

FÜREDI SÁNDOR

A magyarországi bűnügyi DNS-profil nyilvántartás találatkeresési módszerének fejlesztése

Absztrakt

Az igazságügyi DNS-vizsgálatok érzékenységének az elmúlt évtizedben megvalósult folyamatos növekedése a Nemzeti Szakértői és Kutató Központban is lehetővé tette nagyon kis mennyiségű, akár látens vagy bomlott biológiai anyagmaradványok rutinszerű szakértői vizsgálatát. A kirendelő hatóságok részéről is egyre nagyobb az igény az ilyen típusú, például érintési, érintkezési, megragadási nyomok DNS-elemzésére. Ugyanakkor az ismeretlen személyi eredetű DNS-profilok egyre nagyobb hányada a bűnügyi nyilvántartásunkban jelenleg automatikusan nem kerestethető. Ennek fő oka az, hogy a nyilvántartási szoftverünk (CODIS) kizárólag csak a DNS-profilalkotó tulajdonságok (allélok) számszerű egyezését vizsgálja. Az általam kidolgozott újszerű eljárással a találatkeresés és -hitelesítés módszerét továbbfejlesztettem oly módon, hogy az elemzés már nemcsak az allélszámok egyezését, hanem az egyezést mutató alléltulajdonságok népességben való előfordulásának gyakoriságát és az allélok intenzitását is vizsgálja. E valószínűségi alapú (probabilisztikus) találatkeresési módszer alkalmazásával a nyilvántartásba kerülő ismeretlen, egy vagy több személytől származó DNS-profilokra kapott, személyazonosító CODIS-találatok számát jelentősen meg lehetett növelni.

Kulcsszavak: bűnügyi DNS-vizsgálat, többszemélyi eredetű (kevert) biológiai anyagmaradványok, DNS-profil nyilvántartás, találatkeresés

Abstract

The steady increase in the sensitivity of forensic DNA analysis over the past decade has also enabled routine forensic testing of very small amounts

of even latent or degraded biological traces at the Hungarian Institute for Forensic Sciences. There is also an expanding demand from the authorities for DNA analysis of this type of evidence, e.g. touch, contact, grasp samples. However, an increasing proportion of unidentified DNA profiles are currently not automatically searchable in our criminal records. The main reason for this is that our databasing software (CODIS) only checks for numerical matches of DNA profile characteristics (alleles). The presented novel procedure for searching and validating matches now also looks at the peak intensity and the population frequency of matching alleles. By applying this probabilistic searching method, the number of CODIS matches obtained for unknown single source and mixed DNA profiles could be significantly increased.

Keywords: forensic DNA analysis, mixed biological traces, DNA database, database search

Bevezetés

A Rendőrség Tudományos Tanácsának 2023. évi pályázatán III. helyezést elért pályaművem szerkesztett változatát tartja a kezében vagy látja a képernyőjén a Tisztelt Olvasó. Formai okoknál fogva és a gördülékenyebb olvashatóság érdekében az igazságügyi DNS-profil vizsgálatok és a magyarországi DNS-profil nyilvántartások bemutatásáról szóló fejezeteket e tanulmány függelékébe helyeztem át. A függelék tanulmányozása segítheti a laikus olvasót a fő szövegben használt szakkifejezések, szakmai fordulatok értelmezésében.

Az igazságügyi DNS-vizsgálatok érzékenységének az elmúlt évtizedben megvalósult folyamatos növekedése a Nemzeti Szakértői és Kutató Központban (NSZKK) is lehetővé tette nagyon kis mennyiségű, akár szabad szemmel nem látható (látens) vagy bomlott biológiai anyagmaradványok rutinszerű szakértői vizsgálatát. A kirendelő hatóságok részéről is egyre jobban növekszik az igény az ilyen típusú, például érintési, érintkezési,

megragadási nyomok DNS-elemzésére. A vizsgált bűnjeltárgyak sokszor tartalmaznak már a bűncselekmény előtt rájuk kerülő, úgynevezett háttér vagy rendszeres használói DNS-t is (Gill et al., 2020). Ezek miatt azonban a szakértői vizsgálatra kerülő biológiai anyagmaradványok jelentős részéből egyre növekvő arányban rossz minőségű egyszemélyi, illetve több személytől származtatható (úgynevezett kevert) DNS-profil mutatható ki. Ennek megfelelően az ismeretlen személyi eredetű DNS-profilok egyre nagyobb hányada a bűnügyi nyilvántartásunkban jelenleg automatikusan nem kerestethető. Ennek fő oka, hogy a nyilvántartási szoftverünk (CODIS) (URL1) kizárólag csak a DNS-profilalkotó tulajdonságok (allélok) számszerű egyezését vizsgálja. Az ilyen típusú egyszerű keresés azonban rossz minőségű, illetve kevert biológiai anyagmaradványok esetén nagyon sok véletlen (téves) egyezéssel járna együtt, amely problémára az Európai Forenzikus Intézetek Hálózatának (ENFSI) jelenleg érvényben lévő szakmai ajánlása (URL2) is felhívja a figyelmet. A téves találatok várható magas gyakorisága miatt a nyilvántartásunkban sem a múltban, sem jelenleg rossz minőségű egyszemélyi, illetve többszemélyi ismeretlen DNS-profilra, profilkomponensre nem végeztünk és nem végzünk automatikus találatkeresést. A probléma megoldása érdekében az elmúlt évtizedben a szakterület kutatói kidolgoztak és folyamatosan fejlesztenek olyan módszereket, amelyek a találatkeresés során már nemcsak az allélszámok egyezését, hanem az egyezést mutató alléltulajdonságok népszerűségben való előfordulásának gyakoriságát is vizsgálják (Bright et al., 2014; Benschop et al., 2017; Taylor et al., 2021). Ezen módszerek alkalmazása során minél ritkább allélokat érint egy találat, értelemszerűen annál valószínűbb a származás valódisága, és így az ilyen egyezés – mint a legvalószínűbb – a találati lista elejére kerülhet. E valószínűségi alapú találatkeresési módszerek [például Smart-Rank (Benschop et al., 2017)] felhasználhatóságát már az előbb hivatkozott ENFSI-ajánlás is említi. Jelenleg az NSZKK Genetikai Szakértői Intézetében két olyan szoftverrel rendelkezünk {Familias (URL3), DNAXs (Benschop et al., 2019)}, amelyek alkalmasak lehetnek valószínűségi (úgynevezett likelihood ratio, LR) alapú találatkeresésre. A Familias szoftver –

a rutin rokonságvizsgálataink alapszoftvere – képes elvégezni a valószínűségi találatkeresést félfolytonos (kvalitatív) statisztikai modellű LR-számítással. Ehhez viszont az szükséges, hogy mind a vizsgált ügyből, mind a CODIS-ből ki kell exