RPF}} - C_{T\text{ C4A}} + 1}}{q_{C4A}} \qquad n_{C4B} = \frac{2^{C_{T\text{ RPV}} - C_{T\text{ C4B}} + 1}}{q_{C4B}}$$

A tényleges génszám meghatározása szükségessé tette egy korrekciós faktor (q_{C4A} és q_{C4B}) bevezetését, mivel – annak ellenére, hogy a PCR-ek viszonylag széles koncentráció tartományban megfelelő pontossággal működtek (15. ábra) – az eltérő primerek és jelölés miatt azonos *C4*, illetve *RNáz P* génszám nem teljesen megegyező C_T -értéket eredményezett. A korrekciós faktorokat 65 személy (a 2 reakcióelegynek és 3 párhuzamos elemzésnek megfelelően $65 \cdot 6 = 390$ mérési pont) adatai alapján a következő szempontokat figyelembe véve határoztuk meg: a *C4A* és *C4B* génszám csak egész szám lehet, melyek összege csak kivételes esetben lehet 6-nál nagyobb, és a génszám nem lehet 0, ha a real-time PCR során a megfelelő reakcióban DNS-amplifikáció tapasztalható. Ennek figyelembe vételével a génszámok meghatározása során a $q_{C4A} = 0,39 \pm 0,04$, valamint a $q_{C4B} = 0,75 \pm 0,11$ bizonyult eredményesnek [VII].

A mennyiségi mérés megbízhatóságát egy 2 *C4A* és 2 *C4B* génnel rendelkező személy genomi DNS-éből készített hígítási sor (0,3–300 ng, 3 párhuzamos mérési pont) elemzésével ellenőriztük. A 15. ábrán látható, hogy a kalibrációs egyenesek meredeksége közel azonos, ami

a reakciók megfelelő mértékben megegyező hatékonyságát tükrözi, s a minta-DNS koncentráció logaritmusa és a mérhető C_T -értékek között csaknem tökéletes lineáris összefüggés figyelhető meg.

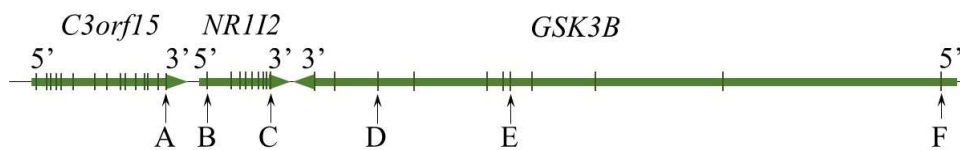


15. ábra. A C4A és C4B génszám meghatározása real-time PCR-rel. Az ábra felső része a C4A és C4B közötti szekvencia különbséget (vastag, zöld betűk, illetve sötétzöld körök) és az alkalmazott TaqMan próbákat mutatja. A negyedelő hígítási sor alkalmazásával készített kalibrációs egyenesek szemléltetik, hogy a C4A (A) és C4B (B) génszám meghatározása 3 nagyságrendnyi koncentráció-tartományban pontosan és megbízhatóan megvalósítható.

A beállított technikával 118 személy mintáját elemeztük, és az egyes C4A, illetve C4B génszám gyakorisági értékeket Southern-blot alapú mérés eredményével vetettük össze [239]. Bár az egyes csoportok közötti eltérés mértéke statisztikailag nem bizonyult szignifikánsnak, az eredmények mégsem egyeztek teljesen, és megfigyelhető volt, hogy a Southern-blot alapú technika szerint a 3 C4A, illetve C4B gén előfordulása számottevően ritkább. Az ellentmondás feloldása érdekében 33 mintát egy független, eltérő PCR rendszeren és azt követő kapilláris gélelektroforézissel történő kvantitatív elemzésen alapuló eljárással is megvizsgáltunk [XX]. Három minta esetén tapasztaltunk eltérést: a génszám elemzését mindkét technikával újra elvégeztük. Az ismétlés mindhárom esetben a real-time PCR-rel kapott eredményt erősítette meg.

5.1.2.2 A glikogénszintáz-kináz 3 β gént érintő CNV vizsgálata

Hasonló, real-time PCR alapú eljárást állítottunk be a glikogénszintáz-kináz 3 β (*GSK3B*) gént érintő CNV elemzésére. Noha a gén által kódolt szerin–treonin-kináz az enzimes család többi tagjával együtt számos alapvető anyagcsere és jelátviteli út résztvevője, a *GSK3B* fokozott aktivitása főképp neuropszichiátriai rendellenességekkel (szkizofrénia [240], Alzheimer-kór [241], hangulatzavar, ADHD [242]) hozható összefüggésbe, és ebből adódóan merül fel a gént érintő CNV szerepe ezen kórképek genetikai hátterében.



16. ábra. A *GSK3B*-t érintő CNV elemzése. A vastag zöld nyilak a *GSK3B*, az *NR112* és a *C3orf15* gének elhelyezkedését mutatják. A függőleges vonalak az exonokat, az A–F betűk a real-time PCR-rel elemzett régiókat jelölik.

A génszámvariáció elemzését első lépésben a *GSK3B* gén 9. exonjára tervezett TaqMan rendszer segítségével (16. ábra: „D” régió) kezdtük meg 170 major depresszióban, 106 bipoláris hangulatzavarban szenvedő beteg és 321 pszichiátriai rendellenességgel nem diagnosztizált kontroll személy bevonásával. Eredményeink alapján ebben a régióban a CNV előfordulása ritka: 5 személy (3 major depresszióban, 1 bipoláris hangulatzavarban szenvedő beteg és egy kontroll) esetén 3, egy major depresszióval diagnosztizált páciensnél pedig 4 kópiát azonosítottunk, ennek megfelelően az amplifikáció (2-nél nagyobb génszám) gyakorisága 1%-nak bizonyult. A régiót érintő deléció egyetlen esetben sem találtunk. A CNV kiterjedésének elemzése céljából a *GSK3B* gén 5. exonjára és 5' szakaszára (16. ábra: „E” és „F”), valamint a szomszédos *NR112* és *C3orf15* génekre (16. ábra: „A”, „B” és „C”) további TaqMan rendszereket terveztünk (Függelék S1. táblázat), illetve vásároltunk a Thermo Fisher Scientific kínálatából. A CNV kiterjedése a 6 mintában nem bizonyult azonosnak. A *C3orf15* 3' és az *NR112* 5' régióján a kópiaszám minden esetben 2 volt, azaz az amplifikáció erre a szakaszra (16. ábra: „A”, „B”) nem terjedt ki. A megemelkedett génszám az *NR112* 3' végén és a *GSK3B* 5. exonjában mind a hat esetben kimutatható volt (16. ábra: „C”, „E”), ugyanakkor az amplifikáció

csupán két minta esetén foglalta magában a teljes *GSK3B* gént kiterjedve annak 5' régiójára (16. ábra: „F”) is [VIII].

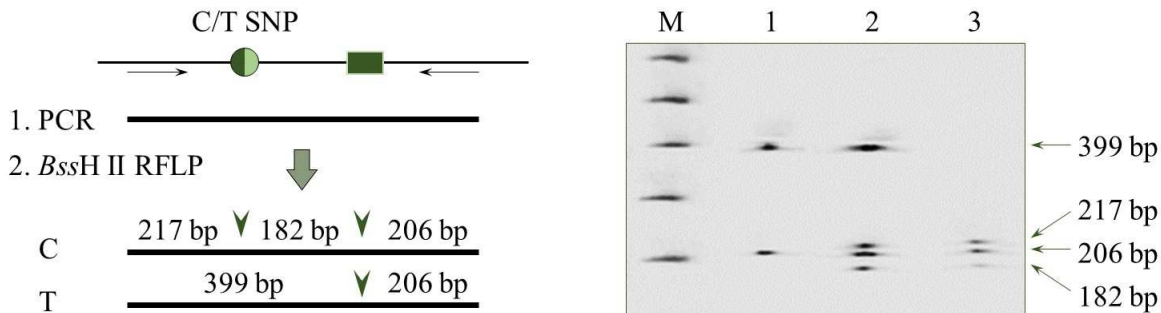
5.1.3 Hatékony elektroforetikus eljárások alkalmazása

A hagyományos PCR alkalmazásán alapuló genotipizáló módszerek befejező lépése a keletkezett termékek gélelektroforézissel történő elválasztása és láthatóvá tétele. Ez a lépés bár sok esetben hagyományos, alámerülő agaróz gélelektroforézissel is megvalósítható, a technika viszonylag munka- és időigényes, kevésbé automatizálható, így nem ritkán a genotipizáló technikák szűk keresztmetszetét jelenti.

5.1.3.1 Ultravékony gélelektroforézis

Az ultravékony gélelektroforézis eljárást eredményesen alkalmaztuk számos genotipizáló módszer befejező lépéseként: a technika a hagyományos, horizontális agaróz gélelektroforézis és a kapilláris elektroforézis számos előnyös tulajdonságát egyesíti. Az üvegekazettában lévő 190 μm vastagságú agaróz–lineáris poliakrilamid gél keverék jó felbontóképességet, a lézer indukálta fluoreszcencia elvén alapuló detektálás a kapilláris elektroforézishez hasonló érzékeny analízist biztosít. A rendszer ugyanakkor lehetővé teszi a hagyományos elektroforézishez hasonlóan az alkalmazott üvegekazetta méretétől függően akár 96 minta párhuzamos elemzését is. A minták egy fésű alakú membránról jutnak a gélbe az elektroforézis során az elektromos tér erő hatására, az elektroforézis időtartama kb. 5 perc. A rendszer a gyors és érzékeny DNS-fragmentum analízis mellett a többlépéses genotipizáló technika automatizálását, hatékonyságának növelését is lehetővé tette. A *DRD4* gén rs1800955 „-521 C/T” SNP-jének PCR–RFLP-vel történő genotipizálása során a restriktív emésztés és a gélelektroforézis egyetlen lépésben megvalósíthatóvá vált. A PCR-t (Függelék S1. táblázat) követően a *BssH* II restriktív endonukleázzal történő emésztést 0,4 μl végtérfogatban a mintafelvitelre használt membrán fogain valósítottuk meg. A reakció – részben a magas *BssH* II koncentráció miatt – 15 perc alatt végbement. A PCR-termék egy nem polimorf, kontroll felismerő helyet is tartalmazott, T allél esetén kettő, C allél jelenlétekor pedig három hasított termék keletkezett, így az emésztés hatékonyságát minden reakcióban ellenőrizni tudtuk. Az ultravékony elektroforézis segítségével az

egymástól csupán 11 bázispár különbséget mutató termékek is egyértelműen elkülöníthetők voltak (17. ábra) [IX].



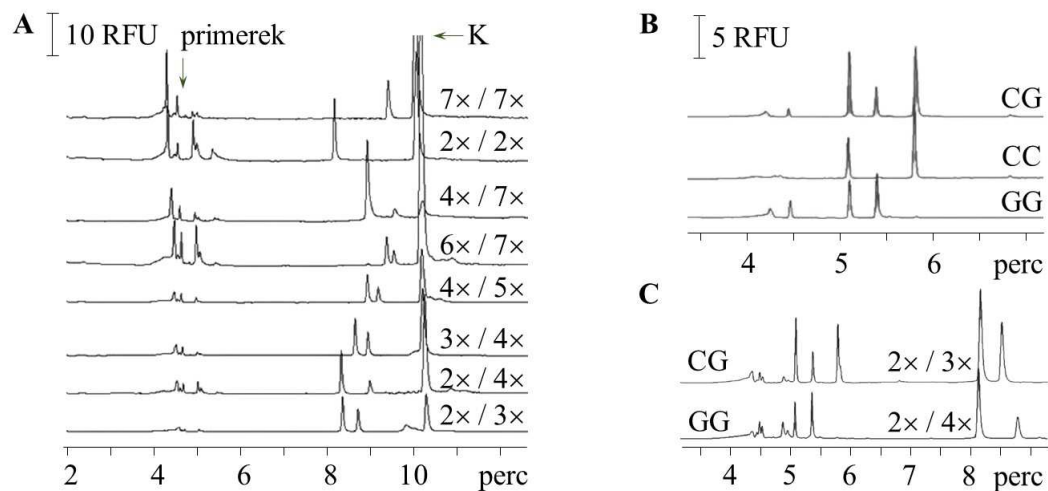
17. ábra. A *DRD4* gén rs1800955 „-521 C/T” SNP PCR–RFLP analízise ultravékony gélelektroforézissel. A PCR-termék egy kontroll (zöld téglalap) *BssH* II hasító helyet tartalmazott, az enzim a PCR-terméket C allél jelenléte esetén az SNP-nél is elhasította (nyílhegyek). Genotípusok: 1: TT, 2: CT, 3: CC

5.1.3.2 Kapilláris gélelektroforézis

A *DRD4* gén polimorfizmusainak elemzése során a keletkező PCR-termékek elemzésére kapilláris gélelektroforézist is alkalmaztunk. A 3. exonban lévő 48 bp-os VNTR ismétlődési száma (n) leggyakrabban 2 és 7 között változik, a régió PCR-rel történő amplifikálása (Függelék S1. táblázat) során ennek megfelelően egy $283 + 48n$ nagyságú termék keletkezett. A gén promóterében lévő rs1800955 „-521 C/T” és rs747302 „-616 C/G” SNP-k haplotípusának elemzésére kombinált allél-specifikus PCR–RFLP eljárást alkalmaztunk, melynek során az rs1800955 SNP genotípusától függő módon képződő PCR-termékben (Függelék S1. táblázat) az rs747302 polimorfizmust *Sau96* I endonukleáz segítségével elemeztük. A 235 bp hosszúságú PCR-termék egy kontroll hasítóhelyet is tartalmazott: rs747302 C allél esetén csak itt (87 bp és 148 bp), G allél jelenlétekor pedig az SNP-nél is hasított az enzim (87 bp, 35 bp és 113 bp). Mivel a VNTR és az SNP-k elemzésére alkalmazható PCR-rendszereket úgy terveztük meg, hogy a keletkező fragmentum mérettartománya ne fedjen át, így – a hatékony elektroforetikus elemzés céljából – lehetőség kínálkozott arra is, hogy a keletkező termékek elválasztását közösen, egy lépésben végezzük el (18. ábra) [I].

A 11. kromoszóma rövid karján, a *DRD4* gén közelében található a D11S1984 mikroszatellita. Asszociáció elemzések eredményéből kiindulva felvetődött annak lehetősége, hogy ez a

4 bp-os ismétlődési polimorfizmus a *DRD4* gén variációinak markereként szolgálhat [243]. A mikroszatellita elemzését denaturáló kapilláris gélelektroforézissel végeztük el, és egy 106 magyar személyből álló populáción meghatároztuk a D11S1984 marker és a *DRD4* gén 5 polimorfizmusa (120 bp-os duplikáció, rs747302 „-616 C/G”, rs796487306 „-615 A/G”, rs1800955 „-521 C/T”, 3. exon 48 bp VNTR) között mérhető LD értéket. Eredményeink a korábbi feltételezést nem támasztották alá: a mikroszatellita és a *DRD4* polimorfizmusok között minden esetben 0,375-nél alacsonyabb Lewontin-féle D' értéket kaptunk [XVI].

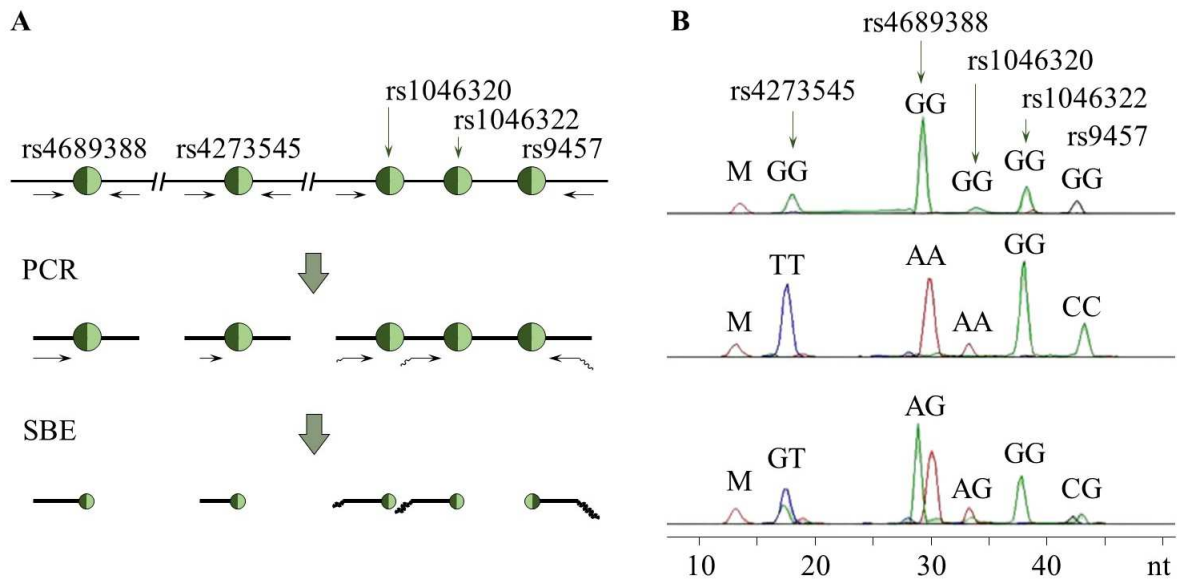


18. ábra. A *DRD4* gén 3. exon VNTR és rs747302 „-616 C/G” SNP analízise kapilláris gélelektroforézissel. *A*: A VNTR analízise során az ismétlődési számnak megfelelően $283 + 48n$ nagyságú termék keletkezett. *K*: az 1066 bp-os kontroll fragmentum az elektroferogramok egymáshoz illesztését tette lehetővé. *B*: az rs747302 SNP *Sau96* I RFLP-termékeinek elválasztása. *C*: a két polimorfizmus együttes elemzése. Az ábra az elválasztások mellett a minták genotípusát is mutatja.

A gyors, hatékony, érzékeny elemzés, valamint a nagy felbontóképesség mellett a kapilláris elektroforézis további előnye a különböző színű fluoreszcens festékek egyidejű detektálásának lehetősége. Ezt használtuk ki a *WFS1* gén promoterében (rs4689388 és rs4273545), illetve 3' UTR-ében (rs1046320, rs1046322 és rs9457) elhelyezkedő 5 SNP multiplex primerextenzióval történő genotipizálása során [VI].

A módszer első lépése a polimorfizmusokat tartalmazó régió PCR amplifikációja. Mivel a promoterben lévő SNP-k egymástól 1360 bp távolságra helyezkednek el, célravezetőbbnek tűnt

a két polimorfizmusnak megfelelően két külön PCR alkalmazása. A 3' UTR-ben lévő három lókuszt ugyanakkor egy mindössze 455 bp hosszúságú génszakaszon belül helyezkedik el, így megvalósítható volt ezek egyetlen közös PCR-rel történő felsokszorozása. A három PCR primereit úgy terveztük meg, hogy a fragmentumok egyszerre, egyetlen közös reakcióban létrehozhatók legyenek.



19. ábra. A *WFS1* génben lévő 5 SNP genotipizálása multiplex primerextenzióval. *A:* A módszer elve és általunk alkalmazott elrendezése. *B:* 3 minta elemzésének elektroferogramja. M: az elválasztások illesztésére szolgáló kontroll fragmentum. Az ábra az elválasztások fölött a minták genotípusát is mutatja.

A megfelelő tisztítási lépéseket követően a technika elvét az SBE reakció jelenti, melynek lényege az SNP-k melletti nukleotidhoz hibridizáló SBE primerek és a négy különböző fluoreszcens festékkel jelölt, láncterminátor, aciklo-nukleotidok alkalmazása (19. ábra). A genotípus leolvasása a jelen lévő allélnak megfelelő színű nukleotid detektálásán alapul, multiplex analízis pedig – a kapilláris elektroforézis nagy felbontású elválasztóképességét kihasználva – eltérő hosszúságú SBE primerek alkalmazásával valósítható meg. Az optimális reakció érdekében az SBE primer tetszőleges irányú lehet. A rendszert úgy kívántuk beállítani, hogy az első PCR-hez hasonlóan az 5 polimorfizmus SBE reakciója is egyetlen elegyben kivitelezhető legyen. Az SNP-eket övező szekvenciák figyelembevételével a promoterben lévő rs4689388 és

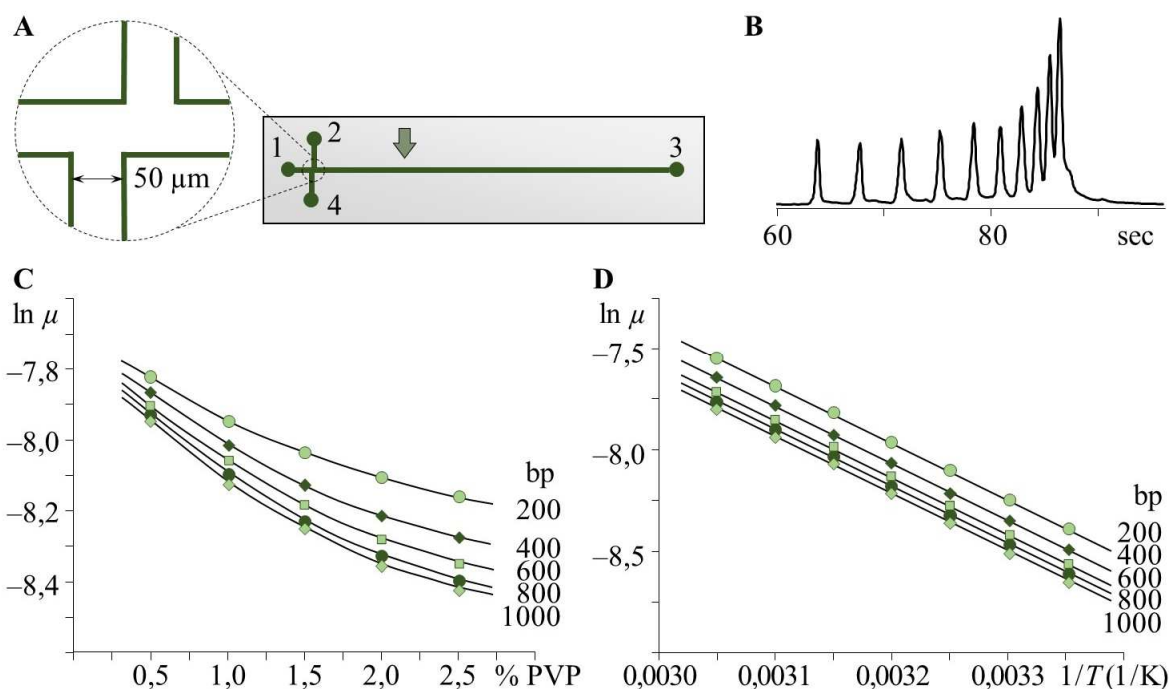
rs4273545 SNP-kre egy 30 és egy 22 nukleotidból álló, „forward” primert tudtunk tervezni. A 3' UTR polimorfizmusai esetén a nem kívánatos másodlagos struktúrák elkerülése érdekében az eltérő hosszúságú (rs1046320: 36 nt, rs1046322: 42 nt, rs9457: 48 nt) primereket úgy hoztuk létre, hogy azok a specifikus régió mellett az 5' végükön eltérő hosszúságú oligo-adenilát (A)_n szakaszt tartalmaztak. A leghatékonyabb működés érdekében az rs1046320 és az rs1046322 SNP-t „forward”, az rs9457 polimorfizmust viszont „reverse” irányú primerrel (*Függelék SI. táblázat*) genotipizáltuk az SBE reakció során (*19. ábra*). A denaturáló kapilláris gélelektroforézis segítségével a 6 nukleotid hosszúságkülönbséget mutató termékek egymástól megbízhatóan elkülöníthetőnek mutatkoztak [VI]. Az 55 nt feletti mérettartomány még további legalább 3 SNP egyidejű genotipizálásának lehetőségét kínálja, ami még abban az esetben is hatékony megoldást jelent, ha a megelőző lépések (1. PCR és SBE reakció) egy-egy multiplex reakcióban történő kivitelezése nem valósítható meg. Ennek megfelelően 96-os formátummal számolva felhasználói beavatkozás nélkül egyetlen mérés során a mi rendszerünkben 480 genotípus határozható meg, ami a további multiplexálási kapacitást kihasználva akár 768-ra is növelhető.

5.1.3.3 DNS-fragmentum analízis mikrofluidikai eszköz alkalmazásával

Noha az ultravékony és a kapilláris gélelektroforézis – jó elválasztóképességének, érzékenységeinek és automatizálhatóságának köszönhetően – a mi rendszereinkben a genotipizálás hatékony eszközének bizonyult, a mikrofluidikai eszközök alkalmazása újabb lehetőségeket kínál. A további miniatürizálás eredményeként az elválasztás időtartama akár 3 perc alá csökkenthető, és a chipen kialakítható csatornahálózatok segítségével az elektroforézist megelőző lépések integrációja is megvalósítható. Munkánk során a gélelektroforézis egyes paramétereit egy egyszerű felépítésű, két egymást keresztező csatornát tartalmazó mikrofluidikai eszköz felhasználásával jellemeztük. Az üvegből készült chipben az elválasztás 50 µm átmérőjű félkör keresztmetszetű csatornában történt, a 100 µm-es „dupla-T” injektor térfogata kb. 10⁻⁴ µl volt. Az elválasztó csatorna hossza az injektortól a 3-as tartályig $L = 80$ mm, az effektív hossz a detektorig $l = 30$ mm volt (*20. ábra*). Az elemzések során polivinilpirrolidon (PVP, $M_w = 1\,300\,000$ g/mol) elválasztó mátrixot alkalmaztunk, melynek segítségével az elektroosmotikus áramlás is hatékonyan minimalizálhatónak bizonyult. Ebben a rendszerben a 100 bp-os DNS létra (100

bp, 200 bp..., 1000 bp) fragmentumainak elválasztása kevesebb, mint 90 másodperc alatt kivitelezhető volt. Bár a 20. ábrán látható, hogy alapvonal-elválasztás csak 700 bp-ig valósul meg, ugyanakkor a rendszer még jelentős tartalékokkal rendelkezik: mindenekelőtt az effektív elválasztási hossz emelésével a felbontóképesség nagymértékben fokozható.

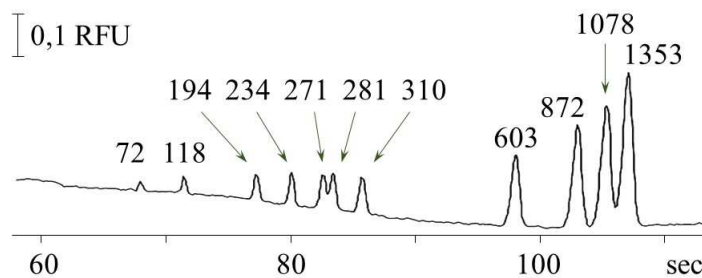
A Ferguson-görbe a PVP-koncentráció és az egyes fragmentumok elektroforetikus mobilitása ($\mu = \frac{v}{E}$; v : sebesség, E : elektromos térerő) közötti összefüggést ábrázolja (20. ábra). Minden adatpont 5 párhuzamos mérés eredményének átlagát mutatja, az ezek közötti 1% alatti eltérés az ábrákon nem jeleníthető meg. A görbék nem lineáris lefutása nem egyedi: korábban mind hagyományos, mind kapilláris poliakrilamid-gélelektroforézis során megfigyelték már [244]. A jelenség hátterében a PVP hígítás során végbemenő kismértékű konformáció változása állhat. Az Arrhenius-görbe az egyes fragmentumok mobilitásának (μ) változását mutatja a hőmérséklet hatására (20. ábra).



20. ábra. Gélelektroforézis mikrofluidikai eszköz alkalmazásával. *A:* Az elektroforetikus chip felépítése. A nyíl a detektálás helyét mutatja. *B:* Az elektroforézis paramétereinek elemzéséhez alkalmazott 100 bp-os DNS létra reprezentatív elválasztási képe. *C:* Ferguson-görbe: Az alkalmazott PVP-koncentráció és az elektroforetikus mobilitás (μ) összefüggése. *D:* Arrhenius-görbe: A hőmérséklet és az elektroforetikus mobilitás (μ) összefüggése.

A vártak megfelelően az elektroforetikus mobilitás a hőmérséklet növelésével emelkedik, az egyenesek párhuzamos lefutása alapján a 20–50 °C-os tartományban a felbontóképesség jelentős változást nem mutat. Megfigyeltük, hogy a Sytox Orange festék koncentrációjának emelése az elektroforetikus mobilitás kismértékű csökkenéséhez vezet, amit feltehetően az interkalátor pozitív töltése, illetve ennek megfelelően a DNS-sel alkotott komplex töltésének eltolódása okoz. Az általunk alkalmazott rendszerben – az elméletileg várttal szemben – az elektromos mező emelkedésével a fragmentumok mobilitásának (μ) növekedése volt tapasztalható. Ennek háttérében a DNS-molekulák magas (70–440 V/cm) elektromos mező hatására bekövetkező térszerkezetváltozása állhat [X].

Az elektroforézis és a megelőző lépések integrációjának modelljeként a kétszálúvá alakított Φ X 174 bakteriofág eredetű DNS-t az elektroforetikus chipen *Hae* III restriktív endonukleázzal emésztettük, majd a kapott fragmentumokat egymástól elválasztottuk. A reakcióelegyet a mikrofluidikai eszköz 2-es, a Φ X 174 DNS-t pedig a 4-es tartályba (20. ábra) mértük, majd a minta-DNS-t megfelelő elektromos mező ($U_1 = 240$ V, $U_2 = 480$ V, $U_3 = 240$ V, $U_4 = 0$ V) létrehozásával a reakcióeleggyel összekevertük. Az inkubációt követően az emésztett terméket a „dupla-T” injektorba juttattuk ($U_1 = 240$ V, $U_2 = 0$ V, $U_3 = 240$ V, $U_4 = 480$ V), majd egymástól elválasztottuk ($U_1 = 0$ V, $U_2 = 250$ V, $U_3 = 1650$ V, $U_4 = 250$ V). A jobb felbontóképesség elérése érdekében a fragmentumokat az injektortól $l = 50$ mm távolságban detektáltuk (21. ábra).



21. ábra. *Hae* III enzimmel emésztett Φ X 174 DNS elválasztása elektroforetikus chipen.

A csúcsok feletti számok a fragmentumok méretét (bp) mutatják.

5.1.4 Genetikai polimorfizmusok haplotípus analízise

A komplex jellegek háttérében szerepet játszó genetikai polimorfizmusok hatása nem additív, az egyes örökletes (és környezeti) faktorok egymással kölcsönhatásban (pl. episztázis) állnak: molekuláris szintű, valamint fenotípusos megnyilvánulásuk komplex hálózatos rendszer működés révén képzelhető el. Ennek legkisebb alkotóeleme az egy génen, illetve kromoszóma régióin belül elhelyezkedő polimorfizmusok haplotípusa, mely sok esetben hatékonyan kiszámítható *in silico* eszközökkel [245], ugyanakkor az egymáshoz viszonylag közeli variációk esetén megfelelő továbbfejlesztett genotipizáló technikákkal is meghatározható.

5.1.4.1 A *DRD4* gén SNP-inek haplotípus elemzése

A humán *DRD4* gén 5' régiójában lévő 3 SNP (rs1800955 „-521 C/T”, rs747302 „-616 C/G” és rs796487306 „-615 A/G”) haplotípusát RFLP és allél-specifikus PCR kombinálásával határoztuk meg. Az utóbbi két polimorfizmus egymás közvetlen szomszédságában helyezkedik el, és a környező szekvencia lehetővé tette olyan restriktív endonukleázok választását, melyek specifikusak valamelyik haplotípusra (2. táblázat). Az elemzést az rs1800955 „-521 C/T” polimorfizmusra úgy terjesztettük ki, hogy az első lépés során az egyik PCR-primer 3' végével az SNP-hez kötődött (allél-specifikus amplifikáció), ennek eredményeként – heterozigóta minták esetén – a C, illetve a T allélt tartalmazó génszakaszt külön reakcióban tudtuk felszorzozni.

2. táblázat. Az rs747302 „-616 C/G” és rs796487306 „-615 A/G” haplotípusának meghatározására alkalmazott restriktív endonukleázok. Az aláhúzott betűk a 2 SNP pozícióját, a zöld vastagbetűk pedig az adott enzim felismerőhelyét mutatják. W: A vagy T, R: A vagy G, Y: C vagy T

<i>a „-616 C/G” és a „-615 A/G” SNP-k haplotípusa a környező szekvenciával</i>	<i>specifikus restriktív endonukleáz</i>
TGAGG <u>A</u> CCA	<i>Ava</i> II (GGWCC)
TGAG <u>C</u> CCA	<i>Bsi</i> HKA I (GWGCWC)
TGAG <u>G</u> CCA	<i>Hae</i> III (GGCC)
TGAG <u>C</u> GCCA	<i>Hae</i> II (RGCGCY)

A haplotípus-elemzésbe a *DRD4* promoterében található 120 bp-os duplikációt is bevontuk oly módon, hogy az ismétlődő régiót valamelyik – az adott minta esetén heterozigóta genotípusú – SNP-re tervezett allél-specifikus PCR-rel amplifikáltuk. Ez az eljárás az ismétlődési

variációtól legnagyobb távolságban lévő rs1800955 „-521 C/T” SNP esetén egy 946 (1× 120 bp), vagy egy 1066 bp (2× 120 bp) nagyságú PCR-termék képződéséhez vezetett. Ezen fragmentumok detektálására eredményesen alkalmaztuk az ultravékony gélelektroforézis technikát is. A 190 µm vastagságú 1% agaróz–2% lineáris poliakrilamid gélben 0,2 µl PCR-termék elválasztása lézeres detektálással 12 perc alatt megvalósítható volt [XXI]. A kidolgozott technikák megbízhatóságát 108 három, illetve négy fős család tagjainak haplotipizálásával ellenőriztük. Az eljárások megfelelő működését igazolta, hogy a kapott haplotípus adatok a Mendel-szabályoknak minden esetben eleget tettek. A haplotípus-gyakoriságok eloszlását 598 független (nem-rokon) személy elemzésével határoztuk meg. A „-616C” és a „-615G” allél által alkotott haplotípus egyetlen esetben sem volt kimutatható az általunk vizsgált populációban, a C–A, G–A és G–G kombinációk gyakorisága 51,6%, 35,2%, illetve 13,2% volt (3. táblázat) [XI].

3. táblázat. A *DRD4* 120 bp-os duplikáció, rs747302 „-616 C/G”, rs796487306 „-615 A/G” és rs1800955 „-521 C/T” SNP-k haplotípus-gyakoriság értékei. Az adatok $N = 598$ személy PCR-alapú haplotípus meghatározásának eredményén alapulnak.

		1× 120 bp		2× 120 bp	
		-615A	-615G	-615A	-615G
-616C	-521C	3,8%	0,0%	16,5%	0,0%
	-521T	4,6%	0,0%	26,7%	0,0%
-616C	-521C	3,7%	0,5%	13,3%	8,8%
	-521T	4,0%	0,6%	14,3%	3,3%

5.1.4.2 MiR-SNP-k haplotípusának vizsgálata

A *WFS1* gén 3' UTR-ében lévő két SNP (rs1046322 A/G és rs9457 C/G) haplotípusát allél-specifikus PCR-rel elemeztük. Az egyes allélok kromoszomális elhelyezkedése ebben az esetben funkcionális szempontból is releváns, mivel eredményeink alapján mindkét polimorfizmus egy-egy miRNS bekötődését modulálja, így ezen szabályozó hatás megnyilvánulására az allélok relatív lokalizációja hatással lehet. Ez a jelenség még kifejezettebben nyilvánul meg a *SNAP-25* gén 3' UTR-ében, ahol az rs3746544 G/T és az rs1051312 C/T SNP-k egymástól csupán 4 bp távolságban vannak, és mindkét polimorfizmus érinti a miR-641 kötőhelyét. A polimorfizmusok genotipizálására PCR–RFLP-t (rs3746544 SNP: *Mnl* I, rs1051312: *Dde* I),

haplotípusuk meghatározására pedig TaqMan próbák alkalmazásán alapuló real-time PCR technikát állítottunk be (Függelék S1. táblázat). A négy haplotípusnak megfelelő próba közül kettőt-kettőt FAM-mal, illetve VIC-kel jelöltünk, az elemzés két reakcióelegyben (1.: FAM-mal jelölt T–T- és VIC-kel jelölt T–C-specifikus; 2.: FAM-mal jelölt G–T- és VIC-kel jelölt G–C-specifikus próba) történt. A polimorfizmusok között $D' = 1$ és $R^2 = 0,21$ LD-értékeket kaptunk, ami az SNP-k közeli elhelyezkedésének megfelelően nem meglepő, ugyanakkor érdekes, hogy a kapcsoltsági egyensúly hiánya mégsem teljes: az egyik lókuszon jelenlévő allélok alapján a másik polimorfizmus genotípusa nem adható meg egyértelműen. A magas Lewontin-féle D' -höz társuló alacsonyabb R^2 érték 3 haplotípus gyakori előfordulása esetén jellemző. Az általunk vizsgált populációban az rs3746544 T alléllal rs1051312 C és T variáns egyaránt előfordult (T–T és T–C haplotípusok), ugyanakkor az rs3746544 G formát csak G–T haplotípusban tudtuk kimutatni, a G–C haplotípus egyetlen esetben sem volt detektálható. Noha ezt az eredményt a Haploview, illetve Phase alkalmazások által számított értékek is megerősítették (4. táblázat), mégis említésre méltó, hogy a számítógépes algoritmusok által számított és a real-time PCR-es direkt haplotípus elemző eljárással meghatározott haplotípus gyakoriságok jelentős eltérést mutattak, ez a különbség – a három gyakori haplotípus csoportot tekintetbe véve – statisztikailag is szignifikánsnak bizonyult (χ^2 -teszt: $p = 0,000395$) [XII].

4. táblázat. A SNAP-25 rs3746544 G/T és rs1051312 C/T SNP-k haplotípus eloszlása.

Számított: az allélfrekvenciák alapján várható értékek kapcsoltsági egyensúlyt feltételezve.

Mért: a real-time PCR-rel meghatározott eredmények, illetve a Haploview és Phase alkalmazásával a genotípusadatok alapján számított várható haplotípus gyakoriságok ($N = 1376$).

	<i>Számított</i>	<i>Mért</i>	<i>Haploview</i>	<i>Phase</i>
T–T	0,48	0,40	0,35	0,41
G–T	0,26	0,36	0,39	0,33
T–C	0,17	0,24	0,26	0,26
G–C	0,09	0,00	0,00	0,000084

Bár a real-time PCR alapján mért, illetve a Haploview-val és Phase-zel számított eredmények a G–C haplotípus hiánya tekintetében nem mondtak ellent egymásnak, mégis – elvben – nem volt teljesen kizárható annak gyanúja, hogy a megfigyelés háttérében valamilyen technikai hiba áll. Ennek vizsgálatára két, független eljárást alkalmaztunk. Az SNP-k genotipizálása során alkalmazott primerpár segítségével hoztuk létre a két polimorfizmust tartalmazó 603 bp

hosszúságú fragmentumot, melyet a *Sau96* I enzim T–T, G–T és T–C haplotípusok esetén egyszer (224 bp + 379 bp), a G–C haplotípus jelenlétekor viszont két helyen hasít (80 bp + 144 bp + 379 bp). A termékben az SNP-ktől függetlenül jelenlévő kontroll hasító hely igazolta a reakció optimális működését: dupla hasítást, azaz G–C haplotípust azonban egyetlen esetben sem tapasztaltunk. A TaqMan próbákkal működő haplotípus meghatározást emellett elvégeztük azon DNS-kostrukciókat alkalmazva is, melyeket a *SNAP-25* 3' UTR-ének funkcionális elemzéséhez készítettünk el, és irányított mutagenezissel a vizsgált személyekben elő nem forduló G–C haplotípus változatot is létrehoztuk. A VIC festékkel jelölt G–C specifikus próba ebben a rendszerben a másik három szignálhoz hasonló intenzitású jelet adott, ami szintén bizonyította a real-time PCR eljárás megbízható működését. A G–C allélkombináció hiánya a haplotipizáló eljárás módosításának lehetőségét kínálta. A G–T haplotípus kimutatására szolgáló próbát NED-del jelölve az elemzés kettő helyett egy reakcióelegyben is elvégezhető volt, ami a haplotipizálás áteresztőképességének jelentős (kétszeres) emelkedéséhez vezetett. Ez az elrendezés azonban természetesen azt jelentette, hogy a rendszer csak a T–T, G–T és T–C kombinációk jelenlétét méri elfogadva a G–C haplotípus teljes hiányát [XII]. Megjegyzendő, hogy ma már nem jelent technikai akadályt 5 fluoreszcens festék (4 haplotípus és ROX belső kontrollként) együttes detektálása sem, a kísérletek elvégzésekor alkalmazott ABI 7300 Real-Time PCR berendezés azonban csak 4 csatornával rendelkezik.

5.2 Asszociáció vizsgálatok

Az ezredforduló óta a genetikai kutatások egyik fő célkitűzése a komplex jellegek és kórképek hátterében álló örökletes faktorok azonosítása. A kitűzött cél elérése nagyobb kihívásnak bizonyult, mint azt a Humán Genom Projekt befejezésekor hitték. Ennek hátterében számos tényező áll. A genotípus hatékony és pontos meghatározása mára már elérhetőnek mondható, az asszociációelemzések kulcsfontosságú része azonban a fenotípus körültekintő megállapítása is. Ez azért nem mindig egyszerű, mert a klinikai diagnózisok definiálása több esetben nem a háttérben meghúzódó, molekuláris szintű folyamatok, hanem a tünetek szerint történik. Az ilyen diagnosztikus kategóriák alkalmazása éppen a tanulmány eredményességét kockáztatja, mivel megeshet, hogy a molekuláris patomechanizmus szintjén eltérő betegek kerülnek egy homo-

gének vélt csoportba. Előrelépést jelenthet a több-lépcsős betegség modell, valamint az endofenotípusok alkalmazása. Az endofenotípus fogalmát legelőször a pszichiátriai genetika területén kezdték használni [246]. Olyan – lehetőség szerint számszerűen, objektíven – mérhető paramétert jelent, mely genetikailag meghatározott, és egyértelmű összefüggésben áll az adott kórképpel, ugyanakkor azzal nem egyezik meg. Az adott endofenotípus az átlagpopulációban is megfigyelhető lehet, ugyanakkor a beteg személyek egészséges családtagjainak körében nagyobb gyakorisággal fordul elő [247, 248]. Bizonyos endofenotípusok jelenléte így esetenként betegségmegelőző állapotnak tekinthető. Az endofenotípus koncepcióját ma már az orvostudomány több más területén is alkalmazzák [249-251], mivel a biológiai alapok felderítése során használatuk eredményesnek tűnik, és ily módon a genetikai háttér azonosítása, és így a célzott kezelés, valamint a megelőzés területén nagy jelentőségűek lehetnek.

A bemutatott genotipizáló és haplotipizáló eljárásokat klinikai együttműködések keretében alkalmaztuk. Számos komplex jelleg, illetve betegség örökletes rizikófaktorait kerestük kandidáns génvizsgálatok során. Ezek közül az agresszióval, a cukorbetegséggel, valamint az impulzivitással kapcsolatos eredményeink a *DRD4*, a *WFS1* és a *SNAP-25* gének egyes polimorfizmusaihoz köthetők.

5.2.1 Agresszió

Az agresszió – a rombolónak szánt viselkedés számos, különböző formája – összefüggésben áll több neuropszichiátriai kórképpel a verbális bántalmazástól az alkohol- és kábítószerfüggőségen, illetve neurodegeneratív kórképeken keresztül az öngyilkosságig [252-254]. Bizonyos mértékű agresszió természetesen egészséges személyek körében is jelen van, ami evolúciós szinten magyarázható. Az állatvilágban a megfelelő szintű agresszió elengedhetetlen az élelem és az élőhely megszerzéséhez, valamint a sikeres szaporodáshoz, ugyanakkor a szociális elszigetelődés és a sérülés kockázatát is magában hordozza. Az agresszió ezen pozitív és negatív következmények között egyensúlyt tartó optimális szintje feltételezetten genetikailag nagy mértékben meghatározott, ugyanakkor megnyilvánulásához – értelemszerűen – a környezeti hatások szintén hozzájárulnak [255]. Az agresszió megnyilvánulása a központi idegrendszer anatómiailag jól körülírt részeihez – az amygdalához, az elülső cinguláris kéreghez és a prefrontális kéreghez – köthető. Ez biokémiai szinten a monoaminerg (dopamin, noradrenalin, szerotonin)

neurotranszmitter rendszerek működésével hozható összefüggésbe. Átfogó asszociáció vizsgálatot végeztünk el a *DRD4*, valamint a monoaminerg ingerület átvivő rendszer számos további receptorának, transzporterének, továbbá a neurotranszmitterek anyagcseréjében szerepet játszó enzimek génjeiben található kiválasztott SNP-k és az agresszió között. A vizsgálatban 887 (54,2% nő, 45,8% férfi) pszichiátriai anamnézissel nem rendelkező önkéntes vett részt. A fenotípus mérésére a Buss–Perry-féle kérdőívet használtuk, amelynek négy alskálája a verbális, illetve a fizikális agresszió, valamint a düh (az agresszió affektív komponense) és a hosztilitás (ellenségesség, rosszindulatúság – az agresszió kognitív összetevője) szintjének mérésére alkalmas [XIII].

Az asszociáció elemzés során a *DRD4* és a monoaminerg neurotranszmitter rendszerekkel összefüggésben álló további fehérjék génjeiben lévő összesen 55 SNP-t (5. táblázat) elemeztünk. A polimorfizmusok kiválasztása során szem előtt tartottuk, hogy – az eredményes elemzés érdekében – olyan nem túl ritka variációkat vizsgáljunk, melyek MAF értéke eléri az 5%-ot. Előnyben részesítettük továbbá azokat a SNP-eket, melyek esetében valamilyen molekuláris biológiai funkció feltételezhető volt (aminosavcsere okozó vagy szabályozó régióban, transzkripció faktor vagy miRNS-kötőhelyét érintő variációk), illetve melyeket korábbi asszociáció elemzések valamilyen személyiségjeggyel vagy pszichiátriai rendellenességgel összefüggésbe hoztak. Az SNP-k genotipizálását a TaqMan próbák alkalmazásán alapuló (12. ábra) nagy átteresztőképességű OpenArray® rendszer alkalmazásával végeztük el. Ezen technika segítségével egyetlen kb. 2×7 cm²-es lemezen 33 nl végtérfogatú reakciókban 48 minta 64 SNP-jének genotípusa határozható meg.

A genotipizálás megbízhatóságának tesztelése céljából minden SNP esetén összevetettük a mért és az allélfrekvencia értékek alapján a Hardy–Weinberg egyensúly szerint várható genotípus gyakoriságokat. Az eredményeket az 5. táblázat foglalja össze, a gyakori allél által alkotott homozigóta genotípust „MM”, a ritka változatot tartalmazót „mm”, a heterozigóta genotípust pedig „Mm” jelöli. Megfigyelhető, hogy a mért és a számolt genotípus-frekvenciák közötti eltérés egyetlen esetben sem bizonyult statisztikailag szignifikánsnak (5. táblázat „HWE”). Az OpenArray® genotipizáló platform az elemzések 77,9%-ában (65–88%) bizonyult alkalmasnak a genotípus megbízható meghatározására. Említésre méltó, hogy egy másik, ugyanezzel a technikával végzett elemzésünk során az eljárás ennél lényegesen eredményesebbnek (83–99%, átlagosan 91,9%) mutatkozott [256].

Az 5. táblázat az SNP-k három genotípus csoportjába tartozó személyek önkítöltős kérdőívvel (Buss–Perry) mért átlagos agresszió pontszám értékét is mutatja. Nominálisan szignifikáns (ANCOVA $p < 0,05$) asszociációt két polimorfizmus, a *DRD4* promoterében lévő rs916455, valamint a szerotonin receptor 2A génjének (*HTR2A*) 2. intronjában elhelyezkedő rs7322347 SNP esetén kaptunk. A többszörös tesztelés miatt a Bonferroni-korrekciónak megfelelően ezen elemzés esetén a $p = 0,05 / 55 = 0,0009$ határértékeket kellett alkalmazni, ennek figyelembevételével csak a *HTR2A* génben található rs7322347 SNP-re vonatkozó eredmény tekinthető statisztikailag szignifikáns asszociációnak [XIII].

5. táblázat. A monoaminerg neurotranszmitter rendszer egyes génjeiben lévő kiválasztott SNP-k asszociáció analízise agresszióval. *M*: az adott SNP gyakori, *m*: a polimorfizmus ritka allélja. *HWE*: A Hardy–Weinberg egyensúlytól való eltérés statisztikai kiértékeléséhez tartozó *p* érték, *CR*: a sikeresen meghatározott genotípusok aránya („call rate”), *p*: A genotípus–fenotípus összefüggés statisztikai elemzéséhez tartozó nominális *p* érték ($N = 887$).

SNP	Gén	Genotípus eloszlás (%)			HWE	CR (%)	Agresszió pontszám			p
		MM	Mm	mm			MM	Mm	mm	
rs1048101	ADRA1A	28,6	50,3	21,1	0,945	86	66,7	66,5	66,5	0,9684
rs3808585	ADRA1A	54,8	38,4	6,8	0,998	81	66,2	68,2	65,9	0,2294
rs2236554	ADRA1D	38,7	45,7	15,6	0,641	85	65,3	67,2	68,3	0,0840
rs553668	ADRA2A	75,0	22,8	2,2	0,770	78	66,5	66,6	70,5	0,8682
rs11030104	BDNF	56,0	37,6	6,4	0,997	79	66,6	66,6	67,7	0,8735
rs2049045	BDNF	60,7	34,9	4,3	0,820	78	66,3	67,2	66,7	0,5703
rs6265	BDNF	60,2	35,3	4,5	0,847	68	66,9	66,8	66,0	0,9220
rs7103411	BDNF	55,0	38,6	6,4	0,966	81	66,6	66,5	67,3	0,9163
rs7094179	CDNF	44,4	44,0	11,6	0,924	77	65,8	66,5	68,3	0,6485
rs7900873	CDNF	55,2	39,2	5,6	0,573	78	67,0	66,5	64,7	0,3912
rs1051730	CHRNA3	42,5	45,8	11,7	0,943	85	67,5	65,5	66,6	0,1190
rs16969968	CHRNA5	42,1	46,3	11,6	0,866	75	67,5	65,8	66,6	0,2138
rs4680	COMT	29,4	48,9	21,7	0,927	68	67,1	66,6	67,6	0,8569
rs135745	CSNK1E	26,0	52,2	21,7	0,460	81	66,0	66,6	66,2	0,7121
rs1997644	CSNK1E	25,6	52,9	21,5	0,291	78	66,8	66,3	65,7	0,7781
rs1611115	DBH	58,2	37,2	4,6	0,482	86	65,7	67,4	70,9	0,0941
rs6271	DBH	84,2	14,9	0,9	0,759	88	66,6	66,7	61,0	0,8731
rs4532	DRD1	37,6	46,9	15,5	0,931	86	66,6	65,9	67,9	0,9000
rs6277	DRD2	29,2	49,1	21,8	0,948	65	66,8	67,1	66,4	0,9148
rs1800497	DRD2	66,0	31,7	2,3	0,261	68	66,7	67,7	61,3	0,7106
rs1079597	DRD2	72,9	26,0	1,2	0,226	69	67,2	66,4	58,0	0,4397

SNP	Gén	Genotípus eloszlás (%)			HWE	CR (%)	Agresszió pontszám			p
		MM	Mm	mm			MM	Mm	mm	
rs1800498	DRD2	36,1	47,1	16,8	0,862	67	67,1	66,5	67,3	0,7979
rs2134655	DRD3	53,9	38,8	7,2	0,981	86	65,7	67,8	66,1	0,1250
rs3732790	DRD3	33,1	49,7	17,2	0,857	83	67,0	66,4	65,9	0,5267
rs6280	DRD3	47,3	43,5	9,2	0,887	84	67,0	66,6	64,5	0,4667
rs963468	DRD3	33,4	49,5	17,1	0,909	83	67,3	67,0	65,1	0,5779
rs11246226	DRD4	25,3	50,7	24,1	0,941	77	67,3	66,3	66,4	0,4831
rs3758653	DRD4	68,1	29,1	2,8	0,923	80	66,5	66,5	69,9	0,9091
rs916455	DRD4	91,7	8,0	0,3	0,803	79	66,9	62,8	46,7	0,0275
rs936460	DRD4	49,4	40,7	9,9	0,655	79	66,6	66,4	67,7	0,9890
rs3733829	EGLN2	38,5	47,0	14,5	0,998	77	67,0	66,8	65,4	0,6238
rs222843	GABARAP	44,9	42,9	12,2	0,601	77	66,5	66,1	68,8	0,9562
rs11111	GDNF	75,1	22,3	2,6	0,241	81	66,6	65,9	73,9	0,9972
rs1549250	GDNF	32,5	49,7	17,7	0,907	80	66,8	65,4	70,3	0,8604
rs1981844	GDNF	55,6	38,7	5,7	0,771	65	66,5	66,9	72,2	0,4727
rs2910702	GDNF	54,9	38,2	7,0	0,971	79	66,3	66,4	68,8	0,5293
rs2973041	GDNF	70,8	26,2	3,0	0,710	78	66,7	66,2	72,0	0,9268
rs2973050	GDNF	41,6	47,3	11,2	0,608	66	66,3	66,5	68,9	0,5259
rs3096140	GDNF	47,7	42,8	9,5	1,000	76	65,8	67,0	68,5	0,1457
rs3812047	GDNF	76,7	21,2	2,1	0,559	77	66,5	67,5	70,9	0,3422
rs6925	HTR1A	27,5	47,6	24,9	0,510	68	66,6	67,6	65,6	0,9441
rs1228814	HTR1B	72,1	25,5	2,3	0,995	68	67,2	66,5	64,0	0,5336
rs130058	HTR1B	55,5	39,0	5,5	0,642	67	67,2	65,6	70,7	0,3419
rs13212041	HTR1B	62,0	34,5	3,5	0,467	68	67,1	66,2	66,9	0,5259
rs11568817	HTR1B	31,2	48,7	20,2	0,937	68	68,7	65,8	67,1	0,0605
rs6296	HTR1B	53,5	38,4	8,1	0,730	68	66,2	67,3	68,6	0,2601
rs6311	HTR2A	31,3	50,3	18,4	0,809	88	66,8	66,3	67,4	0,9130
rs6313	HTR2A	31,2	50,1	18,7	0,893	87	66,4	66,1	67,4	0,9969
rs6314	HTR2A	82,8	16,8	0,4	0,409	87	67,0	64,0	73,7	0,0765
rs7322347	HTR2A	31,6	48,4	20,0	0,866	86	69,2	64,9	66,4	0,0007
rs7984966	HTR2A	54,2	38,7	7,1	0,984	85	67,5	65,7	65,6	0,1356
rs3813929	HTR2C	74,6	15,7	9,7	0,975	84	67,0	63,7	66,4	0,2163
rs518147	HTR2C	52,9	23,2	24,0	0,237	81	66,7	66,4	67,2	0,3203
rs6318	HTR2C	74,1	16,5	9,4	0,737	87	66,4	66,5	67,6	0,2277
rs907094	PPP1R1B	58,0	34,9	7,1	0,308	79	67,0	65,2	68,8	0,2994

Ez a megfigyelés összhangban áll a szerotonin neurotranszmitter rendszer agresszióval és más társuló rendellenességekkel kapcsolatos szerepével, melyet már több évtizede ismerünk

[257]. Egybehangzó bizonyítékok állnak rendelkezésre arra vonatkozóan, hogy az alacsony, illetve károsodott szerotonerg funkció hozzájárul az agresszió és az impulzivitás kialakulásához [258, 259]. A központi idegrendszerben szerotonin a formatio reticularis (hálózatos állomány) raphe magjainak neuronjaiban szintetizálódik, az innen induló pályák az idegrendszer csaknem minden részéhez eljutnak. Emiatt a szerotonin hatása igen sokrétű és szerteágazó, amihez a különféle receptorok specifikus lokalizációja és működése is hozzájárul [260]. A szerotonerg neurotranszmisszió szerepet játszik többek között a veszélyes helyzetek elkerülésében; ennek megfelelően alulműködése ezen folyamatok károsodásához vezethet, ami viszont agresszív reakciókat válthat ki [261]. Ezt támasztja alá az is, hogy a szelektív szerotonin visszavétel-gátló gyógyszerek (fluoxetin, citalopram) az agressziót csökkentik, valamint megfigyelték, hogy a szerotonin-prekursor triptofán hiánya agresszív viselkedést eredményezhet [262, 263].

A *HTR2A* gén által kódolt receptor G_q fehérjén keresztül a foszfolipáz C enzim aktiválódása révén inozitol-1,4,5-triszfoszfát és diacil-glicerin másodlagos hírvivőkkel fejti ki hatását. Számos asszociáció elemzés vizsgálta a gén SNP-i, valamint különböző pszichiátriai rendellenességek, illetve személyiség-zavarok közötti összefüggést: az rs7322347 SNP-t a gyermekkori ADHD és a szkizofrénia lehetséges örökletes komponenseként azonosították [264, 265]. Ez a polimorfizmus a gén 4. intronjában az utolsó exontól kevesebb, mint 350 bp távolságra helyezkedik el. Ebben az exonban található az rs6314 SNP, mely aminosavcserét okoz: a két polimorfizmus közötti távolság 1069 bp, és a két lókuszt között magas LD-érték mérhető. Eredményeink alapján bár az utóbbi SNP az agresszióval csupán tendenciózus ($p = 0,0765$) asszociációt mutatott, a két polimorfizmus által alkotott haplotípus hasonló összefüggésben áll az agresszióval, mint az rs7322347 lókuszt önmagában. Az rs6314 SNP gyakori (91,2%) G alléja hisztidint kódol, ami A változat jelenléte esetén tirozinra módosul. Ez az aminosavcsere a fehérje töltésében is változást eredményez, ami hatással lehet a receptor térszerkezetére és működésére. Post mortem vizsgálatok tanúsága szerint az A allél a prefrontális kéregben alacsonyabb transzkripciós aktivitást eredményez, és bioinformatikai elemzés alapján az SNP szerepet játszhat az mRNS érése során az intronok kivágódásának folyamatában is [266]. Hasonló szabályozó hatás az rs7322347 polimorfizmus esetén is elképzelhető akár a splicing befolyásolása, akár a pre-mRNS stabilitásának megváltoztatása révén.

Az agresszió genetikai hátterének meghatározásában egy másik, a központi idegrendszerben szintén magas expressziót mutató gén, a *WFS1* polimorfizmusai is felmerültek. A gén funkcióvesztését okozó mutációk talaján kialakuló Wolfram-szindrómát ugyanis progresszív neurodegeneráció jellemzi: a betegség gyakran társul pszichiátriai rendellenességekkel: súlyos depresszióval, pszichózissal, valamint verbális és fizikai agresszióval [267]. Azon egészséges személyek, akik valamely Wolfram-szindrómát okozó variánst heterozigóta formában hordoznak, 26-szor gyakrabban szorulnak kórházi kezelésre pszichiátriai kórképpel, mint a mutációt nem hordozó személyek [268], ami rávilágít arra, hogy a gén kismértékű funkcióváltozása is hozzájárulhat pszichiátriai rendellenességek kialakulásához.

6. táblázat. A *WFS1* génben lévő kiválasztott SNP-k asszociáció analízise agresszióval.

M: az adott SNP gyakori, *m*: a polimorfizmus ritka allélja. *HWE*: A Hardy–Weinberg egyensúlytól való eltérés statisztikai kiértékeléséhez tartozó *p* érték, *CR*: a sikeresen meghatározott genotípusok aránya („call rate”), *p*: A genotípus–fenotípus összefüggés statisztikai elemzéséhez tartozó nominális *p* érték (*N* = 801).

SNP	Genotípus eloszlás (%)			HWE	CR (%)	Agresszió pontszám			p
	MM	Mm	mm			MM	Mm	mm	
rs10002743	66,3	29,2	4,5	0,3286	85	66,2	68,1	66,1	0,8711
rs6824720	71,2	25,8	3,0	0,7009	85	66,4	67,1	64,1	0,5821
rs752854	48,0	41,9	10,1	0,8929	68	67,5	66,5	66,8	0,8727
rs4689393	32,9	47,2	19,9	0,5939	83	66,2	67,5	66,6	0,9975
rs10010131	38,3	45,2	16,5	0,4360	79	66,8	66,2	64,8	0,3542
rs13147655	38,7	45,9	15,4	0,7596	81	66,6	67,4	66,1	0,7131
rs4467645	37,9	45,5	16,6	0,4722	86	66,5	67,2	66,2	0,7985
rs13128674	36,0	46,3	17,7	0,6145	67	66,9	67,4	67,2	0,8759
rs6446482	37,1	46,3	16,6	0,6788	82	66,3	67,3	67,0	0,8364
rs4689395	41,9	44,3	13,8	0,6246	83	66,6	66,6	67,0	0,8654
rs28716718	89,9	9,5	0,6	0,3157	86	66,5	66,8	72,2	0,4960
rs1801208	90,5	8,9	0,6	0,2059	87	66,6	68,5	72,5	0,5980
rs734312	29,0	49,7	21,3	1,0000	84	66,0	67,4	66,5	0,9215
rs1046316	44,4	43,0	12,6	0,6125	62	67,3	67,5	66,4	0,8790
rs1046320	36,8	45,8	17,4	0,3979	100	66,9	66,5	66,3	0,8793
rs1046322	81,4	17,5	1,1	0,8911	100	66,4	66,3	85,8	0,0005
rs9457	34,1	45,3	20,6	0,1072	93	66,3	67,3	65,8	0,5462

Az asszociációelemzésbe bevont 17 SNP átlagosan 1764 bp távolságra helyezkedett el egymástól, a köztük mérhető LD-értékeket a 13. ábra mutatja. Egyetlen polimorfizmus, a 3' UTR-ben elhelyezkedő rs1046322 SNP esetén volt megfigyelhető szignifikáns eltérés a különböző genotípuscsoportokba tartozó résztvevők kérdőívvel (Buss–Perry) mért agresszió pontszámai között. Az AA homozigóta genotípusú személyek agresszió értéke közel 20 ponttal meghaladta a G allélt hordozók (AG heterozigóták és GG homozigóták) eredményét (6. táblázat). Ez az asszociáció a többszörös tesztelés miatt alkalmazott Bonferroni-korrekciót követően is statisztikailag szignifikánsnak ($p < 0,0029$) bizonyult [V].

Megemlítendő, hogy a magas agresszióval jellemezhető AA homozigóta genotípus előfordulása meglehetősen ritka (6. táblázat), az általunk elemzett 801 fős populációban ez 9 személyt jelentett. Az eredmények megbízhatóságának további ellenőrzése céljából összehasonlítottuk az egyes genotípuscsoportokba tartozó személyek átlagpontszámát a kérdőív alsóskálái esetén külön-külön is. A genotípus–fenotípus asszociáció mind a négy kategóriában: a verbális, illetve a fizikális agresszió, a düh és a hosztilitás terén is megmutatkozott [V].

5.2.2 Cukorbetegség

Bár a pszichiátriai rendellenességek és a cukorbetegség között első pillantásra nem sok hasonlóság mutatkozik, mégis közös bennük, hogy mindegyikben szerepet játszhat valamilyen „szignál-molekula” – a neurotranszmitterek, illetve az inzulin – megváltozott exocitózisa. Ezt támasztja alá az a tény, hogy a *WFS1* gén funkciókiesése talaján kialakuló Wolfram-szindróma a neurodegeneráció mellett nem autoimmun eredetű, inzulinfüggő diabetes mellitus-szal és a β -sejtek progresszív pusztulásával is társul. Ennek alapján vetődött fel annak lehetősége, hogy a *WFS1* gén működésének kisebb mértékű változását eredményező polimorfizmusok szerepet játszhatnak a cukorbetegség gyakori, poligénes formáinak (T1DM és T2DM) kialakulásában [93-95].

Bár a diabetes két típusa egymástól sokban különbözik, közös bennük az inzulin hatásának elmaradása. Ennek hátterében T1DM esetén a β -sejtek autoimmun pusztulása áll, a 2-es típusú kórformára pedig inzulin-rezisztencia, majd ennek következtében az inzulin szekréció progresszív károsodása jellemző.

Az asszociáció vizsgálatba a *WFS1* gén hét kiválasztott SNP-jét vontuk be. Az allél alapú összehasonlítás eredményeit a 7. táblázat foglalja össze. Az elemzést az 1-es és a 2-es típusú cukorbetegségre vonatkozóan is elvégeztük, mindkét esetben nemben és korban illesztett kontroll csoportot alkalmaztunk az eset–kontroll vizsgálat során. T1DM-ben a Bonferroni-korrekciót követően az rs1046322 SNP hatása bizonyult szignifikánsnak, a polimorfizmushoz tartozó esélyhányados – a komplex jellegekre jellemző értékek tükrében – figyelemreméltóan magas. A 2-es típusú diabetes mellitus-szal a gén 4 polimorfizmusa, a 4. intronban lévő rs10010131, a 6. intronban elhelyezkedő rs13147655, valamint a 3' UTR-ben található rs1046320 és rs9457 SNP mutatott szignifikáns asszociációt (Bonferroni-korrekció: $p < 5,95 \cdot 10^{-4}$). Bár ezek a polimorfizmusok a gén eltérő részében, egy 11 884 bp hosszúságú szakaszon helyezkednek el, megemlítendő, hogy közöttük magas LD-érték mérhető (13. ábra), és feltételezésünk szerint a tényleges, molekuláris biológiai hatás elsősorban a 3' UTR-hez köthető. Érdekes ugyanakkor, hogy a kórkép 1-es típusával asszociációt mutató, szintén a 3' UTR-ben lévő rs1046322 polimorfizmus és a gén többi SNP-je között alacsony LD figyelhető meg [IV].

7. táblázat. A *WFS1* génben lévő kiválasztott SNP-k asszociáció analízise cukorbetegséggel. A táblázat bal oldala (T1) az 1-es típusú ($N = 407$, kontroll (*Ctrl*) csoport: $N = 817$), jobb oldala (T2) a 2-es típusú ($N = 380$, kontroll (*Ctrl*) csoport: $N = 892$) cukorbetegségre vonatkozó eredményeket foglalja össze. *Allélfr.*: a rizikóallél gyakorisága a beteg-, illetve a kontroll csoportban. *p*: A χ^2 -próbaéhoz tartozó nominális *p* érték. *OR*: Esélyhányados a 95%-os konfidencia-intervallummal

SNP, rizikó allél	Allélfr.		<i>p</i>	OR	Allélfr.		<i>p</i>	OR
	T1	Ctrl			T2	Ctrl		
rs6824720 G	0,84	0,81	0,15	1,22 (0,93–1,60)	0,88	0,82	$4,0 \cdot 10^{-3}$	1,51 (1,14–1,99)
rs10010131 G	0,64	0,61	0,13	1,15 (0,96–1,39)	0,69	0,61	$2,5 \cdot 10^{-4}$	1,43 (1,18–1,72)
rs13147655 G	0,65	0,60	0,05	1,23 (1,00–1,52)	0,69	0,61	$4,5 \cdot 10^{-4}$	1,44 (1,18–1,77)
rs1801208 G	0,96	0,96	0,88	1,04 (0,63–1,72)	0,96	0,95	0,55	1,15 (0,73–1,83)
rs1046320 A	0,60	0,59	0,68	1,04 (0,87–1,24)	0,67	0,59	$1,1 \cdot 10^{-4}$	1,43 (1,19–1,71)
rs1046322 A	0,19	0,09	$3,9 \cdot 10^{-11}$	2,33 (1,80–3,00)	0,14	0,09	$1,0 \cdot 10^{-3}$	1,56 (1,19–2,03)
rs9457 C	0,60	0,54	0,01	1,26 (1,06–1,50)	0,63	0,55	$8,1 \cdot 10^{-5}$	1,42 (1,19–1,70)

Az SNP-k és a cukorbetegség két típusa közötti összefüggést a genotípus eloszlások elemzésével is vizsgáltuk. A „recesszív modell” szerint a mindkét homológ kromoszómán rizikóallélt hordozó heterozigóta genotípusú személyek előfordulási gyakoriságát hasonlítottuk össze

a beteg, illetve az egészséges csoportban az eset–kontroll tanulmány során (8. táblázat). Ebben az elemzésben az rs1046322 SNP a korábinál is lényegesen magasabb (7,641) esélyhányados (OR) értékkel volt jellemezhető T1DM-ben. Az allél szerinti elemzés során azonosított négy polimorfizmus közül a 2-es típusú cukorbetegséggel három SNP – az rs10010131, az rs1046320 és az rs9457 – mutatott – a Bonferroni-korrektiót követően is – statisztikailag szignifikáns asszociációt (8. táblázat). Azonos tendenciákat kaptunk a „domináns modell” alkalmazása során is, amikor a rizikóallélt akár homozigóta, akár heterozigóta formában tartalmazó genotípuscsoportokat állítottuk szembe a másik homozigóta csoporttal, ebben az elemzésben azonban egyik SNP hatása sem bizonyul statisztikailag szignifikánsnak a Bonferroni-korrektió alapján számított szignifikancia szintet ($p < 5,95 \cdot 10^{-4}$) figyelembe véve. Ez az eredmény azt támasztja alá, hogy noha a domináns és recesszív öröklődésment elsősorban a monogénes jellegek esetén értelmezhető, mégis ebben az esetben is megfigyelhető, hogy a rizikóallél a betegségre való hajlamot homozigóta formában növeli csak, míg heterozigóta személyek körében a betegség kockázata a másik homozigóta genotípusú csoportéhoz hasonló.

8. táblázat. A WFS1 génben lévő kiválasztott SNP-k asszociáció analízise cukorbetegséggel (recesszív modell). Bal oldalon (T1) az 1-es típusú ($N = 407$, kontroll (*Ctrl*) csoport: $N = 817$), jobb oldalon (T2) a 2-es típusú ($N = 380$, kontroll (*Ctrl*) csoport: $N = 892$) cukorbetegségre vonatkozó eredmények láthatók. *Genotípusfr.*: a rizikógenotípus gyakorisága a beteg-, illetve a kontroll csoportban. *p*: A genotípus–fenotípus összefüggés statisztikai elemzéséhez tartozó nominális *p* érték. *OR*: Esélyhányados a 95%-os konfidencia-intervallummal

SNP, rizikó genotípus	Genotípusfr.		<i>p</i>	OR	Genotípusfr.		<i>p</i>	OR
	T1	Ctrl			T2	Ctrl		
rs6824720 GG	0,73	0,67	0,10	1,30 (0,95–1,78)	0,78	0,69	$4,1 \cdot 10^{-3}$	1,59 (1,16–2,19)
rs10010131 GG	0,43	0,37	0,05	1,31 (1,00–1,70)	0,49	0,37	$1,4 \cdot 10^{-4}$	1,65 (1,27–2,13)
rs13147655 GG	0,45	0,38	0,04	1,35 (1,01–1,82)	0,48	0,37	$1,1 \cdot 10^{-3}$	1,58 (1,19–2,08)
rs1801208 GG	0,92	0,92	0,99	1,00 (0,60–1,69)	0,91	0,90	0,69	1,10 (0,68–1,78)
rs1046320 AA	0,39	0,35	0,22	1,17 (0,91–1,50)	0,47	0,35	$1,8 \cdot 10^{-4}$	1,61 (1,25–2,06)
rs1046322 AA	0,09	0,01	$7,0 \cdot 10^{-10}$	7,64 (3,61–16,1)	0,03	0,01	0,02	2,70 (1,11–6,57)
rs9457 CC	0,38	0,30	0,01	1,38 (1,07–1,78)	0,42	0,30	$7,5 \cdot 10^{-5}$	1,65 (1,29–2,13)

Az elemzések a 3' UTR-ben lévő polimorfizmusok jelentőségét sugallták, ezért az itt elhelyezkedő 3 SNP – az rs1046320, az rs1046322 és az rs9457 – haplotípus alapú asszociáció

analízisét is elvégeztük. Az 1-es típusú cukorbetegségben szenvedő betegek elemzése megerősítette az rs1046322 polimorfizmus meghatározó szerepét. Az SNP „A” allélját, valamint az egyik vagy mindkét polimorfizmus rizikóallélját tartalmazó haplotípusok gyakorisága a betegcsoportban szignifikánsan magasabbnak mutatkozott, az esélyhányados mindhárom esetben 3 felett volt. A T2DM betegcsoport haplotípus elemzése megerősítette az allél, illetve genotípus alapú vizsgálatok eredményét. Bár az rs1046322 SNP önmagában nem mutatott statisztikailag szignifikáns asszociációt a betegséggel, rizikó alléljának jelenléte akár az rs1046320, akár az rs9457 vagy mindkettő SNP-vel haplotípusban jelentősen magasabb – 1,8 fölötti – esélyhányadost eredményezett, mint az rs1046320–rs9457 „AC” haplotípus, melynek OR értéke csupán 1,341 volt. Megjegyzendő, hogy míg ez utóbbi 2 SNP között magas LD mérhető, az rs1046322 a másik két polimorfizmussal kapcsoltsági egyensúlyban van (13. ábra), így ez az eredmény az rs9457 és az rs1046322 SNP független és egymást erősítő molekuláris szintű szerepét sejteti [IV].

Az asszociáció vizsgálatok eredményei alátámasztják a korábbi elemzéseken, illetve klinikai megfigyeléseken nyugvó feltételezést, miszerint a *WFS1* polimorfizmusai szerepet játszanak a T2DM genetikai hátterének meghatározásában. Kiemelendő, hogy vizsgálatunkat viszonylag magas létszámú, 1-es típusú cukorbetegségben szenvedő betegcsoportra is kiterjesztettük, és érdekes, hogy bár a gén szerepe itt is meghatározónak mondható, mégsem ugyanaz a polimorfizmus mutatta a legkifejezettebb asszociációt a két kórforma esetében. Felmerülhet a gondolat, hogy egyáltalán várható-e, hogy a T1DM és a T2DM hátterében valamelyest átfedő örökletes tényezők, illetve molekuláris szintű folyamatok álljanak. Az „akcelerátor hipotézis” [269] – a két kórforma számos lényegi különbözősége ellenére – ezt az álláspontot képviseli. Megfigyeléseink alapján feltételezhető tehát, hogy a wolframín fehérje mennyiségének megváltozása a cukorbetegség mindkét típusában szerepet játszhat. Ehhez járulhat hozzá a T2DM-ben azonosított rs9457 SNP, ami megmagyarázhatja a korábbi elemzések által azonosított lókuszzal (rs10010131, rs1046320) mérhető asszociációt, mivel ezen polimorfizmusok és az rs9457 SNP között magas LD mérhető. Az rs1046322 SNP a 3' UTR-ben hasonlóképpen hatással lehet a képződő fehérje mennyiségére, ugyanakkor a két folyamat értelemszerűen mégis eltérő, hiszen az rs9457 és az rs1046322 polimorfizmus más-más miRNS kötőhelyében található, így eltérő szabályozási hálózat részeként befolyásolhatják a génkifejeződés aktivitását.

5.2.3 Impulzivitás

Az impulzivitás egy többdimenziós személyiségjegy, melyre a tettek átgondolatlansága, a cselekedetek, valamint a viselkedés csökkent önkontrollja jellemző. A háttérben álló neurológiai, biokémiai folyamatok meghatározásáért örökletes tényezők is jelentős mértékben (kb. 45%) felelősek [270]. Az agresszivitáshoz hasonlóan megnyilvánulása széles skálán mozog, és az impulzivitás – bizonyos mértékben – az egészséges személyek természetes jellemzője is. Klinikai szempontból ugyanakkor mégis kiemelő, hogy a fokozott impulzivitás összefüggésben áll különböző pszichiátriai rendellenességekkel: a személyiségjegy az ADHD mellett a bipoláris hangulatzavar endofenotípusának is tekinthető [271, 272]. Háttérben ennek megfelelően a monoaminerg neurotranszmitter rendszer számos komponensét kutatták [273, 274], ezen ingerületátvivők működéséhez ugyanakkor az egyik legjelentősebb központi idegrendszeri fehérje, a SNAP-25 alapvetően szükséges. Ez a fehérje a SNARE-komplex részeként nélkülözhetetlen a neurotranszmittert tartalmazó vezikulumok sejtmembránnal történő fúziójában és exocitózisában. A gén 3' UTR-ében lévő rs3746544 SNP – vélhetően a fehérje általános központi idegrendszeri funkciójának következtében – az ADHD [275] mellett más neuropszichiátriai rendellenességekkel, sclerosis multiplex-szel [276], szkizofréniával [277] is szignifikáns asszociációt mutatott. Feltételeztük, hogy az rs3746544 polimorfizmus az rs1051312 SNP-vel funkcionális egységet képez, mivel egymástól csupán 4 bp távolságra helyezkednek el, és mindkét lokusz érinti a miR-641 kötőhelyét a SNAP-25 3' UTR-ében [XIV]. Vizsgálatainkat emellett kiterjesztettük a gén promoterében elhelyezkedő két kiválasztott SNP-re is: az rs6077690 A/T a transzkripció starthelyétől 2016 bp, az rs6039769 A/C polimorfizmus pedig 523 bp távolságra helyezkedik el.

A 3' UTR-ben lévő két polimorfizmus (rs3746544 és rs1051312) elemzését egy lépésben, TaqMan próbák alkalmazásán alapuló eljárással végeztük el. Az SNP-k közeli elhelyezkedése lehetővé tette, hogy haplotípus specifikus próbákat alkalmazzunk, a polimorfizmusok genotípusa a haplotípus adatok alapján egyértelműen meghatározható volt. A promoterben lévő két SNP (rs6077690 és rs6039769) egymástól 1493 bp távolságra helyezkedik el, emiatt genotípusukat külön-külön elemeztük, az előbbi esetben PCR-RFLP-t (*Függelék S1. táblázat*), utóbbi esetben allél-specifikus TaqMan próbák alkalmazásán alapuló eljárást használtunk. A két SNP közötti LD-t, valamint az egyes minták haplotípusát HaploView, illetve Phase segítségével ha-

tároztuk meg. Említésre méltó, hogy a 3' UTR-hez hasonlóan (4. táblázat) ezen két polimorfizmus esetében is csupán három haplotípus (T–C: 57,9%, A–A: 26,9%, A–C: 14,6%) fordult elő viszonylag magas gyakorisággal az általunk vizsgált populációban ($N = 901$). Ennek oka az SNP-k között magas LD ($D' = 0,96$, $R^2 = 0,50$). A T–A haplotípus frekvenciája csupán 0,6%-nak bizonyult, ami azért is érdekes, mert ezt a haplotípust az rs6077690 SNP gyakori allélja (T: 58,3%) alkotja a másik polimorfizmus ritka (rs6039769 A: 27,6%) variánsával. Kapcsoltsági egyensúlyt feltételezve ennek megfelelően ezen haplotípus frekvenciája 16,1%-nak adódna. A promoterben, illetve a 3' UTR-ben lévő két-két polimorfizmus között ugyanakkor kapcsoltsági egyensúly volt megfigyelhető.

Az asszociáció vizsgálatban 901 fiatal (átlagos életkor: $21,3 \pm 3,3$ év), egészséges, pszichiátriai kórtörténettel nem rendelkező személy (nők: 54,9%, férfiak: 45,1%) vett részt. Első lépésben a négy SNP hatását egyenként vizsgáltuk: összehasonlítottuk az egyes genotípus csoportba tartozó személyek Barratt-kérdőívvel mérhető átlagos impulzivitás pontszámát. A promoterben lévő rs6077690 és rs6039769 SNP-k nem mutattak összefüggést a vizsgált jelleggel. Azokra a személyekre viszont, akik a 3' UTR-ben lévő rs1051312 polimorfizmusra nézve homozigóta CC genotípusúak voltak, magasabb impulzivitás volt jellemző ($p = 0,042$), ez a különbség azonban a Bonferroni-korrekciót követően nem bizonyult statisztikailag szignifikánsnak. Az asszociáció elemzést a továbbiakban a promoterben, illetve a 3' UTR-ben lévő két-két SNP haplotípusának vizsgálatával folytattuk. A promoter polimorfizmusai ebben a megközelítésben sem mutattak összefüggést az impulzivitással. Azon személyek impulzivitás pontszáma viszont szignifikánsan alacsonyabb volt ($p = 0,009$), akik a 3' UTR-ben elhelyezkedő rs3746544 és rs1051312 SNP-k T–T haplotípusát hordozták szemben a G–T, valamint T–C haplotípussal rendelkezőkkel [XIV]. Bár a két csoport Barratt-kérdőívvel mérhető átlagos impulzivitás értéke közötti különbség viszonylag kicsi (1,34 pont) volt, lényeges hangsúlyozni, hogy vizsgálatunkban kizárólag egészséges személyek vettek részt, akik körében az impulzivitás szélsőséges értékei nem fordulnak elő. Ennek ellenére a jelleg genetikai meghatározottsága megfigyelhető volt, ami megerősítette, hogy a *SNAP-25* az impulzivitás endofenotípussal összefüggésben álló pszichiátriai rendellenességek egyik kandidáns génjének tekinthető.

A humán tanulmányok mellett az elmúlt években egyre több kutatás foglalkozik különböző állatok viselkedését meghatározó molekuláris folyamatok, örökletes faktorok feltárásával. A

kutyák elemzése több okból is különleges érdeklődésre tart számot. A kutya az első nagyragadozó házasításával létrejött faj [278], mely évszázadok óta számos különböző funkciót tölt be az emberi társadalomban. A faj elemzése ennek következtében túl is mutat a kutyafajták tulajdonságainak, a rájuk jellemző hasonló és eltérő vonások megismerésén: számos tekintetben az emberi viselkedés és személyiségjegyek, akár pszichiátriai rendellenességek feltárását célzó tanulmányokban modellként alkalmazható. Számos hasonlóság figyelhető meg nem csak a szociális viselkedés, de még a kommunikáció terén is az emberek és a kutyák között [279]. A kisgyermek szüleikhez való kötődése, valamint a kutyák és gazdáik közötti kapcsolat számos párhuzamot mutat [280], és hasonló folyamatok játszódnak le az emberek és a kutyák központi idegrendszerében, amikor fajtársaik, illetve a másik faj hangját hallják [281]. Munkacsoportunk az impulzivitás endofenotípus és a *DRD4* 3. exonjában lévő VNTR (4. ábra) feltételezett kapcsolatát kutatta 189 kutya elemzésével. Eredményeink rávilágítottak a komplex jellegek, valamint azok tanulmányozásának több fő aspektusára. A családi kutyák csoportjában ugyanis nem volt asszociáció a jelleg és a polimorfizmus között, ugyanakkor a 2/2 genotípusú rendőr-kutyákat szignifikánsan alacsonyabb impulzivitás jellemezte, mint a 2/3a, valamint a 3a/3a genotípus csoportba tartozó társaikat. A megfigyelés háttérében vélhetően a gén–környezet kölcsönhatás állhat. A családi kutyák esetén a jelentősen eltérő környezeti hatások mellett a *DRD4* VNTR genetikai hatása nem volt kimutatható szemben a rendőr-kutyákkal, akik hasonló környezetben, hasonló körülmények között éltek, és azonos kiképzésben részesültek [XXIII]. A megfigyelés egyik magyarázata az lehet, hogy a komplex jellegek háttérében az örökletes és a környezeti faktorok bonyolult kölcsönhatása áll. A rendszerint kis hatású genetikai polimorfizmusok sikeres azonosítása ennek megfelelően a lehető leghomogénebb populációk elemzését teszi szükségessé.

5.3 Polimorfizmusok molekuláris funkciójának elemzése

A komplex jellegekkel statisztikai asszociációt mutató polimorfizmusok esetén felvetődik a kérdés, hogy feltárható-e az adott genetikai variációhoz rendelhető molekuláris biológiai funkció, és ha igen, ez milyen kapcsolatban állhat az adott fenotípussal.

Kutatócsoporthozunk korábban kapilláris zónaelektroforézis alkalmazásán alapuló EMSA eljárást optimalizált transzkripció faktorok DNS-kötésének *in vitro* elemzésére [XXIV]. Ezt a technikát használtuk a *DRD4* gén promoterében lévő 120 bp-os duplikáció funkcionális vizsgálata során. Ismert (homozigóta 1×/1×, illetve 2×/2×) genotípusú személyek mintáját alkalmazva PCR-rel hoztuk létre a jelölt próbákat, valamint a kompetitoroként szolgáló jelöletlen DNS-fragmentumokat. A DNS–fehérje kölcsönhatás elemzése során az emelkedő mennyiségű HeLa sejtkivonat az alacsonyabb mobilitású DNS–fehérje-komplex mennyiségének növekedését eredményezte, ami arra utalt, hogy a vizsgált DNS-régióban transzkripció faktor kötőhely található. Szekvencia analízis alapján felmerült az Sp 1 fehérje lehetséges szerepe, amit egy 29 nt hosszúságú tisztított Sp 1 kötőhelyet tartalmazó kompetitor segítségével elemeztünk. Ez a DNS-molekula a 120 bp-os ismétlődést tartalmazó DNS-szakaszról a kötődő transzkripció faktort hatékonyan leszorította. A kísérletet „fordított felállásban” is elvégeztük: a fehérje-kötődést jelölt Sp 1-specifikus oligonukleotid és jelöletlen 1×, valamint 2× 120 bp szekvenciát tartalmazó kompetitor jelenlétében elemeztük. Az eredmények alapján a hosszú változat szignifikánsan erősebben kötődött a transzkripció faktorhoz, mint az 1× forma [XXV]. Megfigyeléseink tükrében felmerült annak lehetősége, hogy az ADHD-val asszociációt mutató *DRD4* gén 3. exonjában lévő 48 bp VNTR és az rs1800955 „–521 C/T” SNP mellett ez a polimorfizmus is hozzájárul a dopamin válasz egyéni variabilitásához.

Kutatócsoporthozunk további munkájának harmadik, talán legfőbb pillére az egyes jellegekkel asszociációt mutató polimorfizmusok funkcionális vizsgálata volt. Izgalmas kérdésnek ígért, hogy a kiválasztott célgének – főként a *WFS1* és a *SNAP-25* – promoterében, illetve 3' UTR-ében lévő polimorfizmusok eltérő mechanizmussal szerepet játszhatnak-e a génkifejeződés szabályozásában.

5.3.1 A *SNAP-25* gén polimorfizmusainak funkcionális elemzése

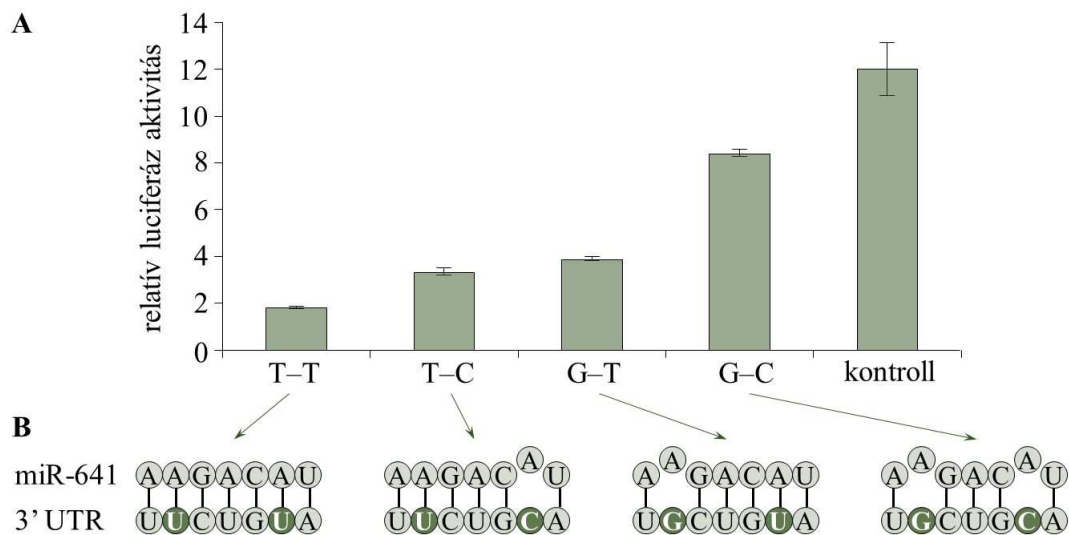
Szakirodalmi adatok, illetve saját asszociáció vizsgálatunk eredményei alapján a *SNAP-25* gén az impulzivitás endofenotípus és vele kapcsolatban álló pszichiátriai rendellenességek – ADHD, hangulatzavar – kandidáns génje, ennek tükrében a gén kiválasztott polimorfizmusainak molekuláris biológiai funkcióját is tanulmányozni kívántuk. Megfigyeléseink alapján a gén

promoterében lévő SNP-k (rs6077690 és rs6039769) nem mutattak asszociációt az impulzivitással, így az *in vitro* elemzésbe a 3' UTR két polimorfizmusát (rs3746544 és rs1051312) vontuk be. Ez a két SNP lokalizációjuk szempontjából is különleges: egymástól csupán 4 bp távolságra helyezkednek el, és az elméletileg lehetséges négy haplotípusból – a közöttük mérhető magas LD miatt – csupán három fordult elő az általunk vizsgált populációban (4. táblázat). A két polimorfizmus által alkotott haplotípusok jelentőségére az *in silico* szekvencia analízis is felhívta a figyelmet: az SNP-k a miR-641 „seed” régiójának kötőhelyében találhatók, így felmerül annak lehetősége, hogy együttesen járulnak hozzá a SNAP-25 kifejeződésének szabályozásához.

Az *in vitro* funkcionális elemzéseket HEK293T sejtek, valamint luciferáz riportter rendszer segítségével végeztük el. Ez a – vese eredetű – sejtvonal megbízhatóan fenntartható, hatékonyan transzfektálható, így a molekuláris biológiai kísérletek széles körben használt és elfogadott sejtvonala. Noha az általunk vizsgált fenotípus szempontjából a vese nem releváns sejttípus, korábbi mérések [282], illetve saját eredményeink alapján a SNAP-25 kifejeződik a HEK293T sejtekben. Kvantitatív PCR-es méréseink során a miR-641 hasonlóképpen detektálhatónak mutatkozott. Az rs3746544 és rs1051312 SNP-k szabályozó hatásának vizsgálatához a SNAP-25 teljes 3' UTR szakaszát pMIR Report luciferáz konstrukcióba, a természetes lokalizációnak megfelelően a kódoló szekvencia mögé klónoztuk (Függelék S1. táblázat). Irányított mutagenézissel hoztuk létre az egyes haplotípusoknak megfelelő konstrukciókat (Függelék S1. táblázat), elkészítettük a G–C variánst tartalmazó rekombináns DNS-t is annak ellenére, hogy ez az a változat, amely az általunk elemzett populációban nem fordult elő. Kontrollként olyan konstrukciót is létrehoztunk, amely a SNAP-25 3' UTR-ével megegyező hosszúságú inzertet tartalmazott, amiben azonban a miR-641 kötőhelye nem volt megtalálható. A mérés során a sejteket a megfelelő luciferáz konstrukció mellett miR-641-gyel, valamint belső kontrollként β -galaktozidázt kódoló konstrukcióval kotranszfektáltuk.

Legalacsonyabb relatív luciferáz aktivitást a T–T haplotípus jelenléte estén figyeltünk meg. Ez összhangban állt azzal, hogy ezen haplotípus esetén a miR-641 „seed” régiója tökéletesen komplementer a SNAP-25 3' UTR-ével (22. ábra). A G–T, valamint a T–C haplotípus a kötőhely egy-egy nukleotidját módosította, ami a relatív luciferáz aktivitás statisztikailag szignifikáns ($p < 0,05$), 1,8, illetve 2,1-szeres emelkedéséhez vezetett. A két haplotípushoz tartozó

érték között ugyanakkor nem mutatkozott jelentős eltérés, ami azzal magyarázható, hogy egyetlen bázis megváltozása a „seed” régió 5', illetve 3' végétől egy-egy nukleotid távolságban hasonló mértékben befolyásolta a két RNS kölcsönhatását. A G–C haplotípus esetén a miR-641 „seed” szekvenciájának két nukleotidja nem komplementer az mRNS-sel, ami szignifikáns ($p < 0,05$), 4,6-szeres relatív luciferáz aktivitást eredményezett a T–T haplotípushoz képest (22. ábra). A megfigyelés háttérében vélhetően az áll, hogy két nukleotid eltérése már jelentősen csökkenti a miRNS bekötődésének hatékonyságát. A G–C haplotípusra jellemző aktivitás már nem volt szignifikánsan alacsonyabb a kontroll konstrukcióhoz tartozó értéknél, melyben a miR-641 kötőhelye nem található meg, így annak translációt gátló hatása nem tud megvalósulni [XIV].



22. ábra. A *SNAP-25* 3' UTR-ében lévő rs3746544 és rs1051312 SNP-k és a miR-641 kölcsönhatásának elemzése. *A*: Az egyes haplotípusokat tartalmazó konstrukciók relatív luciferáz aktivitása. A „kontroll” konstrukcióban a miR-641 kötőhelye nincs jelen. *B*: A miR-641 „seed” szekvencia és a *SNAP-25* 3' UTR kapcsolódása. A sötétzöld háttérű pozíciók az SNP-eket mutatják. T–T haplotípus esetén a miR-641 „seed” szekvenciája tökéletesen komplementer az mRNS-sel. T–C és G–T haplotípus esetén egy-egy, míg G–C variáns jelenlétekor kettő nukleotid nem kötődik a 3' UTR-hez.

A miRNS-ek révén megvalósuló összetett, hálózatos jellegű szabályozási folyamatok megismerése lényeges előrelépést jelent a komplex jellegek molekuláris háttérének feltárása során,

ugyanakkor rávilágít a rendszer – korábban véltnél is nagyobb – összetettségére. Érdekes, hogy egy évtizeddel ezelőtt még úgy vélték, hogy a miRNS-eket kódoló, evolúciósan konzervált régiókban az SNP-k előfordulási gyakorisága meglehetősen alacsony [283]. Mára azonban már a miRNS-eket kódoló génekben is számos polimorfizmust azonosítottak, és megfigyelték, hogy ezek különböző kóros állapotok, pl. T2DM [284], transzplantációt követő tüdőfibrozis [285], autoimmun kórképek és sclerosis multiplex [286], valamint hepatocelluláris karcinóma [287] genetikai rizikófaktorai lehetnek. Ez a megfigyelés felhívja a figyelmet arra, hogy a betegségek hátterében álló polimorfizmusok nem ritkán nem csak a rendellenesség kialakulásában közvetlen szerepet játszó fehérvérjék, hanem az azokat szabályozó faktorok génjeiben lehetnek, ami a kép komplexitását hihetetlen mértékben megnöveli.

Mindezek tükrében a *SNAP-25* gén 3' UTR-ében lévő rs3746544 és rs1051312 polimorfizmusok molekuláris szinten szerepet játszhatnak kóros folyamatok létrejöttében, mivel ismert, hogy a *SNAP-25* expressziójának megváltozása összefüggésben áll az ADHD-val és kimutatható Alzheimer-betegségben is [288]. A *SNAP-25* működésének variabilitásához hozzájárul továbbá az is, hogy a génről két eltérő transzkripciós variáns jöhet létre. Az izoformák ötödik exonja különböző, ami fehérje szinten 9 aminosav eltérést eredményez, ehhez eltérő agyi, illetve sejten belüli lokalizáció [106], valamint valamelyest különböző funkció köthető [105].

Kutatócsoportunk két független eljárást dolgozott ki az izoformák expressziós szintjének elemzésére. Ehhez kontrollként létrehoztuk a két izoformát kifejező expressziós konstrukciót: a *SNAP-25* kódoló szakaszát pcDNA 3.1(–) vektorba klónoztuk (*Függelék S1. táblázat*), majd szekvenálás alapján választottunk ki egy-egy konstrukciót, melyben az *a*, illetve a *b* izoforma volt megtalálható. A két variáns szemikvantitatív mérése azon alapult, hogy az *a* izoformát a *Dde* I restrikciós endonukleáz szelektíven hasítja, az enzim felismerő helye *b* változat 5. exonjában viszont nincs jelen. A vizsgált cDNS-mintákból a régió PCR-amplifikációját (*Függelék S1. táblázat*) követően a hasított, illetve hasítatlan termékek denzitometriás analízise alapján az izoformák relatív mennyisége meghatározható volt. A kapilláris gélelektroforézis alkalmazása a mennyiségi mérés során a hagyományos, alámerülő gélelektroforézishez képest reprodukálhatóbb eredményt adott. A két izoforma pontosabb és hatékonyabb meghatározására Sybr Green interkalátor alkalmazásán alapuló real-time PCR-t állítottunk be. A primereket úgy terveztük meg (*Függelék S1. táblázat*), hogy mindkét oligonukleotid számos bázisa, valamint a 3' vég 2–2 nukleotidja a megfelelő izoformára specifikus legyen, így a két transzkripciós variáns

a real-time PCR során – specifikus próba nélkül is – tökéletesen elkülöníthetőnek bizonyult. A kidolgozott módszerek segítségével a *SNAP-25* variánsok mRNS-szintjét néhány humán szövetben elemeztük. Legmagasabb expresszió – a fehérje működésének megfelelően – hasnyálmirigyben és a központi idegrendszerben volt mérhető, a többi általunk vizsgált szövetben a *SNAP-25* expressziója 2–4 nagyságrenddel alacsonyabb volt. Érdekes, hogy a központi idegrendszer minden vizsgált területén (frontális, occipitális, parietális kortex) a *b* izoforma dominált, ezzel szemben a többi szervben (szív, vese, máj, vázizom, lép, hasnyálmirigy) az *a* variáns magasabb expressziója volt megfigyelhető [XXVI]. A különböző transzkripciós izoformák jelenléte a genetikai variabilitás sajátos formájának tekinthető. A humán genomban a fehérje kódoló gének száma jóval alacsonyabb, mint azt kezdetben hitték, aminek oka éppen az, hogy – pl. alternatív transzkripciós starthely, alternatív splicing, RNS szerkesztés – révén egyetlen gén igen gyakran több különböző fehérjét kódolhat. Jól ismert, hogy az exon–intron határt megváltoztató mutációk az mRNS hibás érése révén sok esetben nem működő fehérje keletkezéséhez, és így monogénes kórképek kialakulásához vezethetnek. Kevesebb adat áll rendelkezésre ugyanakkor azzal kapcsolatban, hogy a fiziológián előforduló, egymástól kisebb-nagyobb mértékben eltérő izoformák arányának eltolódása hozzájárulhat-e betegségek molekuláris patomechanizmusához. A *SNAP-25* esetében a transzkripciós variánsokat állatokon vizsgálták csak [289, 290], aminek nyilvánvalóan részben technikai oka is van, hiszen a génexpresszió elemzése (szemben a polimorfizmusok tanulmányozásával) csak a megfelelő szövetből lehetséges. Módosult alternatív splicing, egyebek mellett a p53 izoformák arányának eltolódása [291] összefüggésben áll különböző daganatos megbetegedések kialakulásával [292]. Az apoptózis szabályozásában szerepet játszó fehérjék (Bcl-XL, BIM, kaspázok) keletkezése során az alternatív splicing megváltozása szerepet játszik a T1DM kialakulásában [293]. A fibronectin izoformák arányának módosulása kapcsolatba hozható a rheumatoid arthritis kialakulásával [294]. A *SNAP-25* izoformák idegrendszeri és idegrendszeren kívüli eltérő aránya minden bizonnyal összefüggésben áll azzal, hogy a splicing szabályozása a központi idegrendszer sejtjeiben kissé eltér az extraneurális sejtek intronkivágódási mechanizmusától. Ennek zavara több neurológiai kórképpel (Alzheimer-betegség, Parkinson-kór, szkizofrénia) áll összefüggésben [295]. Ezen megfigyelések tükrében elképzelhető az is, hogy akár a fiziológián jelen lévő transzkripciós izoformák arányának megváltozása is részt vehet a különböző kórképek kialakulásáért felelős molekuláris mechanizmusokban.

5.3.2 A *WFS1* gén polimorfizmusainak funkcionális vizsgálata

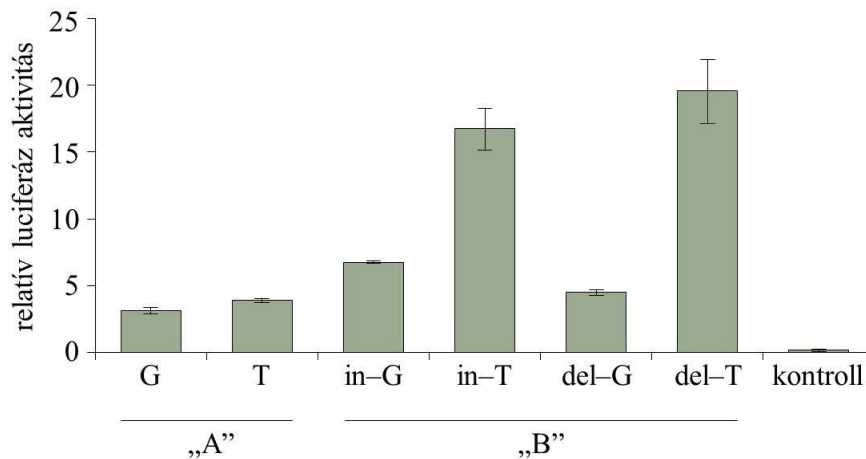
A wolframin és a cukorbetegség kapcsolatát állatkísérletek, illetve a *WFS1* gén funkcióvesztő mutációi talaján kialakuló Wolfram-szindróma bizonyítja [86, 87]. A gén polimorfizmusai között megfigyelhető viszonylag magas LD (13. ábra) ugyanakkor felveti a kérdést, hogy a betegséggel az asszociáció vizsgálatok során összefüggést mutató polimorfizmusok közül melyek esetén mutatható ki a wolframin kifejeződését vagy működését befolyásoló molekuláris biológiai hatás.

5.3.2.1 Az 5' régió polimorfizmusainak elemzése

A *WFS1* gén 5' régiójában elhelyezkedő rs148797429 és rs4273545 polimorfizmusokat *in vitro* sejtes rendszerben, pGL3 Basic luciferáz riportter konstrukció alkalmazásával elemeztük. A gén minimál promotere a –233. és a +104. nukleotidok között található [296], az ennek megfelelően elkészített, „A”-val jelölt, rövidebb (338 bp-os) konstrukció azonban csak az rs4273545 SNP-t foglalta magában. Létrehoztunk emiatt egy hosszabb („B” 554 bp) rekombináns DNS-t is (Függelék SI. táblázat), mely az SNP mellett az rs148797429 in/del polimorfizmust is tartalmazta. Az egyes alléloknak, illetve haplotípusoknak megfelelő konstrukciókat irányított mutagenézis segítségével készítettük el (Függelék SI. táblázat). Méréseink igazolták, hogy a régió promoter aktivitással rendelkezik: mind a rövidebb, mind a hosszabb konstrukciók szignifikánsan magasabb ($p < 0,05$) relatív luciferáz aktivitást mutattak, mint az üres pGL3 Basic vektor (23. ábra). A rövidebb („A”) konstrukció alkalmazása esetén az rs4273545 G, illetve T allélt tartalmazó promoter konstrukció aktivitása között nem tapasztaltunk jelentős különbséget [VI].

A hosszabb, rs148797429 in/del polimorfizmust is magában foglaló („B”) konstrukció elemzése alapján több következtetés is levonható. Úgy tűnik, hogy – hasonlóan az SNP-hez – az rs148797429 hosszúság variációnak önmagában nincs hatása a képződő mRNS mennyiségére, mivel akár G, akár T alléllal együtt az inzerciót, illetve a deléció tartalmazó konstrukciók (23. ábra 3. (in–G) és 5. (del–G), illetve 4. (in–T) és 6. (del–T) oszlopok) hasonló relatív luciferáz aktivitást eredményeztek. Érdekes ugyanakkor, hogy bár a korábbi megfigyelések szerint a –233. nukleotidtól 5' irányban lévő szakasz nem szükséges a gén kifejeződéséhez [296], eredményeink alapján ez a – rs148797429 polimorfizmust is magában foglaló – régió mégis hozzájárul a transzkripció aktivitásának szabályozásához. Ebben a rendszerben ugyanis az rs4273545

T allél jelenléte 2,5–4× magasabb mennyiségű mRNS képződéséhez vezetett a G változathoz képest [VI], és ez a különbség az rs148797429 del variáns jelenlétében (statisztikailag nem szignifikáns mértékben) kifejezettebbnek mutatkozott (23. ábra).

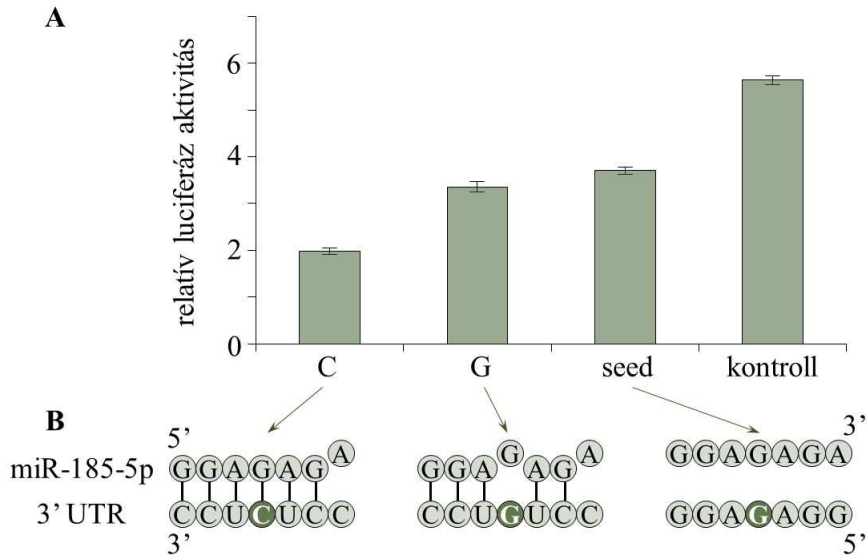


23. ábra. A *WFS1* promoterében lévő rs148797429 és rs4273545 SNP-k hatásának elemzése. A rövidebb („A”) konstrukció csak az rs4273545 G/T SNP-t tartalmazta, a hosszabb („B”) rendszerben az SNP mellett a rs148797429 in/del polimorfizmus is jelen volt. Kontroll: üres pGL3 Basic vektor.

5.3.2.2 A *WFS1* 3' UTR-ében lévő SNP-k funkcionális elemzése

A képződő (illetve jelen lévő) fehérje mennyiségét a transzkripció aktivitásán túl számos további tényező befolyásolja. Következő szint a transláció szabályozása, melynek egyik lényeges komponense a miRNS-ek révén megvalósuló reguláció. A *WFS1* gén 3' UTR-ében két polimorfizmus (rs9457 C/G és rs1046322 A/G) molekuláris hatását elemeztük luciferáz riporter konstrukció alkalmazásával. Ez a két SNP – saját eredményeink alapján – asszociációt mutatott az 1-es, illetve 2-es típusú cukorbetegséggel, valamint az agresszióval [IV, V]. Szekvenca elemzés, illetve a PolymiRTS adatbázis [297] alapján az rs9457 SNP a miR-185, az rs1046322 polimorfizmus pedig a miR-668 kötőhelyében található, ugyanakkor a polimorfizmusok tényleges biológiai hatását korábban nem elemezték. Az SNP-k molekuláris funkciójának vizsgálata relevánsnak tűnt abból a szempontból is, hogy az rs9457 SNP a gén számos további polimorfizmusával magas LD-t mutatott (13. ábra) – köztük olyanokkal is, melyek szakirodalmi adatok alapján cukorbetegséggel asszociációt mutattak, ugyanakkor molekuláris funkciójuk nem volt

tetten érhető [94]. Érdekes, hogy az rs1046322 SNP és a gén általunk elemzett többi polimorfizmusa között ezzel szemben nem figyelhető meg magas LD.



24. ábra. A *WFS1* 3' UTR-ében lévő rs9457 és a miR-185 kötődésének vizsgálata. *A:* Az egyes allélokat, illetve a „seed mutáns” tartalmazó konstrukciók relatív luciferáz aktivitása. A „kontroll” konstrukcióban a miR-185-5p kötőhely nincs jelen. *B:* A miR185-5p „seed” szekvencia és a *WFS1* 3' UTR kapcsolódása. A sötétzöld háttérű pozíciók az SNP helyét jelölik. C allél esetén a miR-185-5p „seed” szekvenciájának 6, G variáns jelenlétekor csupán 5 nukleotidja komplementer a *WFS1* mRNS-sel.

A diabetes mellitus-szal asszociációt mutató rs9457 polimorfizmus miRNS kötődését befolyásoló szerepét *in vitro* luciferáz riporter rendszer alkalmazásával elemeztük. A miR-185 hajtú szerkezetű pre-miRNS előalakjából két különböző érett miRNS keletkezhet. A miR-185-5p különböző sejtvonalak elemzése alapján kb. 200× magasabb expressziós szintet mutat [298], ez az az érett miRNS, amely a *WFS1* 3' UTR-ében az rs9457 SNP régiójához kötődik. A miRNS 5' végén elhelyezkedő „seed” szekvenciából (2–8. nukleotidok) csak az első hat komplementer a *WFS1* 3' UTR szakaszával rs9457 C allél jelenléte esetén (24. ábra), ugyanakkor a miRNS legelső nukleotidja szintén bázispárt hoz létre az mRNS-sel. A *WFS1* gén teljes 3' UTR-ét ismert genotípusú személy genomi DNS-éről PCR-rel felsokszoroztuk, majd pMIR-Report vektorba klónoztuk a luciferázt kódoló szakasz mögé. A G allélt hordozó konstrukciót irányított mutagenézissel készítettük el. Létrehoztunk emellett egy ún. „seed mutáns” rekombináns

DNS-t is, melyben a miR-185 „seed” szekvenciájának kötőhelyét teljes egészében megszüntettük oly módon, hogy a megfelelő nukleotidokat azok komplementerére cseréltük. Ennek következtében a miR-185 „seed” régiójának egyetlen nukleotidja sem tudott bázispárt képezni a DNS-konstrukcióról átíródó mRNS-sel. A kísérleteink során kontrollként alkalmaztunk emellett egy olyan konstrukciót, amely egy azonos hosszúságú, de eltérő bázissorrendű inzertet tartalmazott, amelyhez – a nukleinsavak szekvenciája alapján – a miR-185 nem tudott bekötődni [IV].

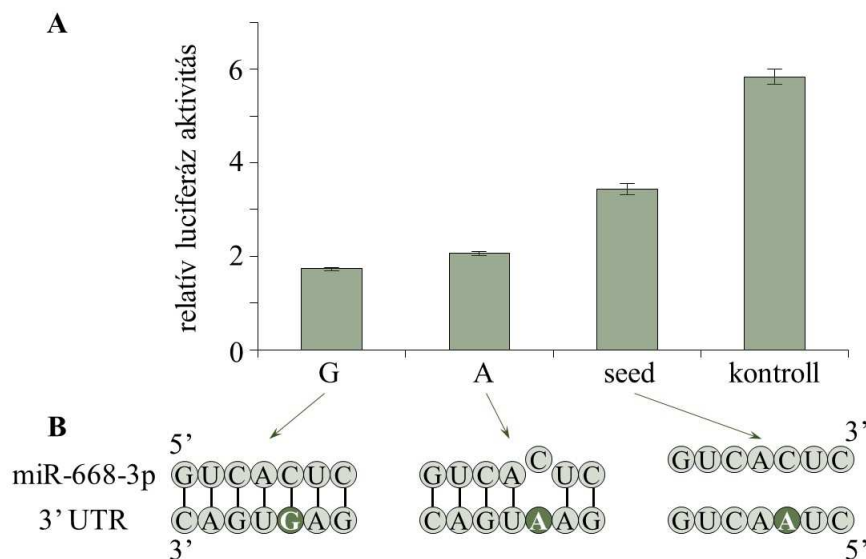
Legalacsonyabb relatív luciferáz aktivitást az rs9457 SNP C allélját hordozó konstrukció mutatott, ez az érték csupán 35%-a volt a kontroll konstrukció aktivitásnak (24. ábra). A mindössze egyetlen bázisban eltérő G variáns 1,7-szeres, statisztikailag szignifikáns enzimaktivitás emelkedéshez vezetett. Figyelemre méltó, hogy ez az érték lényegében azonos volt a „seed mutáns” konstrukció aktivitásával, melyhez a miR-185-5p 2–8. nukleotidjai közül egy sem tudott bekötődni [IV].

Korábbi kutatások eredményei alapján a miR-185 kapcsolatban áll az onkogenezissel: több, különböző mechanizmus révén tumorszupresszor hatást fejt ki, szerepe számos daganatos kórképben (emlődaganat, prosztatatarák, rosszindulatú kolorektális megbetegedés) [299-301], valamint policisztás ovárium szindrómában [302] is kimutatható. Ezek a megfigyelések rávilágítanak arra, hogy a miR-185 hozzájárul a sejtproliferáció, illetve az apoptózis szabályozásához. Ezek a folyamatok a hasnyálmirigy β -sejtjeinek fejlődése és működése során is kulcsfontosságúak. Gesztációs diabetes-ben, T2DM-ben, valamint streptozotocinnal indukált diabeteses egerekben a miR-185 szintjének csökkenése figyelhető meg. A miR-185 kötődik a SOCS3 (suppressor of cytokine signalling 3) 3' UTR-éhez, ami a Jak / STAT jelátviteli úton keresztül hatással van az inzulin szignál kialakulására májban és zsírszövetben [303, 304]. A miRNS-ek révén megvalósuló szabályozás komplexitását és hálózatos jellegét jól példázza, hogy ugyanez a miRNS egy másik, a cukorbetegséggel teljesen eltérő molekuláris mechanizmus révén kapcsolatban álló fehérje, a wolframin kifejeződésére is hatással van. A miR-185 és a *WFS1* interakciója több szempontból is sajátos. Ismert, hogy a miRNS kötődésében az 5' végen elhelyezkedő 7–8 nukleotid hosszúságú „seed” szekvencia játszik döntő szerepet, amit azonban a két nukleinsav további szakaszainak komplementaritása befolyásolhat. Megfigyelték ugyanakkor azt is, hogy a „seed” szekvencia tökéletes bázispárosodása önmagában nem mindig elegendő: nem okozza minden esetben a génkifejeződés jelentős változását, és ismert az is, hogy a 3'

UTR-ben lévő AU-gazdag szekvenciák, valamint a miRNS-kötőhely 3' UTR-en belüli lokalizációja szintén befolyásolhatja a szabályozó hatás megnyilvánulását [305]. Ezen már önmagában is sokváltozós rendszert az SNP-k tovább modulálják: az rs9457 SNP viszonylag kifejezett hatásához vélhetően hozzájárul, hogy a polimorfizmus egy rövid „seed” kötőhely közepén helyezkedik el. Mindezek alapján érthető, hogy bár az *in silico* predikció értékes kiindulást jelent a miRNS-ek hatásának elemzése során, a szabályozó hatás feltérképezésében az *in vitro* és *in vivo* mérések elengedhetetlenül fontosak.

Az rs1046322 A/G SNP a *WFS1* gén 3' UTR-ében a kódoló szakaszhoz 351 bp-ral közelebb helyezkedik el, vizsgálataink alapján az agresszió endofenotípussal és az 1-es típusú cukorbetegséggel mutat asszociációt [IV, V]. A polimorfizmus körüli génszakasz a miR-668-3p kötőhelye: a régió G allél jelenléte esetén tökéletesen komplementer a miR-668-3p 5' végi 8 nukleotidjával. A polimorfizmus miRNS-kötődésre, illetve ennek révén a transláció aktivitására gyakorolt hatását luciferáz riporter rendszer alkalmazásával elemeztük. A mérések során ugyanazt a riporter konstrukciót használtuk, melyet az rs9457 SNP vizsgálatakor, mely a *WFS1* teljes 3' UTR-ét tartalmazta, így az SNP-k egymáshoz, illetve a kódoló szakaszhoz viszonyított elhelyezkedése a természetes lokalizációval megegyezett. Ebben az esetben is irányított mutagenézissel készítettük el az A allélt, valamint a „seed mutációt” tartalmazó konstrukciókat (Függelék S1. táblázat), ez utóbbiban a miR-668 2–8. nukleotidjának kötőhelyét változtattuk meg, így a miRNS és az mRNS ezen szakaszon nem tudott kötődni egymáshoz (25. ábra).

Eredményeink alapján a miR-668 kötődik a *WFS1* 3' UTR-éhez: a tökéletes kötőhelyet (G allélt) tartalmazó konstrukció 50,7%-os ($p < 0,001$) luciferáz aktivitást mutatott a „seed mutáns” képest. A csupán az SNP-nek megfelelő egyetlen nukleotidban eltérő konstrukciók között kisebb különbség volt megfigyelhető, ami összefüggésben állhat azzal, hogy ebben az esetben a G allél 8 nukleotid hosszúságú komplementer szakaszt eredményez, s ebben a szekvenciában az SNP kissé excentrikusan (6. nukleotid) helyezkedik el (25. ábra). A vizsgálatot a pMIR-Report mellett pGL3-Control vektor alkalmazásával is megismételtük, ebben a rendszerben alapvetően hasonló tendenciák voltak megfigyelhetők, ugyanakkor az A, illetve a G allélt tartalmazó konstrukció között jelentősebb, közel 50%-os, statisztikailag szignifikáns ($p < 0,05$) különbség volt kimutatható [V].



25. ábra. A *WFS1* gén rs1046322 SNP miR-668 kötődésére kifejtett hatásának elemzése.

A: Az egyes allélokat, illetve a „seed mutánst” tartalmazó konstrukciók relatív luciferáz aktivitása. A „kontroll” konstrukcióban a miR-668-3p kötőhely nem található meg. *B:* A miR-668-3p „seed” szekvencia és a *WFS1* 3' UTR kapcsolódása. A sötétzöld háttérű betűk az rs1046322 SNP pozícióját mutatják. G allél esetén a miR-668-3p „seed” szekvenciája tökéletesen komplementer a *WFS1* 3' UTR-rel.

Érdekes, hogy eredményeink alapján három egymástól többé-kevésbé eltérő molekuláris mechanizmus (promoter polimorfizmusok, miR-185, illetve miR-668 kötődése) is feltételezhető, mely a wolframin szintjének modulálása révén szerepet játszhat a T1DM és a T2DM örökletes hajlamának meghatározásában, és bizonyos, hogy a *WFS1* gén kifejeződését emellett még számtalan tényező több különböző szinten befolyásolja. A wolframin alapvető szerepet játszik az ER-stressz szabályozásában. Számos hatás, többek között az akár étkezést követően, fiziológiai körülmények között kialakuló fokozott inzulin termelés a hibásan feltekeredett fehérjék mennyiségének megemelkedéséhez vezethet, ami ER-stresszt okoz. Ezen állapot helyreállítására az endoplazmás retikulum a selejtfehérje válasszal (UPR) reagál. A három fő (IRE1, PERK és ATF6) útvonal végső soron chaperonok átíródása, a selejtfehérjék helyes feltekeredése vagy lebontása, illetve a transzláció gátlása révén az ER túltöltését megszünteti, ugyanakkor súlyos károsodás esetén apoptózist indukál. A wolframin mindhárom UPR-útvonallal összefüggésben áll: ER-stresszt okozó hatások a *WFS1* gén transzkripcióját fokozzák, a wolf-

ramin csökkent mennyisége vagy hiánya esetén pedig az ER-stressz markerek szintjének emelkedése (BiP, Chop, hasított XBP-1) figyelhető meg [85, 86]. A β -sejtek normál körülmények között is – időről időre – nagy mennyiségű inzulint termelnek. Inzulínrezisztencia esetén ez tovább fokozódik: a β -sejtek az egyre emelkedő inzulinszükségletet idővel nem tudják biztosítani, ami krónikus ER-stresszhez, és végül sejtpusztuláshoz vezet. Érthető ez alapján a wolframin és a cukorbetegség kapcsolata: az ER-homeosztázis fenntartásában a wolframin fontos szerepet játszik, és ez az egyensúly igen érzékeny. A wolframin szintjének csökkenése [306] mellett mennyiségének emelkedése is kedvezőtlen lehet az UPR lecsengésének késése miatt [85], így belátható, hogy a *WFS1* kifejeződésének polimorf, miRNS-ek révén megvalósuló módosulása a kényes szimmetria megbillentése révén a T1DM és a T2DM molekuláris komponense lehet.

6 KITEKINTÉS

Az *információ* fogalma, elérhetősége, az adatok kezelésének és feldolgozásának lehetőségei szinte követetetlen tempóban fejlődtek az elmúlt évtizedek során. Vonatkozik ez a *genetikai* információra is, jól példázza ezt, ha a Humán Genom Program 14 évig tartó munkáját összevetjük az új generációs szekvenálási technikákkal, melyek segítségével kb. 5–50 milliárd bázispárnyi szekvencia meghatározása ma már órák vagy legfeljebb napok alatt megvalósítható. A „molekuláris biológia” két-három évtizeddel ezelőtt a genetikai információ kifejeződésének elméletét, a centrális dogma alapján a replikáció, transzkripció, transláció általános lépéseit jelentette elsősorban. A módszerek robbanásszerű fejlődése, valamint a DNS, illetve az RNS szerkezetének, működésének és egyedi variabilitásának megismerése ugyanakkor mára lehetővé tette, hogy ez a tudás és eszköztár a betegségek molekuláris mechanizmusának megértésétől a diagnosztikán és szűrésen át a terápiáig a gyakorlati klinikai orvoslás részévé váljon.

Mi sem bizonyítja ezt jobban, mint az elmúlt évek világméretű egészségügyi problémája. A koronavírus járvány kapcsán a „PCR-teszt” mindennapos fogalommá vált. A kórokozók specifikus, szenzitív és akár mennyiségi jellegű kimutatása azonban nem újkeletű, és természetesen csaknem bármilyen vírus, illetve baktérium detektálására alkalmazható [307]. Bár az alapelv – a specifikus primerek, illetve próbákkal történő amplifikáció – egyszerű, mégis a SARS-CoV-2 vírus kimutatására többszáz különböző real-time PCR rendszer érhető el, melyek érzékenységükben és pontosságukban valamelyest eltérnek egymástól [308]. További lehetőségeket kínálnak az antigén-gyorsteszték, melyek nyilvánvaló előnye egyszerű, eszközt nem igénylő kivitelezhetőségük, az alkalmazott antitesttől függően azonban ezek szenzitivitása még szélesebb skálán mozog, és rendszerint nem éri el a real-time PCR alapú módszerek érzékenységét [309].

A molekuláris szintű klinikai vizsgálatok ugyanakkor a fertőzések kimutatásánál messze általánosabbak, és minden ember életében a megszületést követően csupán egy-két napot váratnak magukra. 2022. január 1. óta Magyarországon – az idén bevezetett cisztás fibrózis újszülöttkori szűrésével – 27 örökletes anyagcsere betegség vizsgálata kötelező. A technikai lehetőségek fejlődése ezen a téren is alapvető volt: a tandem tömegspektrometria (MS-MS) bevezetése tette lehetővé, hogy a klasszikusan vizsgált néhány kórkép (fenilketonuria, hypothyreosis, biotinidáz hiány, galactosaemia) mellett ma már számos aminosav anyagcsere

zavar, a zsírsavak oxidációjának különféle rendellenességei, valamint az urea-ciklus és a szerves savak metabolizmusának bizonyos hibái is a rutin szűrés részét képezik. Olyan betegségek ezek, melyeknél a korai felismeréssel és kezeléssel a gyerekek életminősége szignifikánsan javítható, mivel a különböző károsodások, elsősorban a középsúlyos–súlyos neuropszichológiai diszfunkciók és a mentális retardáció kialakulása jelentősen csökkenthető. 2005-ben az American College of Medical Genetics 29 rendellenesség esetén tartotta célszerűnek és kivitelezhetőnek a rutinszerű újszülöttkori szűrés elvégzését, a ténylegesen vizsgált betegségek száma az egyes országokban kisebb-nagyobb eltérést mutat [310]. Bár a tömegspektrometrián alapuló elemzés a legtöbb esetben megfelelően hatékony és pontos, rendszerint ismertek a betegségek háttérében álló mutációk, ami a DNS-szekvenciát elemző szűrővizsgálatok lehetőségét kínálja. Figyelembe kell azonban venni, hogy egy-egy betegség kialakulását nem ritkán egy génen belül számos genetikai variáció okozhatja. A célzott real-time PCR, illetve allélspecifikus amplifikáción alapuló eljárások helyett emiatt bizonyos esetekben a hagyományos vagy az új generációs szekvenálás nyújt csak megfelelő információt: a jelenleg a cisztikus fibrózis szűrésére elérhető NGS-alapú rendszer 324 mutáció egyidejű elemzését teszi lehetővé [311].

A molekuláris genetikai technikák azonban még ennél is korábban, már jóval a születés előtt helyt kaphatnak a diagnosztikában. A terhességek kb. 3%-ában mutatható ki a rutin ultrahang vizsgálatok során valamilyen – izolált vagy több szervet érintő – eltérés, melyek interpretálása gyakran nem csak a család, hanem a szakemberek számára is kihívást jelent. A genetikai vizsgálatok elterjedésének egyebek mellett lendületet adott annak felismerése, hogy a negyedik gesztációs héttől kezdve az anyai keringésben magzati eredetű DNS („cell-free” – cfDNS) van jelen, így az elemzések invazív mintavétel nélkül elvégezhetők. Noha az anyai és a magzati eredetű DNS között leírtak apróbb eltéréseket, ezek egymástól nem különíthetők el, a magzati DNS mennyisége mégis néhány hét elteltével már elegendő a genetikai elemzések elvégzéséhez [312]. Real-time PCR alapú eljárással legkorábban a magzat nemének meghatározását végezték el hemofiliával, Duchenne-féle izomdisztrófiával, illetve kongenitális adrenális hiperpláziával érintett családokban, mivel ez az információ a további (diagnosztikus és terápiás) beavatkozások szükségességét ezen kórképek esetén alapvetően meghatározza. A későbbiekben az eljárást sikeresen alkalmazták olyan betegségek kimutatása során, melyeket *de novo* vagy apai eredetű mutációk okoznak, ebben az esetben ugyanis a nagy mennyiségben jelenlévő anyai DNS az elemzést kevésbé zavarja, mivel az adott genetikai variáció abban ismertén nincs jelen. A PCR,

illetve real-time PCR alapú eljárások – a módszer elvének megfelelően – célzottan egy adott mutáció jelenlétét vizsgálják: ez megfelelő megközelítés pozitív családi anamnézis esetén, ugyanakkor nem deríthet fényt az ultrahang vizsgálattal kimutatott strukturális rendellenesség okára [313]. Az exom, illetve a genom szekvenálás számos ilyen esetben informatív megoldást jelenthet, a háttérben álló genetikai rendellenesség több mint 800 magzatot magában foglaló vizsgálatok eredményei alapján az esetek 8,5–10%-ában azonosítható volt [314, 315]. Az NGS-alapú DNS-kvantifikálás elérhető a terhesség alatt az aneuploidiák hatékony, pontos és kockázatmentes szűrésére. A NIFTY (Verifi[®]) teszt érzékenysége és specifitása megegyezik a 0,5–1% vetélés kockázattal járó chorion boholy biopsziával nyert minta analízisével: megbízhatóan alkalmas az autoszomális (Down-, Edwards-, Patau-szindróma) és a nemi kromoszómák (Turner-, Klinefelter-szindróma) számbeli rendellenességeinek kimutatása mellett a magzat nemének meghatározására is [316].

Napjainkban az örökletes rendellenességek kimutatására akár már a terhesség előtt is sor kerülhet: a preimplantációs genetikai vizsgálat azonban számos technikai és etikai kérdést felvet. A preimplantációs fejlődés 3–5. napján nyert minta teljes genom amplifikációt követően ma már nem csak multiplex PCR, hanem SNP-chip- vagy NGS-alapú elemzést is lehetővé tesz, így a monogénes kórképek és a kromoszómák számának, illetve szerkezetének rendellenességei egyaránt kimutathatók. Az eljárást jelenleg legfőképp gyakori, monogénes kórképek (cisztás fibrózis, Huntington-betegség, örökletes hemoglobinopátiák, Duchenne-féle izomdisztrófia) kizárása céljából használják. A módszer alkalmazásának törvényi szabályozása a különböző országokban igen széles skálán mozog: az eljárást Máltán és Boszina-Hercegovinában egyáltalán nem engedélyezik, az Egyesült Királyságban és Franciaországban szigorúan szabályozzák, ugyanakkor az USA-ban a módszernek nincsenek törvényi korlátai [317].

Legyen szó valamely kórokozó kimutatásáról, vagy egy monogénes betegséget okozó mutáció – bármely életkorban célzott vagy genom szintű eljárással történő – azonosításáról, a molekuláris biológiai eljárások ilyen irányú használata az orvosi alkalmazások csupán egyik szelétét jelenti. Egy másik vetület az az óriási kirakó, ami a komplex kórképek genetikai, illetve molekuláris szintű hátterének feltárását jelenti, és amelynek az elmúlt évtizedek alatt egyre több eleme a helyére került; reményeink szerint ehhez valamelyest kutatócsoportunk munkái is hozzájárultak. Ismereteink olyan szintre jutottak el, ami lehetővé teszi, hogy a molekuláris biológia

eszköztárát a betegségek elméleti hátterének feltárásán túl a diagnosztika és a kezelés szolgálatába állítsuk. Létezik néhány olyan gyakori, komplex öröklődésű betegség, melynek kialakulásában viszonylag alacsony allélfrekvenciával rendelkező mutáció is szerepet játszik, és ez a variáció a betegség kialakulásának kockázatát önmagában jelentősen fokozza. Ilyen például a *HNF1A* transzkripció faktor génjében lévő aminosavcsere okozó p.E508K mutáció, ami a T2DM kockázatát 5-szörösére emeli. A legtöbb poligénes kórkép esetén azonban ilyen kiemelkedő genetikai variáció nem azonosítható, hanem számos, egyenként igen kis hatású polimorfizmus együttesen járul hozzá a betegség kialakulásához. A DNS szekvenciájának meghatározása, illetve az SNP-k genotipizálása azonban ma már olyan hatékonyan megvalósítható, hogy bizonyos komplex kórképek esetén a betegség kialakulásának rizikója a monogénes kórképekkel összemérhető pontossággal meghatározható. A poligénes rizikó pontszám (polygenic risk score: PRS vagy genome-wide polygenic score: GPS) kiszámítása pitvarfibrilláció, koszorúérbetegség, T2DM és gyulladós bélbetegség (Crohn-betegség és colitis ulcerosa) esetén több mint 6 millió SNP alapján történik, emlőrák esetén ez a szám jóval alacsonyabb (5218), míg az Alzheimer-kór kialakulása 39 polimorfizmus alapján hatékonyan jósolható [318]. Érthető ugyanakkor, hogy egyetlen teljes genomra kiterjedő szekvencia elemzés adataiból bármely betegség rizikója akár a születést követően meghatározható. A UK Biobank adatai, valamint mintái alapján a populáció 19,8%-ában a koszorúérbetegség, a pitvarfibrilláció, a T2DM, a gyulladós bélbetegség, valamint az emlőrák közül legalább egy kialakulásának kockázata több mint háromszoros. Ez a tudás a célzott szűrés és prevenció lehetőségét kínálja, ugyanakkor komoly etikai kérdéseket is felvet [319].

A tényleges génműködés elemzése további lépést jelent a fenotípus felé, ennek megfelelően a szakirodalomban megjelenik a genomszintű génexpresszió vizsgálaton alapuló transzkripció rizikó pontszám fogalma is [320], emellett mind elméleti, mind gyakorlati szempontból jelentős a miRNS-ek betegségekkel való kapcsolatának elemzése. A sejtekből aktívan vagy passzívan kijutó miRNS-ek fehérjékhez, illetve lipoproteinekhez kötve igen stabilak, így – technikai szempontból is – megfelelő biomarkerek lehetnek [321]. Emellett működésük több esetben egyértelműen összefügg egyes kórképek kialakulásával. A miR-375 a foszfoinozítid-dependens kináz 1, a miR-143 pedig a protein kináz-B kifejeződését befolyásolja, de emellett számos további miRNS ismert, melyek részt vesznek a MAP-kináz kaszkád tagjainak, valamint

ciklinek és ciklinfüggő kinázok szabályozásában, így közvetlen kapcsolatba hozhatók daganatos kórképek kialakulásával [322, 323]. Érdekes, hogy a miRNS-ek szerepe a major depresszió és más betegségek – pl. sclerosis multiplex, migrén – komorbiditásával kapcsolatban is tetten érhető [324, 325]. A miRNS-ek és különböző kórképek asszociációjának feltérképezése kezelési célpontokat is kínál. A miR-23B szintje összefügg a sugárkezelés hatékonyságával hasnyálmirigyrákban, a miR-29c fokozott expressziója pedig gemcitabin terápia hatékonyságát növeli, ami a rendkívül rossz prognózisú betegség kezelésében lényeges előrelépést jelenthet [326]. A miR-92a, a miR-129, a miR-491 és a miR-20a részt vesz a tumorsejtek apoptózisának szabályozásában, ezen folyamatok által kínált kezelési lehetőségeket kolorektális tumorok terápiaja során vizsgálták [52], de a miRNS-ek, illetve az általuk szabályozott folyamatok több más kórkép – asztma, szívinfarktus, szívelégtelenség [327, 328] – esetén is a kezelés célpontjai lehetnek. De a terápia akár a DNS szintjén is történhet. A hiányzó gén adenovírus eredetű vektorral történő bejuttatása a cisztikus fibrózis kezelése során ismert eljárás [329]. De emellett a CRISPR/Cas9 technológia fejlődésének köszönhetően ma már megvalósítható a DNS-ben lévő hiba tényleges kijavítása is [330], ez a megközelítés alkalmazása azonban elvi szinten is csak a monogénes kórképek terápiaja esetén jön szóba, amikor a betegség kialakulásáért egyetlen mutáció felel.

Három milliárd bázispár: „A-k”, „C-k”, „G-k” és „T-k” hosszú sora. Négy jel – megszámlálhatatlanul sokféle jelentés. Mindenkiben – több mint 99%-ban – azonos: ez az ember biológiai kódja. Mégis minden ember más, mert egyes szakaszok másként vannak megírva. Itt eltér egy „betű”, ott száz, máshol százezer. Olykor a százezer módosulását észre sem vesszük, az egy megváltozása pedig esetenként végzetes. *„Ilyen az ember. Egyedüli példány.”* (Kosztolányi Dezső)

7 KÖSZÖNETNYILVÁNÍTÁS

Szeretnék köszönetet mondani

- Sasvári Máriának, aki már orvostanhallgató koromban megmutatta, hogy egy gélkép milyen szép lehet, és aki utolérhetetlen módon irányította munkámat, illetve kutatócsoportunk munkáját,
- Guttman Andrásnak, aki munkámat PhD-hallgató korom óta feltartóztathatatlan lendülettel támogatja, s aki módszertani ismereteimet több területen megalapozta,
- Kovács-Nagy Rékának, Molnár Zsuzsannának és Németh Nórának, akik a bemutatott kísérleti munka jelentős részét elvégezték, de ami ennél is fontosabb, hogy velük azóta is mindmáig együtt végezhetem munkámat, és minden téren számíthatok rájuk,
- Keszler Gergelynek, aki szinte hihetetlen elméleti tudásával támogatja csoportunkat,
- Bánlaki Zsófiának, aki kimagasló elméleti és gyakorlati ismereteivel járul hozzá kutatásainkhoz,
- Németh Helgának, aki mindig gondoskodik arról, hogy a laborban minden rendelkezésre álljon,
- Szántai Eszternek és Héjjas Krisztinának, akik jelentős részt vállaltak több genotipizáló eljárás beállításában,
- munkacsoportunk, valamint a Biokémiai és Molekuláris Biológiai Intézet Molekuláris Biológiai Tanszék korábbi és jelenlegi tagjainak a közös munkáért,
- Müllner Nándornak, akitől nemcsak a kutatás és az oktatás, hanem az élet egyéb területeire vonatkozóan is mindig bölcs és megnyugtató útmutatást kapok,
- Tanszékünk korábbi és jelenlegi vezetőinek, akik munkámat mindig támogatták,
- Sahin-Tóth Miklósnak, aki bostoni tanulmányutam során vezette munkámat,
- A vizsgálatokban résztvevő személyeknek,
- A kutatásokban résztvevő munkatársaknak,
- Édesapámnak; feleségemnek, Évának és kisfiaimnak, Máténak és Samunak, akik mindig bíznak és hisznek bennem.

8 SAJÁT KÖZLEMÉNYEK

8.1 A doktori értekezés alapjául szolgáló

közlemények

- I. Szantai E, Szilagyai A, Guttman A, Sasvari-Szekely M, Ronai Z: **Genotyping and haplotyping of the dopamine D4 receptor gene by capillary electrophoresis.** *J Chromatogr A* 2004, **1053**(1-2):241-245.
- II. Szantai E, Szmola R, Sasvari-Szekely M, Guttman A, Ronai Z: **The polymorphic nature of the human dopamine D4 receptor gene: a comparative analysis of known variants and a novel 27 bp deletion in the promoter region.** *BMC Genet* 2005, **6**(39)
- III. Hejjas K, Vas J, Kubinyi E, Sasvari-Szekely M, Miklosi A, Ronai Z. **Novel repeat polymorphisms of the dopaminergic neurotransmitter genes among dogs and wolves.** *Mamm Genome* 2007, **18**(12):871-879.
- IV. Elek Z, Nemeth N, Nagy G, Nemeth H, Somogyi A, Hosszufalusi N, Sasvari-Szekely M, Ronai Z: **Micro-RNA Binding Site Polymorphisms in the WFS1 Gene Are Risk Factors of Diabetes Mellitus.** *PLoS One* 2015, **10**(10)
- V. Kovacs-Nagy R, Elek Z, Szekely A, Nanasi T, Sasvari-Szekely M, Ronai Z: **Association of aggression with a novel microRNA binding site polymorphism in the wolf-ramin gene.** *Am J Med Genet B Neuropsychiatr Genet* 2013, **162B**(4):404-412.
- VI. Elek Z, Denes R, Prokop S, Somogyi A, Yowanto H, Luo J, Souquet M, Guttman A, Ronai Z: **Multicapillary gel electrophoresis based analysis of genetic variants in the WFS1 gene.** *Electrophoresis* 2016, **37**(17-18):2313-2321.
- VII. Szilagyai A, Blasko B, Szilassy D, Fust G, Sasvari-Szekely M, Ronai Z: **Real-time PCR quantification of human complement C4A and C4B genes.** *BMC Genet* 2006, **7**(1)
- VIII. Ronai Z, Kovacs-Nagy R, Szantai E, Elek Z, Sasvari-Szekely M, Faludi G, Benkovits J, Rethelyi JM, Szekely A: **Glycogen synthase kinase 3 beta gene structural variants as possible risk factors of bipolar depression.** *Am J Med Genet B Neuropsychiatr Genet* 2014, **165B**(3):217-222.
- IX. Ronai Z, Sasvari-Szekely M, Guttman A: **Miniaturized SNP detection: quasi-solid-phase RFLP analysis.** *Biotechniques* 2003, **34**(6):1172-1173.
- X. Ronai Z, Barta C, Sasvari-Szekely M, Guttman A: **DNA analysis on electrophoretic microchips: effect of operational variables.** *Electrophoresis* 2001, **22**(2):294-299.
- XI. Szantai E, Kiraly O, Nemoda Z, Kereszturi E, Csapo Z, Sasvari-Szekely M, Gervai J, Ronai Z: **Linkage analysis and molecular haplotyping of the dopamine D4 receptor gene promoter region.** *Psychiatr Genet* 2005, **15**(4):259-270.

- XII. Kovacs-Nagy R, Sarkozy P, Hu J, Guttman A, Sasvari-Szekely M, Ronai Z: **Haplotyping of putative microRNA-binding sites in the SNAP-25 gene.** *Electrophoresis* 2011, **32**(15):2013-2020.
- XIII. Banlaki Z, Elek Z, Nanasi T, Szekely A, Nemoda Z, Sasvari-Szekely M, Ronai Z: **Polymorphism in the serotonin receptor 2a (HTR2A) gene as possible predisposal factor for aggressive traits.** *PLoS One* 2015, **10**(2)
- XIV. Nemeth N, Kovacs-Nagy R, Szekely A, Sasvari-Szekely M, Ronai Z: **Association of impulsivity and polymorphic microRNA-641 target sites in the SNAP-25 gene.** *PLoS One* 2013, **8**(12)

8.2 A doktori értekezéshez kapcsolódó közlemények

- XV. Ronai Z, Lippai Z, Elek Z, Somogyi A: **[Investigation of the genetic background of complex diseases].** *Orv Hetil* 2018, **159**(31):1254-1261.
- XVI. Toth-Petroczy A, Szilagyi A, Ronai Z, Sasvari-Szekely M, Guttman A: **Validation of a tentative microsatellite marker for the dopamine D4 receptor gene by capillary gel electrophoresis.** *J Chromatogr A* 2006, **1130**(2):201-205.
- XVII. Ronai Z, Guttman A, Nemoda Z, Staub M, Kalasz H, Sasvari-Szekely M: **Rapid and sensitive genotyping of dopamine D4 receptor tandem repeats by automated ultrathin-layer gel electrophoresis.** *Electrophoresis* 2000, **21**(10):2058-2061.
- XVIII. Ronai Z, Barta C, Guttman A, Lakatos K, Gervai J, Staub M, Sasvari-Szekely M: **Genotyping the -521C/T functional polymorphism in the promoter region of dopamine D4 receptor (DRD4) gene.** *Electrophoresis* 2001, **22**(6):1102-1105.
- XIX. Ronai Z, Szantai E, Szmola R, Nemoda Z, Szekely A, Gervai J, Guttman A, Sasvari-Szekely M: **A novel A/G SNP in the -615th position of the dopamine D4 receptor promoter region as a source of misgenotyping of the -616 C/G SNP.** *Am J Med Genet B Neuropsychiatr Genet* 2004, **126B**(1):74-78.
- XX. Szilagyi A, Blasko B, Ronai Z, Fust G, Sasvari-Szekely M, Guttman A: **Rapid quantification of human complement component C4A and C4B genes by capillary gel electrophoresis.** *Electrophoresis* 2006, **27**(8):1437-1443.
- XXI. Ronai Z, Guttman A, Nemoda Z, Gervai J, Sasvari-Szekely M: **Direct haplotype detection of adjacent polymorphic sites in the regulatory region of the dopamine D4 receptor (DRD4) gene.** *Electrophoresis* 2002, **23**(10):1512-1516.
- XXII. Kerekgyarto M, Nemeth N, Kerekes T, Ronai Z, Guttman A: **Ultrafast haplotyping of putative microRNA-binding sites in the WFS1 gene by multiplex polymerase chain reaction and capillary gel electrophoresis.** *J Chromatogr A* 2013, **1286**:229-234.

- XXIII. Hejjas K, Vas J, Topal J, Szantai E, Ronai Z, Szekely A, Kubinyi E, Horvath Z, Sasvari-Szekely M, Miklosi A: **Association of polymorphisms in the dopamine D4 receptor gene and the activity-impulsivity endophenotype in dogs.** *Anim Genet* 2007, **38**(6):629-633.
- XXIV. Ronai Z, Guttman A, Keszler G, Sasvari-Szekely M: **Capillary electrophoresis study on DNA-protein complex formation in the polymorphic 5' upstream region of the dopamine D4 receptor (DRD4) gene.** *Curr Med Chem* 2004, **11**(8):1023-1029.
- XXV. Ronai Z, Wang Y, Khandurina J, Budworth P, Sasvari-Szekely M, Wang X, Guttman A: **Transcription factor binding study by capillary zone electrophoretic mobility shift assay.** *Electrophoresis* 2003, **24**(1-2):96-100.
- XXVI. Nemeth N, Kerekgyarto M, Sasvari-Szekely M, Ronai Z, Guttman A: **Rapid identification of human SNAP-25 transcript variants by a miniaturized capillary electrophoresis system.** *Electrophoresis* 2014, **35**(2-3):379-384.

9 IRODALOMJEGYZÉK

1. van Dijk PJ, Ellis TH: **The Full Breadth of Mendel's Genetics.** *Genetics* 2016, **204**(4):1327-1336.
2. Gayon J: **From Mendel to epigenetics: History of genetics.** *Comptes rendus biologies* 2016, **339**(7-8):225-230.
3. **Nobel Prize in Medicine Comes to California.** *Cal West Med* 1933, **39**(6):418.
4. Muller HJ: **Thomas Hunt Morgan 1866-1945.** *Science* 1946, **103**(2679):550-551.
5. Beadle GW, Tatum EL: **Genetic Control of Biochemical Reactions in Neurospora.** *Proc Natl Acad Sci U S A* 1941, **27**(11):499-506.
6. Menezes MR, Sweasy JB: **Mouse models of DNA polymerases.** *Environ Mol Mutagen* 2012, **53**(9):645-665.
7. Gilbert W, Maxam A: **The nucleotide sequence of the lac operator.** *Proc Natl Acad Sci U S A* 1973, **70**(12):3581-3584.
8. Maxam AM, Gilbert W: **A new method for sequencing DNA.** *Proc Natl Acad Sci U S A* 1977, **74**(2):560-564.
9. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors.** *Proc Natl Acad Sci U S A* 1977, **74**(12):5463-5467.
10. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H: **Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.** *Cold Spring Harb Symp Quant Biol* 1986, **51 Pt 1**:263-273.
11. Lawyer FC, Stoffel S, Saiki RK, Myambo K, Drummond R, Gelfand DH: **Isolation, characterization, and expression in Escherichia coli of the DNA polymerase gene from Thermus aquaticus.** *J Biol Chem* 1989, **264**(11):6427-6437.
12. DeLisi C: **Meetings that changed the world: Santa Fe 1986: Human genome baby-steps.** *Nature* 2008, **455**(7215):876-877.
13. Dulbecco R: **A turning point in cancer research: sequencing the human genome.** *Science* 1986, **231**(4742):1055-1056.
14. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W *et al*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**(6822):860-921.
15. Collins FS: **Contemplating the end of the beginning.** *Genome Res* 2001, **11**(5):641-643.
16. Waterston RH, Lander ES, Sulston JE: **On the sequencing of the human genome.** *Proc Natl Acad Sci U S A* 2002, **99**(6):3712-3716.
17. Dovichi NJ, Zhang J: **How Capillary Electrophoresis Sequenced the Human Genome This Essay is based on a lecture given at the Analytica 2000 conference in Munich (Germany) on the occasion of the Heinrich-Emanuel-Merck Prize presentation.** *Angew Chem Int Ed Engl* 2000, **39**(24):4463-4468.
18. Swerdlow H, Gesteland R: **Capillary gel electrophoresis for rapid, high resolution DNA sequencing.** *Nucleic Acids Res* 1990, **18**(6):1415-1419.

19. Ruiz-Martinez MC, Berka J, Belenkii A, Foret F, Miller AW, Karger BL: **DNA sequencing by capillary electrophoresis with replaceable linear polyacrylamide and laser-induced fluorescence detection.** *Anal Chem* 1993, **65**(20):2851-2858.
20. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA *et al*: **The sequence of the human genome.** *Science* 2001, **291**(5507):1304-1351.
21. International Human Genome Sequencing C: **Finishing the euchromatic sequence of the human genome.** *Nature* 2004, **431**(7011):931-945.
22. Collins FS: **Shattuck lecture--medical and societal consequences of the Human Genome Project.** *N Engl J Med* 1999, **341**(1):28-37.
23. Collins FS, McKusick VA: **Implications of the Human Genome Project for medical science.** *JAMA* 2001, **285**(5):540-544.
24. Little J, Khoury MJ, Bradley L, Clyne M, Gwinn M, Lin B, Lindgren ML, Yoon P: **The human genome project is complete. How do we develop a handle for the pump?** *Am J Epidemiol* 2003, **157**(8):667-673.
25. Maller J, George S, Purcell S, Fagerness J, Altshuler D, Daly MJ, Seddon JM: **Common variation in three genes, including a noncoding variant in CFH, strongly influences risk of age-related macular degeneration.** *Nat Genet* 2006, **38**(9):1055-1059.
26. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G *et al*: **Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes.** *Nat Genet* 2008, **40**(5):638-645.
27. Maher B: **Personal genomes: The case of the missing heritability.** *Nature* 2008, **456**(7218):18-21.
28. Hyttinen V, Kaprio J, Kinnunen L, Koskenvuo M, Tuomilehto J: **Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study.** *Diabetes* 2003, **52**(4):1052-1055.
29. Kaprio J: **Twins and the mystery of missing heritability: the contribution of gene-environment interactions.** *J Intern Med* 2012, **272**(5):440-448.
30. Sullivan PF, Daly MJ, O'Donovan M: **Genetic architectures of psychiatric disorders: the emerging picture and its implications.** *Nature reviews Genetics* 2012, **13**(8):537-551.
31. Lundstrom S, Chang Z, Rastam M, Gillberg C, Larsson H, Anckarsater H, Lichtenstein P: **Autism spectrum disorders and autistic like traits: similar etiology in the extreme end and the normal variation.** *Arch Gen Psychiatry* 2012, **69**(1):46-52.
32. International HapMap C: **The International HapMap Project.** *Nature* 2003, **426**(6968):789-796.
33. International HapMap C, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F *et al*: **Integrating common and rare genetic variation in diverse human populations.** *Nature* 2010, **467**(7311):52-58.

34. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA: **An integrated map of genetic variation from 1,092 human genomes.** *Nature* 2012, **491**(7422):56-65.
35. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J: **Exome sequencing as a tool for Mendelian disease gene discovery.** *Nature reviews Genetics* 2011, **12**(11):745-755.
36. Kraatari M, Skarp S, Niinimäki J, Karppinen J, Mannikko M: **A Whole Exome Study Identifies Novel Candidate Genes for Vertebral Bone Marrow Signal Changes (Modic Changes).** *Spine (Phila Pa 1976)* 2017, **42**(16):1201-1206.
37. Hishida A, Nakatochi M, Tamura T, Nagayoshi M, Okada R, Kubo Y, Tsukamoto M, Kadomatsu Y, Suzuki S, Nishiyama T *et al*: **Genome-wide association study of serum prostate-specific antigen levels based on 1000 Genomes imputed data in Japanese: the Japan Multi-Institutional Collaborative Cohort Study.** *Nagoya J Med Sci* 2021, **83**(1):183-194.
38. **The Gnomad Consortium Releases First Studies of Human Genetic Variation.** *Am J Med Genet A* 2020, **182**(9):1999-2000.
39. Glusman G, Caballero J, Mauldin DE, Hood L, Roach JC: **Kaviar: an accessible system for testing SNV novelty.** *Bioinformatics* 2011, **27**(22):3216-3217.
40. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW: **The Database of Genomic Variants: a curated collection of structural variation in the human genome.** *Nucleic Acids Res* 2014, **42**(Database issue):D986-992.
41. Brandon MC, Lott MT, Nguyen KC, Spolim S, Navathe SB, Baldi P, Wallace DC: **MITOMAP: a human mitochondrial genome database--2004 update.** *Nucleic Acids Res* 2005, **33**(Database issue):D611-613.
42. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR: **ClinVar: public archive of relationships among sequence variation and human phenotype.** *Nucleic Acids Res* 2014, **42**(Database issue):D980-985.
43. Lambert SA, Abraham G, Inouye M: **Towards clinical utility of polygenic risk scores.** *Hum Mol Genet* 2019, **28**(R2):R133-R142.
44. Szalai C, Falus A, F. O: **[Orvosi genomika és bioinformatika].** *Semmelweis Egyetem, Budapest* 2012.
45. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C: **Detection of large-scale variation in the human genome.** *Nat Genet* 2004, **36**(9):949-951.
46. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M *et al*: **Large-scale copy number polymorphism in the human genome.** *Science* 2004, **305**(5683):525-528.
47. Hunt R, Sauna ZE, Ambudkar SV, Gottesman MM, Kimchi-Sarfaty C: **Silent (synonymous) SNPs: should we care about them?** *Methods Mol Biol* 2009, **578**:23-39.
48. Zhang Y, Wang D, Johnson AD, Papp AC, Sadee W: **Allelic expression imbalance of human mu opioid receptor (OPRM1) caused by variant A118G.** *J Biol Chem* 2005, **280**(38):32618-32624.

49. Hangauer MJ, Vaughn IW, McManus MT: **Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs.** *PLoS Genet* 2013, **9**(6):e1003569.
50. Pertea M: **The human transcriptome: an unfinished story.** *Genes (Basel)* 2012, **3**(3):344-360.
51. Xu J, Song J, Xiao M, Wang C, Zhang Q, Yuan X, Tian S: **RUNX1 (RUNX family transcription factor 1), a target of microRNA miR-128-3p, promotes temozolomide resistance in glioblastoma multiform by upregulating multidrug resistance-associated protein 1 (MRP1).** *Bioengineered* 2021, **12**(2):11768-11781.
52. Wang H: **MicroRNAs and Apoptosis in Colorectal Cancer.** *Int J Mol Sci* 2020, **21**(15).
53. Dayeh TA, Olsson AH, Volkov P, Almgren P, Ronn T, Ling C: **Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets.** *Diabetologia* 2013, **56**(5):1036-1046.
54. Ha M, Kim VN: **Regulation of microRNA biogenesis.** *Nat Rev Mol Cell Biol* 2014, **15**(8):509-524.
55. Ebstein RP, Novick O, Umansky R, Priel B, Osher Y, Blaine D, Bennett ER, Nemanov L, Katz M, Belmaker RH: **Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking.** *Nat Genet* 1996, **12**(1):78-80.
56. Seeman P, Guan HC, Van Tol HH: **Dopamine D4 receptors elevated in schizophrenia.** *Nature* 1993, **365**(6445):441-445.
57. Van Tol HH, Bunzow JR, Guan HC, Sunahara RK, Seeman P, Niznik HB, Civelli O: **Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine.** *Nature* 1991, **350**(6319):610-614.
58. Van Tol HH, Wu CM, Guan HC, Ohara K, Bunzow JR, Civelli O, Kennedy J, Seeman P, Niznik HB, Jovanovic V: **Multiple dopamine D4 receptor variants in the human population.** *Nature* 1992, **358**(6382):149-152.
59. Asghari V, Sanyal S, Buchwaldt S, Paterson A, Jovanovic V, Van Tol HH: **Modulation of intracellular cyclic AMP levels by different human dopamine D4 receptor variants.** *J Neurochem* 1995, **65**(3):1157-1165.
60. Jovanovic V, Guan HC, Van Tol HH: **Comparative pharmacological and functional analysis of the human dopamine D4.2 and D4.10 receptor variants.** *Pharmacogenetics* 1999, **9**(5):561-568.
61. Benjamin J, Li L, Patterson C, Greenberg BD, Murphy DL, Hamer DH: **Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking.** *Nat Genet* 1996, **12**(1):81-84.
62. Jonsson EG, Nothen MM, Gustavsson JP, Neidt H, Brene S, Tylec A, Propping P, Sedvall GC: **Lack of evidence for allelic association between personality traits and the dopamine D4 receptor gene polymorphisms.** *Am J Psychiatry* 1997, **154**(5):697-699.
63. Sullivan PF, Fifeild WJ, Kennedy MA, Mulder RT, Sellman JD, Joyce PR: **No association between novelty seeking and the type 4 dopamine receptor gene (DRD4) in two New Zealand samples.** *Am J Psychiatry* 1998, **155**(1):98-101.

64. He Y, Martin N, Zhu G, Liu Y: **Candidate genes for novelty-seeking: a meta-analysis of association studies of DRD4 exon III and COMT Val158Met.** *Psychiatr Genet* 2018, **28**(6):97-109.
65. Noble EP, Ozkaragoz TZ, Ritchie TL, Zhang X, Belin TR, Sparkes RS: **D2 and D4 dopamine receptor polymorphisms and personality.** *Am J Med Genet* 1998, **81**(3):257-267.
66. Strobel A, Wehr A, Michel A, Brocke B: **Association between the dopamine D4 receptor (DRD4) exon III polymorphism and measures of Novelty Seeking in a German population.** *Mol Psychiatry* 1999, **4**(4):378-384.
67. Halley AC, Boretsky M, Puts DA, Shriver M: **Self-Reported Sexual Behavioral Interests and Polymorphisms in the Dopamine Receptor D4 (DRD4) Exon III VNTR in Heterosexual Young Adults.** *Arch Sex Behav* 2016, **45**(8):2091-2100.
68. Michalowska-Sawczyn M, Niewczas M, Krol P, Czarny W, Rzeszutko A, Chmielowiec K, Chmielowiec J, Grzywacz A, Huminska-Lisowska K, Lachowicz M *et al*: **Associations between the dopamine D4 receptor gene polymorphisms and personality traits in elite athletes.** *Biol Sport* 2019, **36**(4):365-372.
69. Uysal MA, Sever U, Nursal AF, Yedikule Smoking Cessation Study G, Pehlivan S: **Dopamine D4 Receptor Gene Exon III VNTR Variant Influences Smoking Status in Turkish Population.** *Noro Psikiyatrs Ars* 2019, **56**(4):248-252.
70. Alshogran OY, Al-Eitan LN, Altawalbeh SM, Aman HA: **Association of DRD4 exon III and 5-HTTLPR VNTR genetic polymorphisms with psychiatric symptoms in hemodialysis patients.** *PLoS One* 2021, **16**(3):e0249284.
71. Bonvicini C, Cortese S, Maj C, Baune BT, Faraone SV, Scassellati C: **DRD4 48 bp multiallelic variants as age-population-specific biomarkers in attention-deficit/hyperactivity disorder.** *Transl Psychiatry* 2020, **10**(1):70.
72. Al-Eitan LN, Alshudaifat KM, Anani JY: **Association of the DRD4 exon III and 5-HTTLPR VNTR polymorphisms with substance abuse in Jordanian Arab population.** *Gene* 2020, **733**:144267.
73. Stolf AR, Cupertino RB, Muller D, Sanvicente-Vieira B, Roman T, Vitola ES, Grevet EH, von Diemen L, Kessler FHP, Grassi-Oliveira R *et al*: **Effects of DRD2 splicing-regulatory polymorphism and DRD4 48 bp VNTR on crack cocaine addiction.** *J Neural Transm (Vienna)* 2019, **126**(2):193-199.
74. Niimi Y, Inoue-Murayama M, Murayama Y, Ito S, Iwasaki T: **Allelic variation of the D4 dopamine receptor polymorphic region in two dog breeds, Golden retriever and Shiba.** *J Vet Med Sci* 1999, **61**(12):1281-1286.
75. Ito H, Nara H, Inoue-Murayama M, Shimada MK, Koshimura A, Ueda Y, Kitagawa H, Takeuchi Y, Mori Y, Murayama Y *et al*: **Allele frequency distribution of the canine dopamine receptor D4 gene exon III and I in 23 breeds.** *J Vet Med Sci* 2004, **66**(7):815-820.
76. Lichter JB, Barr CL, Kennedy JL, Van Tol HH, Kidd KK, Livak KJ: **A hypervariable segment in the human dopamine receptor D4 (DRD4) gene.** *Hum Mol Genet* 1993, **2**(6):767-773.

77. Okuyama Y, Ishiguro H, Toru M, Arinami T: **A genetic polymorphism in the promoter region of DRD4 associated with expression and schizophrenia.** *Biochem Biophys Res Commun* 1999, **258**(2):292-295.
78. Thomson CJ, Hanna CW, Carlson SR, Rupert JL: **The -521 C/T variant in the dopamine-4-receptor gene (DRD4) is associated with skiing and snowboarding behavior.** *Scand J Med Sci Sports* 2013, **23**(2):e108-113.
79. Wang F, Lu L, Wang SB, Zhang L, Ng CH, Ungvari GS, Cao XL, Lu JP, Hou CL, Jia FJ *et al*: **The prevalence of autism spectrum disorders in China: a comprehensive meta-analysis.** *Int J Biol Sci* 2018, **14**(7):717-725.
80. Xu FL, Wu X, Zhang JJ, Wang BJ, Yao J: **A meta-analysis of data associating DRD4 gene polymorphisms with schizophrenia.** *Neuropsychiatr Dis Treat* 2018, **14**:153-164.
81. D'Souza UM, Russ C, Tahir E, Mill J, McGuffin P, Asherson PJ, Craig IW: **Functional effects of a tandem duplication polymorphism in the 5'flanking region of the DRD4 gene.** *Biol Psychiatry* 2004, **56**(9):691-697.
82. Li KJ, Solomon HV, DeLisi LE: **Clozapine pharmacogenomics: a review of efficacy, pharmacokinetics, and agranulocytosis.** *Current opinion in psychiatry* 2018, **31**(5):403-408.
83. Mill J, Fisher N, Curran S, Richards S, Taylor E, Asherson P: **Polymorphisms in the dopamine D4 receptor gene and attention-deficit hyperactivity disorder.** *Neuroreport* 2003, **14**(11):1463-1466.
84. Takeda K, Inoue H, Tanizawa Y, Matsuzaki Y, Oba J, Watanabe Y, Shinoda K, Oka Y: **WFS1 (Wolfram syndrome 1) gene product: predominant subcellular localization to endoplasmic reticulum in cultured cells and neuronal expression in rat brain.** *Hum Mol Genet* 2001, **10**(5):477-484.
85. Fonseca SG, Ishigaki S, Osowski CM, Lu S, Lipson KL, Ghosh R, Hayashi E, Ishihara H, Oka Y, Permutt MA *et al*: **Wolfram syndrome 1 gene negatively regulates ER stress signaling in rodent and human cells.** *J Clin Invest* 2010, **120**(3):744-755.
86. Fonseca SG, Fukuma M, Lipson KL, Nguyen LX, Allen JR, Oka Y, Urano F: **WFS1 is a novel component of the unfolded protein response and maintains homeostasis of the endoplasmic reticulum in pancreatic beta-cells.** *J Biol Chem* 2005, **280**(47):39609-39615.
87. Ishihara H, Takeda S, Tamura A, Takahashi R, Yamaguchi S, Takei D, Yamada T, Inoue H, Soga H, Katagiri H *et al*: **Disruption of the WFS1 gene in mice causes progressive beta-cell loss and impaired stimulus-secretion coupling in insulin secretion.** *Hum Mol Genet* 2004, **13**(11):1159-1170.
88. Hatanaka M, Tanabe K, Yanai A, Ohta Y, Kondo M, Akiyama M, Shinoda K, Oka Y, Tanizawa Y: **Wolfram syndrome 1 gene (WFS1) product localizes to secretory granules and determines granule acidification in pancreatic beta-cells.** *Hum Mol Genet* 2011, **20**(7):1274-1284.
89. Osman AA, Saito M, Makepeace C, Permutt MA, Schlesinger P, Mueckler M: **Wolframin expression induces novel ion channel activity in endoplasmic reticulum membranes and increases intracellular calcium.** *J Biol Chem* 2003, **278**(52):52755-52762.

90. Rando TA, Horton JC, Layzer RB: **Wolfram syndrome: evidence of a diffuse neurodegenerative disease by magnetic resonance imaging.** *Neurology* 1992, **42**(6):1220-1224.
91. Chapla A, Johnson J, Korula S, Mohan N, Ahmed A, Varghese D, Rangasamy P, Ravichandran L, Jebasingh F, Agrawal KK *et al*: **WFS1 gene associated diabetes phenotypes and identification of a founder mutation in Southern India.** *J Clin Endocrinol Metab* 2022.
92. Kumar S: **Wolfram syndrome: important implications for pediatricians and pediatric endocrinologists.** *Pediatr Diabetes* 2010, **11**(1):28-37.
93. Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proenca C, Bacot F, Balkau B, Belisle A, Borch-Johnsen K *et al*: **Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia.** *Nat Genet* 2009, **41**(10):1110-1115.
94. Fawcett KA, Wheeler E, Morris AP, Ricketts SL, Hallmans G, Rolandsson O, Daly A, Wasson J, Permutt A, Hattersley AT *et al*: **Detailed investigation of the role of common and low-frequency WFS1 variants in type 2 diabetes risk.** *Diabetes* 2010, **59**(3):741-746.
95. Kong X, Zhang X, Xing X, Zhang B, Hong J, Yang W: **The Association of Type 2 Diabetes Loci Identified in Genome-Wide Association Studies with Metabolic Syndrome and Its Components in a Chinese Population with Type 2 Diabetes.** *PLoS One* 2015, **10**(11):e0143607.
96. Torkamandi S, Bastami M, Ghaedi H, Tarighi S, Shokri F, Javadi A, Mirfakhraie R, Omrani MD: **Association of CpG-SNP and 3'UTR-SNP of WFS1 with the Risk of Type 2 Diabetes Mellitus in an Iranian Population.** *Int J Mol Cell Med* 2017, **6**(4):197-203.
97. Koido K, Koks S, Nikopensus T, Maron E, Altmae S, Heinaste E, Vabrit K, Tammekivi V, Hallast P, Kurg A *et al*: **Polymorphisms in wolframin (WFS1) gene are possibly related to increased risk for mood disorders.** *Int J Neuropsychopharmacol* 2005, **8**(2):235-244.
98. Kawamoto T, Horikawa Y, Tanaka T, Kabe N, Takeda J, Mikuni M: **Genetic variations in the WFS1 gene in Japanese with type 2 diabetes and bipolar disorder.** *Mol Genet Metab* 2004, **82**(3):238-245.
99. Martorell L, Zaera MG, Valero J, Serrano D, Figuera L, Joven J, Labad A, Vilella E, Nunes V: **The WFS1 (Wolfram syndrome 1) is not a major susceptibility gene for the development of psychiatric disorders.** *Psychiatr Genet* 2003, **13**(1):29-32.
100. Daraio T, Bombek LK, Gosak M, Valladolid-Acebes I, Klemen MS, Refai E, Berggren PO, Brismar K, Rupnik MS, Bark C: **SNAP-25b-deficiency increases insulin secretion and changes spatiotemporal profile of Ca(2+)oscillations in beta cell networks.** *Sci Rep* 2017, **7**(1):7744.
101. Sollner T, Whiteheart SW, Brunner M, Erdjument-Bromage H, Geromanos S, Tempst P, Rothman JE: **SNAP receptors implicated in vesicle targeting and fusion.** *Nature* 1993, **362**(6418):318-324.

102. Ramakrishnan NA, Drescher MJ, Drescher DG: **The SNARE complex in neuronal and sensory cells.** *Mol Cell Neurosci* 2012, **50**(1):58-69.
103. Hess EJ, Jinnah HA, Kozak CA, Wilson MC: **Spontaneous locomotor hyperactivity in a mouse mutant with a deletion including the Snap gene on chromosome 2.** *J Neurosci* 1992, **12**(7):2865-2874.
104. Bark IC, Hahn KM, Ryabinin AE, Wilson MC: **Differential expression of SNAP-25 protein isoforms during divergent vesicle fusion events of neural development.** *Proc Natl Acad Sci U S A* 1995, **92**(5):1510-1514.
105. Johansson JU, Ericsson J, Janson J, Beraki S, Stanic D, Mandic SA, Wikstrom MA, Hokfelt T, Ogren SO, Rozell B *et al*: **An ancient duplication of exon 5 in the Snap25 gene is required for complex neuronal development/function.** *PLoS Genet* 2008, **4**(11):e1000278.
106. Yamamori S, Itakura M, Sugaya D, Katsumata O, Sakagami H, Takahashi M: **Differential expression of SNAP-25 family proteins in the mouse brain.** *J Comp Neurol* 2011, **519**(5):916-932.
107. Kim JW, Biederman J, Arbeitman L, Fagerness J, Doyle AE, Petty C, Perlis RH, Purcell S, Smoller JW, Faraone SV *et al*: **Investigation of variation in SNAP-25 and ADHD and relationship to co-morbid major depressive disorder.** *Am J Med Genet B Neuropsychiatr Genet* 2007, **144B**(6):781-790.
108. Gizer IR, Ficks C, Waldman ID: **Candidate gene studies of ADHD: a meta-analytic review.** *Hum Genet* 2009, **126**(1):51-90.
109. Scarr E, Gray L, Keriakous D, Robinson PJ, Dean B: **Increased levels of SNAP-25 and synaptophysin in the dorsolateral prefrontal cortex in bipolar I disorder.** *Bipolar Disord* 2006, **8**(2):133-143.
110. Carroll LS, Kendall K, O'Donovan MC, Owen MJ, Williams NM: **Evidence that putative ADHD low risk alleles at SNAP25 may increase the risk of schizophrenia.** *Am J Med Genet B Neuropsychiatr Genet* 2009, **150B**(7):893-899.
111. Dawkins R, Leelayuwat C, Gaudieri S, Tay G, Hui J, Cattley S, Martinez P, Kulski J: **Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease.** *Immunol Rev* 1999, **167**:275-304.
112. Yu CY, Chung EK, Yang Y, Blanchong CA, Jacobsen N, Saxena K, Yang Z, Miller W, Varga L, Fust G: **Dancing with complement C4 and the RP-C4-CYP21-TNX (RCCX) modules of the major histocompatibility complex.** *Prog Nucleic Acid Res Mol Biol* 2003, **75**:217-292.
113. Milner CM, Campbell RD: **Genetic organization of the human MHC class III region.** *Front Biosci* 2001, **6**:D914-926.
114. Volanakis JE: **Transcriptional regulation of complement genes.** *Annu Rev Immunol* 1995, **13**:277-305.
115. Wu Z, Zhang S, Li P, Zhang F, Li Y: **Association between complement 4 copy number variation and systemic lupus erythematosus: a meta-analysis.** *Clin Exp Med* 2020, **20**(4):627-634.

116. Ansari AA, Tipu HN, Ahmad D, Farhan M: **Impact of Homozygous C4A Deficiency on Clinical Presentation of Systemic Lupus Erythematosus.** *J Coll Physicians Surg Pak* 2020, **30**(8):790-795.
117. Lefvert AK, Hamsten A, Holm G: **Association between circulating immune complexes, complement C4 null alleles, and myocardial infarction before age 45 years.** *Arterioscler Thromb Vasc Biol* 1995, **15**(5):665-668.
118. de Messias IJ, Santamaria J, Brenden M, Reis A, Mauff G: **Association of C4B deficiency (C4B*Q0) with erythema nodosum in leprosy.** *Clin Exp Immunol* 1993, **92**(2):284-287.
119. Welch TR, Brickman C, Bishof N, Maringhini S, Rutkowski M, Frenzke M, Kantor N: **The phenotype of SLE associated with complete deficiency of complement isotype C4A.** *J Clin Immunol* 1998, **18**(1):48-51.
120. Wang H, Liu M: **Complement C4, Infections, and Autoimmune Diseases.** *Front Immunol* 2021, **12**:694928.
121. Liu YH, Wan L, Chang CT, Liao WL, Chen WC, Tsai Y, Tsai CH, Tsai FJ: **Association between copy number variation of complement component C4 and Graves' disease.** *J Biomed Sci* 2011, **18**:71.
122. Grassmann F, Cantsilieris S, Schulz-Kuhnt AS, White SJ, Richardson AJ, Hewitt AW, Vote BJ, Schmied D, Guymer RH, Weber BH *et al*: **Multiallelic copy number variation in the complement component 4A (C4A) gene is associated with late-stage age-related macular degeneration (AMD).** *J Neuroinflammation* 2016, **13**(1):81.
123. Lee H, Martinez-Agosto JA, Rexach J, Fogel BL: **Next generation sequencing in clinical diagnosis.** *Lancet Neurol* 2019, **18**(5):426.
124. Jorgenson JW, Lukacs KD: **Free-zone electrophoresis in glass capillaries.** *Clin Chem* 1981, **27**(9):1551-1553.
125. Kok WT: **Capillary electrophoresis: Instrumentation and operation.** *Chromatographia* 2000, **51**:S5-S89.
126. Bruin GJ, Huisden R, Kraak JC, Poppe H: **Performance of carbohydrate-modified fused-silica capillaries for the separation of proteins by zone electrophoresis.** *J Chromatogr* 1989, **480**:339-349.
127. Bruin GJ: **Recent developments in electrokinetically driven analysis on microfabricated devices.** *Electrophoresis* 2000, **21**(18):3931-3951.
128. Xie Y, Dai L, Yang Y: **Microfluidic technology and its application in the point-of-care testing field.** *Biosens Bioelectron X* 2022, **10**:100109.
129. Wang H, Diao B, Cui Z, Yan M, Kan B: **Genotyping of Salmonella Typhi using 8-loci multi locus VNTR analysis.** *Gut Pathog* 2016, **8**:14.
130. Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF: **Mutation detection using Surveyor nuclease.** *Biotechniques* 2004, **36**(4):702-707.
131. Little S: **Amplification-refractory mutation system (ARMS) analysis of point mutations.** *Curr Protoc Hum Genet* 2001, **Chapter 9**:Unit 9 8.

132. Bunyan DJ, Eccles DM, Sillibourne J, Wilkins E, Thomas NS, Shea-Simonds J, Duncan PJ, Curtis CE, Robinson DO, Harvey JF *et al*: **Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification.** *Br J Cancer* 2004, **91**(6):1155-1159.
133. Konecny M, Zavodna K, Vranova V, Vizvaryova M, Weismanova E, Mlkva I, Kuglik P, Kausitz J, Bartosova Z: **Identification of rare complete BRCA1 gene deletion using a combination of SNP haplotype analysis, MLPA and array-CGH techniques.** *Breast Cancer Res Treat* 2008, **109**(3):581-583.
134. Dai X, Shi F, Cheung CKY, Li J, Lin S: **Abnormal Y chromosome detection in infertile males using multiplex ligation-dependent probe amplification.** *Andrologia* 2022, **54**(2):e14316.
135. Shang S, Mei Y, Wang T, Zheng X, Chen K, Xiong S, Dong Y, Chang Y, Wu X, Kong X *et al*: **Diagnosis and genotype-phenotype correlation in patients with PKD1/TSC2 contiguous gene deletion syndrome.** *Clin Nephrol* 2022.
136. Kim KB, Park S, Ha JS, Ryoo N, Kim DH: **Three Cases of False-positive Multiplex Ligation-dependent Probe Amplification of BRCA1.** *Ann Lab Med* 2022, **42**(4):497-499.
137. Verlow JAM, Clemens E, de Vries JH, Zolk O, Verkerk A, Am Zehnhoff-Dinnesen A, Medina-Gomez C, Lanvers-Kaminsky C, Rivadeneira F, Langer T *et al*: **A comparison of genotyping arrays.** *Eur J Hum Genet* 2021, **29**(11):1611-1624.
138. Urtiaga GO, Domingues WB, Komninou ER, Martins AWS, Blodorn EB, Dellagostin EN, Woloski RDS, Pinto LS, Brum CB, Tovo-Rodrigues L *et al*: **DNA microarray for forensic intelligence purposes: High-density SNP profiles obtained directly from casework-like samples with and without a DNA purification step.** *Forensic Sci Int* 2022, **332**:111181.
139. Dalma-Weiszhausz DD, Warrington J, Tanimoto EY, Miyada CG: **The affymetrix GeneChip platform: an overview.** *Methods Enzymol* 2006, **410**:3-28.
140. Esteves LM, Bulhoes SM, Brilhante MJ, Mota-Vieira L: **Three multiplex snapshot assays for SNP genotyping in candidate innate immune genes.** *BMC Res Notes* 2013, **6**:54.
141. De Luca P, Iaconis D, Biffali E, Enza C, de Magistris L, Riegler G, Pappalardo D, Amato MR, Iardino P, Montanino C *et al*: **Development of a novel SNP assay to detect lactase persistence associated genetic variants.** *Mol Biol Rep* 2021, **48**(11):7087-7093.
142. Ortiz M, Jauset-Rubio M, Skouridou V, Machado D, Viveiros M, Clark TG, Simonova A, Kodr D, Hocek M, O'Sullivan CK: **Electrochemical Detection of Single-Nucleotide Polymorphism Associated with Rifampicin Resistance in Mycobacterium tuberculosis Using Solid-Phase Primer Elongation with Ferrocene-Linked Redox-Labeled Nucleotides.** *ACS Sens* 2021, **6**(12):4398-4407.
143. Gudnason H, Dufva M, Bang DD, Wolff A: **Comparison of multiple DNA dyes for real-time PCR: effects of dye concentration and sequence composition on DNA amplification and melting temperature.** *Nucleic Acids Res* 2007, **35**(19):e127.

144. De la Vega FM, Lazaruk KD, Rhodes MD, Wenz MH: **Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP Genotyping Assays and the SNPlex Genotyping System.** *Mutat Res* 2005, **573**(1-2):111-135.
145. Abravaya K, Huff J, Marshall R, Merchant B, Mullen C, Schneider G, Robinson J: **Molecular beacons as diagnostic tools: technology and applications.** *Clin Chem Lab Med* 2003, **41**(4):468-474.
146. Tapp I, Malmberg L, Rennel E, Wik M, Syvanen AC: **Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan assay and Molecular Beacon probes.** *Biotechniques* 2000, **28**(4):732-738.
147. Pomeroy R, Duncan G, Sunar-Reeder B, Ortenberg E, Ketchum M, Wasiluk H, Reeder D: **A low-cost, high-throughput, automated single nucleotide polymorphism assay for forensic human DNA applications.** *Anal Biochem* 2009, **395**(1):61-67.
148. Bautista-Medina MA, Gallardo-Blanco HL, Martinez-Garza LE, Cerda-Flores RM, Lavallo-Gonzalez FJ, Villarreal-Perez JZ: **Association study in Mexican patients with thyrotoxic hypokalemic periodic paralysis.** *Biomed Rep* 2020, **13**(4):24.
149. Patel SN, Wu Y, Bao Y, Mancebo R, Au-Young J, Grigorenko E: **TaqMan(R) OpenArray(R) high-throughput transcriptional analysis of human embryonic and induced pluripotent stem cells.** *Methods Mol Biol* 2013, **997**:191-201.
150. Tucker EJ, Huynh BL: **Genotyping by high-resolution melting analysis.** *Methods Mol Biol* 2014, **1145**:59-66.
151. Bruzzone CM, Steer CJ: **High-resolution melting analysis of single nucleotide polymorphisms.** *Methods Mol Biol* 2015, **1310**:5-27.
152. Ito S, Hirota T, Yanai M, Muto M, Watanabe E, Taya Y, Ieiri I: **Effects of Genetic Polymorphisms of Cathepsin A on Metabolism of Tenofovir Alafenamide.** *Genes (Basel)* 2021, **12**(12).
153. Knapp LA: **Single nucleotide polymorphism screening with denaturing gradient gel electrophoresis.** *Methods Mol Biol* 2009, **578**:137-151.
154. Viglasky V: **Polyacrylamide temperature gradient gel electrophoresis.** *Methods Mol Biol* 2013, **1054**:159-171.
155. Dong Y, Zhu H: **Single-strand conformational polymorphism analysis: basic principles and routine practice.** *Methods Mol Med* 2005, **108**:149-157.
156. Haraksingh RR, Abyzov A, Urban AE: **Comprehensive performance comparison of high-resolution array platforms for genome-wide Copy Number Variation (CNV) analysis in humans.** *BMC Genomics* 2017, **18**(1):321.
157. Ceulemans S, van der Ven K, Del-Favero J: **Targeted screening and validation of copy number variations.** *Methods Mol Biol* 2012, **838**:311-328.
158. Sorrentino E, Daja M, Cristofoli F, Paolacci S, Bertelli M, Marceddu G: **CNV analysis in a diagnostic setting using target panel.** *Eur Rev Med Pharmacol Sci* 2021, **25**(1 Suppl):7-13.
159. Anderson MW, Schrijver I: **Next generation DNA sequencing and the future of genomic medicine.** *Genes (Basel)* 2010, **1**(1):38-69.

160. McKernan KJ, Peckham HE, Costa GL, McLaughlin SF, Fu Y, Tsung EF, Clouser CR, Duncan C, Ichikawa JK, Lee CC *et al*: **Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding.** *Genome Res* 2009, **19**(9):1527-1541.
161. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z *et al*: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437**(7057):376-380.
162. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M *et al*: **An integrated semiconductor device enabling non-optical genome sequencing.** *Nature* 2011, **475**(7356):348-352.
163. Nair J, Shetty S, Kasi CI, Thondehalmath N, Ganesh D, Bhat VR, Mannadia S, Ranganath A, Nayak R, Gunasheela D *et al*: **Preimplantation genetic testing for aneuploidy (PGT-A)-a single-center experience.** *J Assist Reprod Genet* 2022.
164. Hemissi I, Boussetta S, Dallali H, Hellal F, Durand G, Voegelé C, Ayed H, Zaghbib S, Naimi Z, Ayadi M *et al*: **Development of a custom next-generation sequencing panel for the determination of bladder cancer risk in a Tunisian cohort.** *Mol Biol Rep* 2022, **49**(2):1233-1258.
165. Suppan C, Graf R, Jahn S, Zhou Q, Klocker EV, Bartsch R, Terbuch A, Kashofer K, Regitnig P, Lindenmann J *et al*: **Sensitive and robust liquid biopsy-based detection of PIK3CA mutations in hormone-receptor-positive metastatic breast cancer patients.** *Br J Cancer* 2022, **126**(3):456-463.
166. Huang CC, Liu CY, Huang CJ, Hsu YC, Lien HH, Wong JU, Tai FC, Ku WH, Hung CF, Lin JT *et al*: **Deciphering Genetic Alterations of Taiwanese Patients with Pancreatic Adenocarcinoma through Targeted Sequencing.** *Int J Mol Sci* 2022, **23**(3).
167. Guo J, Xu N, Li Z, Zhang S, Wu J, Kim DH, Sano Marma M, Meng Q, Cao H, Li X *et al*: **Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides.** *Proc Natl Acad Sci U S A* 2008, **105**(27):9145-9150.
168. Ju J, Kim DH, Bi L, Meng Q, Bai X, Li Z, Li X, Marma MS, Shi S, Wu J *et al*: **Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators.** *Proc Natl Acad Sci U S A* 2006, **103**(52):19635-19640.
169. Nemashkalo A, Phipps ME, Hennelly SP, Goodwin PM: **Real-time, single-molecule observation of biomolecular interactions inside nanophotonic zero mode waveguides.** *Nanotechnology* 2022, **33**(16).
170. Chaisson MJ, Huddleston J, Dennis MY, Sudmant PH, Malig M, Hormozdiari F, Antonacci F, Surti U, Sandstrom R, Boitano M *et al*: **Resolving the complexity of the human genome using single-molecule sequencing.** *Nature* 2015, **517**(7536):608-611.
171. McCoy RC, Taylor RW, Blauwkamp TA, Kelley JL, Kertesz M, Pushkarev D, Petrov DA, Fiston-Lavier AS: **Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements.** *PLoS One* 2014, **9**(9):e106689.
172. David S: **A current guide to candidate gene association studies.** *Trends Genet* 2021, **37**(12):1056-1059.

173. Wang K, Zhang H, Kugathasan S, Annese V, Bradfield JP, Russell RK, Sleiman PM, Imielinski M, Glessner J, Hou C *et al*: **Diverse genome-wide association studies associate the IL12/IL23 pathway with Crohn Disease.** *Am J Hum Genet* 2009, **84**(3):399-405.
174. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, de Leeuw CA, Benjamins JS, Munoz-Manchado AB, Nagel M *et al*: **Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways.** *Nat Genet* 2019, **51**(3):394-403.
175. Watanabe K, Stringer S, Frei O, Umicevic Mirkov M, de Leeuw C, Polderman TJC, van der Sluis S, Andreassen OA, Neale BM, Posthuma D: **A global overview of pleiotropy and genetic architecture in complex traits.** *Nat Genet* 2019, **51**(9):1339-1348.
176. Purcell S, Cherny SS, Sham PC: **Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits.** *Bioinformatics* 2003, **19**(1):149-150.
177. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE: **Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population.** *Am J Epidemiol* 2017, **186**(9):1026-1034.
178. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, Shang L, Boisson B, Casanova JL, Abel L: **Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants.** *Proc Natl Acad Sci U S A* 2015, **112**(17):5473-5478.
179. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al*: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007, **81**(3):559-575.
180. Lawson DJ, Davies NM, Haworth S, Ashraf B, Howe L, Crawford A, Hemani G, Davey Smith G, Timpson NJ: **Correction to: Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity?** *Hum Genet* 2020, **139**(1):43-44.
181. Hamer D, Sirota L: **Beware the chopsticks gene.** *Mol Psychiatry* 2000, **5**(1):11-13.
182. Yamamoto A, Shibuya T: **Efficient Differentially Private Methods for a Transmission Disequilibrium Test in Genome Wide Association Studies.** *Pac Symp Biocomput* 2022, **27**:85-96.
183. Neumann C, Taub MA, Younkin SG, Beaty TH, Ruczinski I, Schwender H: **Analytic power and sample size calculation for the genotypic transmission/disequilibrium test in case-parent trio studies.** *Biom J* 2014, **56**(6):1076-1092.
184. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing.** *J R Stat Soc B* 1995, **57**(1):289-300.
185. Watanabe K, Taskesen E, van Bochoven A, Posthuma D: **Functional mapping and annotation of genetic associations with FUMA.** *Nature communications* 2017, **8**(1):1826.

186. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ: **LocusZoom: regional visualization of genome-wide association scan results.** *Bioinformatics* 2010, **26**(18):2336-2337.
187. Hemminki K, Forsti A, Bermejo JL: **The 'common disease-common variant' hypothesis and familial risks.** *PLoS One* 2008, **3**(6):e2504.
188. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A *et al*: **Finding the missing heritability of complex diseases.** *Nature* 2009, **461**(7265):747-753.
189. Wellek S, Ziegler A: **Cochran-Armitage Test versus Logistic Regression in the Analysis of Genetic Association Studies.** *Hum Hered* 2012, **73**(1):14-17.
190. Das PM, Ramachandran K, vanWert J, Singal R: **Chromatin immunoprecipitation assay.** *Biotechniques* 2004, **37**(6):961-969.
191. Kim TH, Ren B: **Genome-wide analysis of protein-DNA interactions.** *Annual review of genomics and human genetics* 2006, **7**:81-102.
192. Daniels DL, Urh M: **Isolation of intracellular protein--DNA complexes using HaloCHIP, an antibody-free alternative to chromatin immunoprecipitation.** *Methods Mol Biol* 2013, **977**:111-124.
193. Greil F, Moorman C, van Steensel B: **DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase.** *Methods Enzymol* 2006, **410**:342-359.
194. Hellman LM, Fried MG: **Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions.** *Nat Protoc* 2007, **2**(8):1849-1861.
195. Galas DJ, Schmitz A: **DNase footprinting: a simple method for the detection of protein-DNA binding specificity.** *Nucleic Acids Res* 1978, **5**(9):3157-3170.
196. Cardew AS, Fox KR: **DNase I footprinting.** *Methods Mol Biol* 2010, **613**:153-172.
197. Mendauletova A, Latham JA: **Biosynthesis of the redox cofactor mycofactocin is controlled by the transcriptional regulator MftR and induced by long-chain acyl-CoA species.** *J Biol Chem* 2022, **298**(1):101474.
198. Bailly C, Kluza J, Martin C, Ellis T, Waring MJ: **DNase I footprinting of small molecule binding sites on DNA.** *Methods Mol Biol* 2005, **288**:319-342.
199. Pfeifer GP, Riggs AD: **Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR.** *Genes Dev* 1991, **5**(6):1102-1113.
200. Liu GL, Yin Y, Kunchakarra S, Mukherjee B, Gerion D, Jett SD, Bear DG, Gray JW, Alivisatos AP, Lee LP *et al*: **A nanoplasmonic molecular ruler for measuring nuclease activity and DNA footprinting.** *Nat Nanotechnol* 2006, **1**(1):47-52.
201. Wang J, Li T, Guo X, Lu Z: **Exonuclease III protection assay with FRET probe for detecting DNA-binding proteins.** *Nucleic Acids Res* 2005, **33**(2):e23.
202. He HJ, Pires R, Zhu TN, Zhou A, Gaigalas AK, Zou S, Wang L: **Fluorescence resonance energy transfer-based method for detection of DNA binding activities of nuclear factor kappaB.** *Biotechniques* 2007, **43**(1):93-98.

203. Blouin S, Craggs TD, Lafontaine DA, Penedo JC: **Functional studies of DNA-protein interactions using FRET techniques.** *Methods Mol Biol* 2009, **543**:475-502.
204. Glickman JF, Wu X, Mercuri R, Illy C, Bowen BR, He Y, Sills M: **A comparison of ALPHAScreen, TR-FRET, and TRF as assay methods for FXR nuclear receptors.** *J Biomol Screen* 2002, **7**(1):3-10.
205. Reece-Hoyes JS, Marian Walhout AJ: **Yeast one-hybrid assays: a historical and technical perspective.** *Methods* 2012, **57**(4):441-447.
206. Vidal M, Legrain P: **Yeast forward and reverse 'n'-hybrid systems.** *Nucleic Acids Res* 1999, **27**(4):919-929.
207. Gustafsdottir SM, Schlingemann J, Rada-Iglesias A, Schallmeiner E, Kamali-Moghaddam M, Wadelius C, Landegren U: **In vitro analysis of DNA-protein interactions by proximity ligation.** *Proc Natl Acad Sci U S A* 2007, **104**(9):3067-3072.
208. Alam J, Cook JL: **Reporter genes: application to the study of mammalian gene transcription.** *Anal Biochem* 1990, **188**(2):245-254.
209. Ciccacci C, Di Fusco D, Cacciotti L, Morganti R, D'Amato C, Greco C, Rufini S, Novelli G, Sangiuolo F, Spallone V *et al*: **MicroRNA genetic variations: association with type 2 diabetes.** *Acta Diabetol* 2013, **50**(6):867-872.
210. Locke JM, Lango Allen H, Harries LW: **A rare SNP in pre-miR-34a is associated with increased levels of miR-34a in pancreatic beta cells.** *Acta Diabetol* 2014, **51**(2):325-329.
211. Yang L, Yang S, Ren C, Liu S, Zhang X, Sui A: **Deciphering the roles of miR-16-5p in Malignant Solid Tumorsmalignant solid tumors.** *Biomed Pharmacother* 2022, **148**:112703.
212. Maralani M, Shanehbandi D, Asadi M, Hashemzadeh S, Hajiasgharzadeh K, Mashhadi Abdolahi H, Baradaran B, Peeters M: **Expression profiles of miR-196, miR-132, miR-146a, and miR-134 in human colorectal cancer tissues in accordance with their clinical significance : Comparison regarding KRAS mutation.** *Wien Klin Wochenschr* 2021, **133**(21-22):1162-1170.
213. Rupaimoole R, Slack FJ: **MicroRNA therapeutics: towards a new era for the management of cancer and other diseases.** *Nat Rev Drug Discov* 2017, **16**(3):203-222.
214. Sticht C, De La Torre C, Parveen A, Gretz N: **miRWalk: An online resource for prediction of microRNA binding sites.** *PLoS One* 2018, **13**(10):e0206239.
215. Cho S, Jang I, Jun Y, Yoon S, Ko M, Kwon Y, Choi I, Chang H, Ryu D, Lee B *et al*: **MiRGator v3.0: a microRNA portal for deep sequencing, expression profiling and mRNA targeting.** *Nucleic Acids Res* 2013, **41**(Database issue):D252-257.
216. Wolter JM, Kotagama K, Pierre-Bez AC, Firago M, Mangone M: **3'LIFE: a functional assay to detect miRNA targets in high-throughput.** *Nucleic Acids Res* 2014, **42**(17):e132.
217. Guo H, Ingolia NT, Weissman JS, Bartel DP: **Mammalian microRNAs predominantly act to decrease target mRNA levels.** *Nature* 2010, **466**(7308):835-840.

218. Lim LP, Lau NC, Garrett-Engle P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM: **Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs.** *Nature* 2005, **433**(7027):769-773.
219. Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP: **The impact of microRNAs on protein output.** *Nature* 2008, **455**(7209):64-71.
220. Karginov FV, Conaco C, Xuan Z, Schmidt BH, Parker JS, Mandel G, Hannon GJ: **A biochemical approach to identifying microRNA targets.** *Proc Natl Acad Sci U S A* 2007, **104**(49):19291-19296.
221. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M *et al*: **Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.** *Cell* 2010, **141**(1):129-141.
222. Helwak A, Tollervey D: **Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH).** *Nat Protoc* 2014, **9**(3):711-728.
223. Orom UA, Nielsen FC, Lund AH: **MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation.** *Mol Cell* 2008, **30**(4):460-471.
224. Lal A, Thomas MP, Altschuler G, Navarro F, O'Day E, Li XL, Concepcion C, Han YC, Thiery J, Rajani DK *et al*: **Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling.** *PLoS Genet* 2011, **7**(11):e1002363.
225. Baigude H, Ahsanullah, Li Z, Zhou Y, Rana TM: **miR-TRAP: a benchtop chemical biology strategy to identify microRNA targets.** *Angew Chem Int Ed Engl* 2012, **51**(24):5880-5883.
226. Guo YE, Steitz JA: **3'-Biotin-tagged microRNA-27 does not associate with Argonaute proteins in cells.** *RNA* 2014, **20**(7):985-988.
227. Li J, Huang L, Xiao X, Chen Y, Wang X, Zhou Z, Zhang C, Zhang Y: **Photoclickable MicroRNA for the Intracellular Target Identification of MicroRNAs.** *J Am Chem Soc* 2016, **138**(49):15943-15949.
228. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps.** *Bioinformatics* 2005, **21**(2):263-265.
229. Stephens M, Smith NJ, Donnelly P: **A new statistical method for haplotype reconstruction from population data.** *Am J Hum Genet* 2001, **68**(4):978-989.
230. Mitsuyasu H, Hirata N, Sakai Y, Shibata H, Takeda Y, Ninomiya H, Kawasaki H, Tashiro N, Fukumaki Y: **Association analysis of polymorphisms in the upstream region of the human dopamine D4 receptor gene (DRD4) with schizophrenia and personality traits.** *J Hum Genet* 2001, **46**(1):26-31.
231. Vila C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J, Wayne RK: **Multiple and ancient origins of the domestic dog.** *Science* 1997, **276**(5319):1687-1689.
232. Jaghutriz BA, Heni M, Lutz SZ, Fritsche L, Machicao F, Staiger H, Peter A, Haring HU, Fritsche A, Wagner R: **Gene x Gene Interactions Highlight the Role of Incretin Resistance for Insulin Secretion.** *Front Endocrinol (Lausanne)* 2019, **10**:72.

233. Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, Chang CT, Wang TY, Chen RH, Shiu CF, Liu YM *et al*: **A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese.** *PLoS Genet* 2010, **6**(2):e1000847.
234. Zalsman G, Mann MJ, Huang YY, Oquendo MA, Brent DA, Burke AK, Ellis SP, Mann JJ: **Wolframin gene H611R polymorphism: no direct association with suicidal behavior but possible link to mood disorders.** *Prog Neuropsychopharmacol Biol Psychiatry* 2009, **33**(4):707-710.
235. Ryu J, Lee C: **Differential promoter activity by nucleotide substitution at a type 2 diabetes genome-wide association study signal upstream of the wolframin gene.** *J Diabetes* 2016, **8**(2):253-259.
236. Man XY, Luo HR, Li XP, Yao YG, Mao CZ, Zhang YP: **Polymerase chain reaction based C4AQ0 and C4BQ0 genotyping: association with systemic lupus erythematosus in southwest Han Chinese.** *Ann Rheum Dis* 2003, **62**(1):71-73.
237. Yu CY, Campbell RD: **Definitive RFLPs to distinguish between the human complement C4A/C4B isotypes and the major Rodgers/Chido determinants: application to the study of C4 null alleles.** *Immunogenetics* 1987, **25**(6):383-390.
238. Blanchong CA, Zhou B, Rupert KL, Chung EK, Jones KN, Sotos JF, Zipf WB, Rennebohm RM, Yung Yu C: **Deficiencies of human complement component C4A and C4B and heterozygosity in length variants of RP-C4-CYP21-TNX (RCCX) modules in caucasians. The load of RCCX genetic diversity on major histocompatibility complex-associated disease.** *J Exp Med* 2000, **191**(12):2183-2196.
239. Yang Y, Chung EK, Zhou B, Blanchong CA, Yu CY, Fust G, Kovacs M, Vatay A, Szalai C, Karadi I *et al*: **Diversity in intrinsic strengths of the human complement system: serum C4 protein concentrations correlate with C4 gene size and polygenic variations, hemolytic activities, and body mass index.** *J Immunol* 2003, **171**(5):2734-2745.
240. Emamian ES, Hall D, Birnbaum MJ, Karayiorgou M, Gogos JA: **Convergent evidence for impaired AKT1-GSK3beta signaling in schizophrenia.** *Nat Genet* 2004, **36**(2):131-137.
241. Van Wauwe J, Haefner B: **Glycogen synthase kinase-3 as drug target: from wallflower to center of attention.** *Drug News Perspect* 2003, **16**(9):557-565.
242. Gould TD, Zarate CA, Manji HK: **Glycogen synthase kinase-3: a target for novel bipolar disorder treatments.** *J Clin Psychiatry* 2004, **65**(1):10-21.
243. Long JC, Knowler WC, Hanson RL, Robin RW, Urbanek M, Moore E, Bennett PH, Goldman D: **Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from an autosome-wide scan in an American Indian population.** *Am J Med Genet* 1998, **81**(3):216-221.
244. Heiger DN, Cohen AS, Karger BL: **Separation of DNA restriction fragments by high performance capillary electrophoresis with low and zero crosslinked polyacrylamide using continuous and pulsed electric fields.** *J Chromatogr* 1990, **516**(1):33-48.
245. Delaneau O, Zagury JF, Robinson MR, Marchini JL, Dermitzakis ET: **Accurate, scalable and integrative haplotype estimation.** *Nature communications* 2019, **10**(1):5436.

-
246. McGuffin P, Farmer A, Gottesman, II: **Is there really a split in schizophrenia? The genetic evidence.** *Br J Psychiatry* 1987, **150**:581-592.
247. Reitz C, Mayeux R: **Endophenotypes in normal brain morphology and Alzheimer's disease: a review.** *Neuroscience* 2009, **164**(1):174-190.
248. Lenzenweger MF: **Endophenotype, intermediate phenotype, biomarker: definitions, concept comparisons, clarifications.** *Depress Anxiety* 2013, **30**(3):185-189.
249. Cherny SS, Williams FMK, Livshits G: **Genetic and environmental correlational structure among metabolic syndrome endophenotypes.** *Ann Hum Genet* 2022.
250. Zalloua P, Kadar H, Hariri E, Abi Farraj L, Brial F, Hedjazi L, Le Lay A, Colleu A, Dubus J, Touboul D *et al*: **Untargeted Mass Spectrometry Lipidomics identifies correlation between serum sphingomyelins and plasma cholesterol.** *Lipids Health Dis* 2019, **18**(1):38.
251. Murthy VL, Naylor M, Carnethon M, Reis JP, Lloyd-Jones D, Allen NB, Kitchen R, Piaggi P, Steffen LM, Vasan RS *et al*: **Circulating metabolite profile in young adulthood identifies long-term diabetes susceptibility: the Coronary Artery Risk Development in Young Adults (CARDIA) study.** *Diabetologia* 2022, **65**(4):657-674.
252. Kachadourian LK, Quigley BM, Leonard KE: **Alcohol expectancies and evaluations of aggression in alcohol-related intimate-partner verbal and physical aggression.** *J Stud Alcohol Drugs* 2014, **75**(5):744-752.
253. Bergamini G, Massinet H, Durkin S, Steiner MA: **Longitudinal assessment of aggression and circadian rhythms in the APPswe mouse model of Alzheimer's disease.** *Physiol Behav* 2022:113787.
254. Moore FR, Doughty H, Neumann T, McClelland H, Allott C, O'Connor RC: **Impulsivity, aggression, and suicidality relationship in adults: A systematic review and meta-analysis.** *EClinicalMedicine* 2022, **45**:101307.
255. Guthman EM, Falkner AL: **Neural mechanisms of persistent aggression.** *Curr Opin Neurobiol* 2022, **73**:102526.
256. Elek Z, Kovacs Z, Keszler G, Szabo M, Csanky E, Luo J, Guttman A, Ronai Z: **High Throughput Multiplex SNP-analysis in Chronic Obstructive Pulmonary Disease and Lung Cancer.** *Curr Mol Med* 2020, **20**(3):185-193.
257. Brown GL, Goodwin FK, Ballenger JC, Goyer PF, Major LF: **Aggression in humans correlates with cerebrospinal fluid amine metabolites.** *Psychiatry Res* 1979, **1**(2):131-139.
258. Winstanley CA: **Gambling rats: insight into impulsive and addictive behavior.** *Neuropsychopharmacology* 2011, **36**(1):359.
259. Pattij T, Vanderschuren L: **The Neuropharmacology of Impulsive Behaviour, an Update.** *Curr Top Behav Neurosci* 2020, **47**:3-22.
260. Yildirim BO, Derksen JJ: **Systematic review, structural analysis, and new theoretical perspectives on the role of serotonin and associated genes in the etiology of psychopathy and sociopathy.** *Neurosci Biobehav Rev* 2013, **37**(7):1254-1296.
261. Tops M, Russo S, Boksem MA, Tucker DM: **Serotonin: modulator of a drive to withdraw.** *Brain Cogn* 2009, **71**(3):427-436.
-

262. Ten Eyck GR, Regen EM: **Chronic fluoxetine treatment promotes submissive behavior in the territorial frog, *Eleutherodactylus coqui*.** *Pharmacol Biochem Behav* 2014, **124**:86-91.
263. Passamonti L, Crockett MJ, Apergis-Schoute AM, Clark L, Rowe JB, Calder AJ, Robbins TW: **Effects of acute tryptophan depletion on prefrontal-amygdala connectivity while viewing facial signals of aggression.** *Biol Psychiatry* 2012, **71**(1):36-43.
264. Morozova A, Zorkina Y, Pavlov K, Pavlova O, Storozheva Z, Zubkov E, Zakharova N, Karpenko O, Reznik A, Chekhonin V *et al*: **Association of rs4680 COMT, rs6280 DRD3, and rs7322347 5HT2A With Clinical Features of Youth-Onset Schizophrenia.** *Front Psychiatry* 2019, **10**:830.
265. Ribases M, Ramos-Quiroga JA, Hervas A, Bosch R, Bielsa A, Gastaminza X, Artigas J, Rodriguez-Ben S, Estivill X, Casas M *et al*: **Exploration of 19 serotonergic candidate genes in adults and children with attention-deficit/hyperactivity disorder identifies association for 5HT2A, DDC and MAOB.** *Mol Psychiatry* 2009, **14**(1):71-85.
266. Blasi G, De Virgilio C, Papazacharias A, Taurisano P, Gelao B, Fazio L, Ursini G, Sinibaldi L, Andriola I, Masellis R *et al*: **Converging evidence for the association of functional genetic variation in the serotonin receptor 2a gene with prefrontal function and olanzapine treatment.** *JAMA psychiatry* 2013, **70**(9):921-930.
267. Rigoli L, Lombardo F, Di Bella C: **Wolfram syndrome and WFS1 gene.** *Clin Genet* 2011, **79**(2):103-117.
268. Swift M, Swift RG: **Psychiatric disorders and mutations at the Wolfram syndrome locus.** *Biol Psychiatry* 2000, **47**(9):787-793.
269. Wilkin TJ: **Diabetes: 1 and 2, or one and the same? Progress with the accelerator hypothesis.** *Pediatr Diabetes* 2008, **9**(3 Pt 2):23-32.
270. Congdon E, Canli T: **A neurogenetic approach to impulsivity.** *J Pers* 2008, **76**(6):1447-1484.
271. Swann AC, Dougherty DM, Pazzaglia PJ, Pham M, Steinberg JL, Moeller FG: **Increased impulsivity associated with severity of suicide attempt history in patients with bipolar disorder.** *Am J Psychiatry* 2005, **162**(9):1680-1687.
272. Nandagopal JJ, Fleck DE, Adler CM, Mills NP, Strakowski SM, DelBello MP: **Impulsivity in adolescents with bipolar disorder and/or attention-deficit/hyperactivity disorder and healthy controls as measured by the Barratt Impulsiveness Scale.** *J Child Adolesc Psychopharmacol* 2011, **21**(5):465-468.
273. Boscutti A, Pignoni A, Delvecchio G, Lazzaretti M, Mandolini GM, Girardi P, Ferro A, Sala M, Abbiati V, Cappucciati M *et al*: **The Influence of 5-HTTLPR, BDNF Rs6265 and COMT Rs4680 Polymorphisms on Impulsivity in Bipolar Disorder: The Role of Gender.** *Genes (Basel)* 2022, **13**(3).
274. Park CI, Kim HW, Hwang SS, Kang JI, Kim SJ: **Influence of dopamine-related genes on craving, impulsivity, and aggressiveness in Korean males with alcohol use disorder.** *Eur Arch Psychiatry Clin Neurosci* 2021, **271**(5):865-872.

275. Bidwell LC, Gray JC, Weafer J, Palmer AA, de Wit H, MacKillop J: **Genetic influences on ADHD symptom dimensions: Examination of a priori candidates, gene-based tests, genome-wide variation, and SNP heritability.** *Am J Med Genet B Neuropsychiatr Genet* 2017, **174**(4):458-466.
276. Yalin OO, Gokdogan Edgunlu T, Karakas Celik S, Emre U, Gunes T, Erdal Y, Eroglu Unal A: **Novel SNARE Complex Polymorphisms Associated with Multiple Sclerosis: Signs of Synaptopathy in Multiple Sclerosis.** *Balkan Med J* 2019, **36**(3):174-178.
277. Guan F, Zhang T, Han W, Zhu L, Ni T, Lin H, Liu D, Chen G, Xiao J, Li T: **Relationship of SNAP25 variants with schizophrenia and antipsychotic-induced weight change in large-scale schizophrenia patients.** *Schizophr Res* 2020, **215**:250-255.
278. Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B *et al*: **Genome sequencing highlights the dynamic early history of dogs.** *PLoS Genet* 2014, **10**(1):e1004016.
279. Topal J, Gergely G, Erdohegyi A, Csibra G, Miklosi A: **Differential sensitivity to human communication in dogs, wolves, and human infants.** *Science* 2009, **325**(5945):1269-1272.
280. Horn L, Huber L, Range F: **The importance of the secure base effect for domestic dogs - evidence from a manipulative problem-solving task.** *PLoS One* 2013, **8**(5):e65296.
281. Andics A, Gacsi M, Farago T, Kis A, Miklosi A: **Voice-sensitive regions in the dog and human brain are revealed by comparative fMRI.** *Curr Biol* 2014, **24**(5):574-578.
282. Woodard GE, Salido GM, Rosado JA: **Enhanced exocytotic-like insertion of Orail into the plasma membrane upon intracellular Ca²⁺ store depletion.** *Am J Physiol Cell Physiol* 2008, **294**(6):C1323-1331.
283. Ziebarth JD, Bhattacharya A, Chen A, Cui Y: **PolymiRTS Database 2.0: linking polymorphisms in microRNA target sites with human diseases and complex traits.** *Nucleic Acids Res* 2012, **40**(Database issue):D216-221.
284. Yan YX, Xiao HB, Zhang J, Wang S, Dong J, Wu LJ: **Pri-miR-144 rs9279 is associated with type 2 diabetes and regulation of stress response.** *J Cell Physiol* 2021, **236**(1):561-569.
285. Wang L, Xing Q, Feng T, He M, Yu W, Chen H: **SNP rs710886 A>G in long noncoding RNA PCAT1 is associated with the risk of endometriosis by modulating expression of multiple stemness-related genes via microRNA-145 signaling pathway.** *J Cell Biochem* 2020, **121**(2):1703-1715.
286. Shareef S, Ebrahimi SO, Reisi S: **Contribution of hsa-miR-146a and hsa-miR-223 gene variations in patients with multiple sclerosis reveals association of rs2910164 and rs1044165 with risk of multiple sclerosis susceptibility.** *J Investig Med* 2021, **69**(5):1015-1021.
287. Ghafouri-Fard S, Honarmand Tamizkar K, Hussen BM, Taheri M: **MicroRNA signature in liver cancer.** *Pathol Res Pract* 2021, **219**:153369.

288. Zhong ZQ, Xiang Y, Hu X, Wang YC, Zeng X, Wang XM, Xia QJ, Wang TH, Zhang X: **Synaptosomal-associated protein 25 may be an intervention target for improving sensory and locomotor functions after spinal cord contusion.** *Neural Regen Res* 2017, **12**(6):969-976.
289. Irfan M, Gopaul KR, Miry O, Hokfelt T, Stanton PK, Bark C: **SNAP-25 isoforms differentially regulate synaptic transmission and long-term synaptic plasticity at central synapses.** *Sci Rep* 2019, **9**(1):6403.
290. Daraio T, Valladolid-Acebes I, Brismar K, Bark C: **SNAP-25a and SNAP-25b differently mediate interactions with Munc18-1 and Gbetagamma subunits.** *Neurosci Lett* 2018, **674**:75-80.
291. Beck J, Turnquist C, Horikawa I, Harris C: **Targeting cellular senescence in cancer and aging: roles of p53 and its isoforms.** *Carcinogenesis* 2020, **41**(8):1017-1029.
292. Murphy AJ, Li AH, Li P, Sun H: **Therapeutic Targeting of Alternative Splicing: A New Frontier in Cancer Treatment.** *Front Oncol* 2022, **12**:868664.
293. Dlamini Z, Mokoena F, Hull R: **Abnormalities in alternative splicing in diabetes: therapeutic targets.** *J Mol Endocrinol* 2017, **59**(2):R93-R107.
294. Shinde AV, Bystroff C, Wang C, Vogelesang MG, Vincent PA, Hynes RO, Van De Water L: **Identification of the peptide sequences within the EIIIA (EDA) segment of fibronectin that mediate integrin alpha9beta1-dependent cellular activities.** *J Biol Chem* 2008, **283**(5):2858-2870.
295. Dredge BK, Polydorides AD, Darnell RB: **The splice of life: alternative splicing and neurological disease.** *Nat Rev Neurosci* 2001, **2**(1):43-50.
296. Ricketts C, Zatyka M, Barrett T: **The characterisation of the human Wolfram syndrome gene promoter demonstrating regulation by Sp1 and Sp3 transcription factors.** *Biochim Biophys Acta* 2006, **1759**(7):367-377.
297. Bhattacharya A, Ziebarth JD, Cui Y: **PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways.** *Nucleic Acids Res* 2014, **42**(Database issue):D86-91.
298. Kozomara A, Birgaoanu M, Griffiths-Jones S: **miRBase: from microRNA sequences to function.** *Nucleic Acids Res* 2019, **47**(D1):D155-D162.
299. Liu X, Meng X, Peng X, Yao Q, Zhu F, Ding Z, Sun H, Liu X, Li D, Lu Y *et al*: **Impaired AGO2/miR-185-3p/NRP1 axis promotes colorectal cancer metastasis.** *Cell Death Dis* 2021, **12**(4):390.
300. Lu G, Li Y, Ma Y, Lu J, Chen Y, Jiang Q, Qin Q, Zhao L, Huang Q, Luo Z *et al*: **Long noncoding RNA LINC00511 contributes to breast cancer tumorigenesis and stemness by inducing the miR-185-3p/E2F1/Nanog axis.** *J Exp Clin Cancer Res* 2018, **37**(1):289.
301. Li Y, Zhong W, Zhu M, Li M, Yang Z: **miR-185 inhibits prostate cancer angiogenesis induced by the nodal/ALK4 pathway.** *BMC Urol* 2020, **20**(1):49.
302. Wei J, Zhao Y: **MiR-185-5p Protects Against Angiogenesis in Polycystic Ovary Syndrome by Targeting VEGFA.** *Front Pharmacol* 2020, **11**:1030.

303. Bao L, Fu X, Si M, Wang Y, Ma R, Ren X, Lv H: **MicroRNA-185 targets SOCS3 to inhibit beta-cell dysfunction in diabetes.** *PLoS One* 2015, **10**(2):e0116067.
304. Qi S, Wang X: **Decreased Expression of miR-185 in Serum and Placenta of Patients with Gestational Diabetes Mellitus.** *Clin Lab* 2019, **65**(12).
305. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP: **MicroRNA targeting specificity in mammals: determinants beyond seed pairing.** *Mol Cell* 2007, **27**(1):91-105.
306. Yamada T, Ishihara H, Tamura A, Takahashi R, Yamaguchi S, Takei D, Tokita A, Satake C, Tashiro F, Katagiri H *et al*: **WFS1-deficiency increases endoplasmic reticulum stress, impairs cell cycle progression and triggers the apoptotic pathway specifically in pancreatic beta-cells.** *Hum Mol Genet* 2006, **15**(10):1600-1609.
307. Bustin SA, Mueller R: **Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis.** *Clin Sci (Lond)* 2005, **109**(4):365-379.
308. Garg A, Ghoshal U, Patel SS, Singh DV, Arya AK, Vasanth S, Pandey A, Srivastava N: **Evaluation of seven commercial RT-PCR kits for COVID-19 testing in pooled clinical specimens.** *J Med Virol* 2021, **93**(4):2281-2286.
309. Yamayoshi S, Sakai-Tagawa Y, Koga M, Akasaka O, Nakachi I, Koh H, Maeda K, Adachi E, Saito M, Nagai H *et al*: **Comparison of Rapid Antigen Tests for COVID-19.** *Viruses* 2020, **12**(12).
310. Pourfarzam M, Zadhoush F: **Newborn Screening for inherited metabolic disorders; news and views.** *J Res Med Sci* 2013, **18**(9):801-808.
311. Furnier SM, Durkin MS, Baker MW: **Translating Molecular Technologies into Routine Newborn Screening Practice (vol 6, 80, 2020).** *Int J Neonat Screen* 2021, **7**(4).
312. Alberry M, Maddocks D, Jones M, Abdel Hadi M, Abdel-Fattah S, Avent N, Soothill PW: **Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast.** *Prenat Diagn* 2007, **27**(5):415-418.
313. Scotchman E, Shaw J, Paternoster B, Chandler N, Chitty LS: **Non-invasive prenatal diagnosis and screening for monogenic disorders.** *Eur J Obstet Gynecol Reprod Biol* 2020, **253**:320-327.
314. Pratt M, Garritty C, Thuku M, Esmailisaraji L, Hamel C, Hartley T, Millar K, Skidmore B, Dougan S, Armour CM: **Application of exome sequencing for prenatal diagnosis: a rapid scoping review.** *Genet Med* 2020, **22**(12):1925-1934.
315. Peretz-Machluf R, Rabinowitz T, Shomron N: **Genome-Wide Noninvasive Prenatal Diagnosis of De Novo Mutations.** *Methods Mol Biol* 2021, **2243**:249-269.
316. Jiang F, Ren J, Chen F, Zhou Y, Xie J, Dan S, Su Y, Xie J, Yin B, Su W *et al*: **Noninvasive Fetal Trisomy (NIFTY) test: an advanced noninvasive prenatal diagnosis methodology for fetal autosomal and sex chromosomal aneuploidies.** *BMC Med Genomics* 2012, **5**:57.
317. De Rycke M, Berckmoes V: **Preimplantation Genetic Testing for Monogenic Disorders.** *Genes (Basel)* 2020, **11**(8).

318. de Rojas I, Moreno-Grau S, Tesi N, Grenier-Boley B, Andrade V, Jansen IE, Pedersen NL, Stringa N, Zettergren A, Hernandez I *et al*: **Common variants in Alzheimer's disease and risk stratification by polygenic risk scores.** *Nature communications* 2021, **12**(1):3417.
319. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT *et al*: **Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations.** *Nat Genet* 2018, **50**(9):1219-1224.
320. Shan N, Xie Y, Song S, Jiang W, Wang Z, Hou L: **A novel transcriptional risk score for risk prediction of complex human diseases.** *Genet Epidemiol* 2021, **45**(8):811-820.
321. Valihrach L, Androvic P, Kubista M: **Circulating miRNA analysis for cancer diagnostics and therapy.** *Mol Aspects Med* 2020, **72**:100825.
322. Alessandrini L, Manchi M, De Re V, Dolcetti R, Canzonieri V: **Proposed Molecular and miRNA Classification of Gastric Cancer.** *Int J Mol Sci* 2018, **19**(6).
323. Volovat SR, Volovat C, Hordila I, Hordila DA, Mirestean CC, Miron OT, Lungulescu C, Scripcariu DV, Stolniceanu CR, Konsoulova-Kirova AA *et al*: **MiRNA and LncRNA as Potential Biomarkers in Triple-Negative Breast Cancer: A Review.** *Front Oncol* 2020, **10**:526850.
324. Wang H: **MicroRNAs, Multiple Sclerosis, and Depression.** *Int J Mol Sci* 2021, **22**(15).
325. Chen YH, Wang H: **The Association between Migraine and Depression based on miRNA Biomarkers and Cohort Studies.** *Curr Med Chem* 2021, **28**(27):5648-5656.
326. Liang S, Li X, Gao C, Zhang L: **microRNA-based autophagy inhibition as targeted therapy in pancreatic cancer.** *Biomed Pharmacother* 2020, **132**:110799.
327. Ramelli SC, Gerthoffer WT: **MicroRNA Targets for Asthma Therapy.** *Adv Exp Med Biol* 2021, **1303**:89-105.
328. Tual-Chalot S, Stellos K: **MicroRNA-based therapy of postmyocardial infarction heart failure.** *Hellenic J Cardiol* 2021, **62**(2):149-151.
329. Lee JA, Cho A, Huang EN, Xu Y, Quach H, Hu J, Wong AP: **Gene therapy for cystic fibrosis: new tools for precision medicine.** *J Transl Med* 2021, **19**(1):452.
330. Zhang B: **CRISPR/Cas gene therapy.** *J Cell Physiol* 2021, **236**(4):2459-2481.

10 FÜGGELÉK

S1. táblázat. A bemutatott kísérletek során alkalmazott primerek. T_a : a PCR során alkalmazott annealási hőmérséklet, F : forward, R : reverse primer.

<i>Alkalmazás</i>	<i>Szekvencia (5'–3')</i>	T_a (°C)
Humán <i>DRD4</i> 27 bp deléció genotipizálása	F : GGAATGGAGGAGGGAGCGGG R : GACGCCAGCGCCATCCTACC	65
Kutya <i>DAT</i> 9. intron VNTR genotipizálása	F : CTCCTGTGTCCCCGCTGTCTT R : GACAGAGCAGGGCAGGGAGG	65
Kutya <i>DBH</i> 4. intron VNTR genotipizálása	F : CCCCTCACCTCCAAGCAG R : AGGGTGATGTGGGCAGGAT	56
Kutya <i>TH</i> 4. intron VNTR genotipizálása	F : GTCTGTCTGCTGTCTGGCTCCC R : TGGAGAGGCTTCCTGACACCC	65
Kutya <i>DRD4</i> 3. intron VNTR genotipizálása 3a és 3b allél elkülönítése:	F : CGCGCGTCGGGCCAAGCTG R : GCGGGGGGCAGGGGGCG R : TGGGCTGGGGGTGCCGTCC	67
Kutya <i>DRD4</i> 1. exon 24 bp in / del genotipizálása	F : CGCCATGGGGAACCGCAG R : CGGCTCACCTCGGAGTAGA	60
<i>WFS1</i> rs148797429 in / del genotipizálása	F : TGTTACAGGAGAAGGCGAGC R : CGAAAGTTTCCGAGCGGACT	58
<i>C4A</i> és <i>C4B</i> génszám meghatározás <i>C4A</i> -specifikus próba: <i>C4B</i> -specifikus próba:	F : GCAGGAGACATCTAACTGGCTTCT R : CCGCACCTGCATGCTCCT <i>VIC</i> –ACCCCTGTCCAGTGTTAG– <i>MGB</i> <i>FAM</i> –ACCTCTCTCCAGTGATAC– <i>MGB</i>	60
<i>NR1I2</i> (5' UTR) génszám meghatározás TaqMan próba:	F : TTGTGAGCCAGAAGGGATTTG R : GGCATGACTCCAGCTCAGGTA <i>FAM</i> –TACGTATGTTACCCATAACC– <i>MGB</i>	60
<i>GSK3B</i> (promoter) génszám meghatározás TaqMan próba:	F : GGCCACTGTTGCCATTGTC R : AGCCGATCAGCCTGAGAAAC <i>VIC</i> –CCAGCCAGCCACCGA– <i>MGB</i>	60
<i>DRD4</i> rs1800955 SNP genotipizálás <i>BssH</i> II RFLP-vel	F : GGAATGGAGGAGGGAGCGGG R : CGCTCCACCGTGAGCCAGTAT	65
<i>DRD4</i> 3. exon 48 bp VNTR genotipizálása	F : GCGACTACGTGGTCTACTCG R : AGGACCCTCATGGCCTTG	60
<i>DRD4</i> rs1800955–rs747302 haplotipizálás: rs747302 C-: illetve T-specifikus primer:	F : GGAATGGAGGAGGGAGCGGG R : GCCTCGACCTCGTGCGCG R : GCCTCGACCTCGTGCGCA	65
D11S1984 mikroszatellita genotipizálása	F : GGGTGACAGAGCAAATTCT R : ACACCTGGATCTTGACTCA	56

Alkalmazás	Szekvencia (5'-3')	T _a (°C)
<i>WFS1</i> rs4689388 SNP genotipizálása SBE-primer:	F: GCTCACCTCTGTTTCCCCTCCAGCTCT R: GCACTTCCCAAACCCTTTGCAGGGGC F: TGCCTGTTGCTTTTTTCAGTCTCTTAGCTCC	63 55
<i>WFS1</i> rs4273545 SNP genotipizálása SBE-primer:	F: GGGGACGCTGAGGCCGGAGTGG R: CTGCGGCCCGGGGGGAGCAATC F: CTCGGGCCCAACCACGCAGGGG	63 55
<i>WFS1</i> rs1046320, 1046322 és rs9457 SNP-k genotipizálása rs1046320 SBE-primer: rs1046322 SBE-primer: rs9457 SBE-primer:	F: GAGCACGACTGGCGCAGCACCG R: CCCACAGCCGCCCCACAGGGT F: (A) ₇ GAGACCTTGCGACCATGTGTAGATTGC F: (A) ₁₅ TCTCCACCCTGAGCCTGACCTTTCTGA R: (A) ₁₉ CTTCACACATTCACCCACTACCTGCTGGA	63 55
<i>SNAP-25</i> rs3746544 és rs1051312 genotipizálása	F: TTGTGTCTTTTGTCTCT R: AAGGATGATTATGATTTAGA	51
<i>SNAP-25</i> rs3746544 és rs1051312 haplotipizálása T-T-specifikus próba T-C-specifikus próba G-T-specifikus próba G-C-specifikus próba	F: TTCCAAAGGTTGTACATAGTGGTCAT R: AAATGCCACCGAGGAGAGAA FAM-CCTTGATGTCTTGAGTT-MGB VIC-TCCTTGATGTCTTGAGTT-MGB FAM vagy NED-CCTTGAGGTCTTGAGTT-MGB VIC-CCTTGAGGTCTTGAGTT-MGB	60
<i>SNAP-25</i> rs6077690 SNP genotipizálása	F: ATGTCAGTGTGGGGCATC R: AGGCATGTTGCTGAAATTTGTT	63
<i>SNAP-25</i> 3' UTR klónozása pMIR-Report vektorba	F: TGTAATACGCGTCTGGGAAGTGGTTAAGTGT R: CCCGACAAGCTTAAACTAGCTACAAAATGTCAATCA	56
<i>SNAP-25</i> 3' UTR mutagenézis: T-C haplotípus létrehozása	F: GAAAAATGAAACTCAAGACGTCAAGGAGTTAGAGC R: GCTCTAACTCCTTGACGTCTTGAGTTTCA-TTTTTC	52
<i>SNAP-25</i> 3' UTR mutagenézis: G-T haplotípus létrehozása	F: GAAAAATGAAACTCACGACATCAAGGAGTTAGAGC R: GCTCTAACTCCTTGATGTCGTGAGTTTCA-TTTTTC	52
<i>SNAP-25</i> 3' UTR mutagenézis: G-C haplotípus létrehozása	F: GAAAAATGAAACTCACGACGTCAAGGAGTTAGAGC R: GCTCTAACTCCTTGACGTGTCGTGAGTTTCA-TTTTTC	52
<i>SNAP-25</i> kódoló szakasz klónozása pcDNA 3.1(-) vektorba	F: TCCCCTCGAGCCATGGCCGAAGACGCAGAC R: CAGCATGGATCCGAGAACACGGGTGGGCACACTTA	60
<i>SNAP-25</i> izoformák mérése <i>Dde</i> I PCR-RFLP-vel	F: TGTCTTTCCTCCCTCCCTGCTC R: CATCTGCTCCCGTTCGTCCACT	62

Alkalmazás	Szekvencia (5'–3')	T_a (°C)
<i>SNAP-25a</i> izoforma specifikus (real-time) PCR	<i>F</i> : GAAGGCATGAACCATATCAACC <i>R</i> : GCCACAGCATTTCCTAAATCTT	60
<i>SNAP-25b</i> izoforma specifikus (real-time) PCR	<i>F</i> : GAAGGGATGGACCAAATCAATA <i>R</i> : CTTGTTACAGGGACACACAC	60
<i>WFS1</i> 5' régió klónozása pGL3 Basic vektorba (<i>F1</i> : „A”, <i>F2</i> : „B” konstrukció létrehozása)	<i>F1</i> : GCAGCTCGAGGTCCGCTCGGAACTTTCGCTGTGG <i>F2</i> : GCCCTTCTCGAGTCTTGCAGCGCCGGAATAGGC <i>R</i> : GCAGAAGCTTAAGTTGTTTCGGGAGCAGCTGAACG	57
<i>WFS1</i> 5' régió mutagenézis rs4273545 G létrehozása	<i>F</i> : CCACGCAGGGGGAGTGCCGGGCC <i>R</i> : GGCCCGGCACTCCCCCTGCGTGG	70
<i>WFS1</i> 5' régió mutagenézis rs148797429 del létrehozása	<i>F</i> : TTCCGCGGGCGGGGCGGGGC <i>R</i> : GCCCCGCCCGCCCGCGGAA	68
<i>WFS1</i> 3' UTR klónozása pMIR-Report vektorba	<i>F</i> : TCGGCGGAGCTCGGATGGTCCGCCACGAGGAGC <i>R</i> : AAAGGAAAGCTTGCCTGCAGGTTCCACAGAGG	57
pMIR-Report– <i>WFS1</i> 3' UTR rs 9457 C → G mutagenézis	<i>F</i> : CCTCATGACCCTCCTGTCCAGCAGGTAG <i>R</i> : CTACCTGCTGGACAGGAGGGTCATGAGG	65,8
pMIR-Report– <i>WFS1</i> 3' UTR rs 9457 „seed mutáns” létrehozása	<i>F</i> : CAGGCTGCCTCATGACCCTCGAGAGGTGCAGGTAGTGGGTGAATGTG <i>R</i> : CAGGCTGCCTCATGACCCTCGAGAGGTGCAGGTAGTGGGTGAATGTG	75
pMIR-Report– <i>WFS1</i> 3' UTR rs1046322 G → A mutagenézis	<i>F</i> : CCTGAGCCTGACCTTCTGAATGACATGGGTG <i>R</i> : CACCCATGTCATTCAGAAAGGTCAGGCTCAGG	66
pMIR-Report– <i>WFS1</i> 3' UTR rs1046322 „seed mutáns” létrehozása	<i>F</i> : CAGGCTGCCTCATGACCCTCGAGAGGTGCAGGTAGTGGGTGAATGTG <i>R</i> : CACATTCACCCACTACCTGCACCTCTCGAGGGTCATGAGGCAGCCTG	75
<i>WFS1</i> 3' UTR klónozása pGL3-Control vektorba	<i>F</i> : GCGTCTAGAGGATGGTCCGCCACGAGGAGC <i>R</i> : AAAGGATCTAGAGCGCTGCAGGTTCCACCAAGAGG	57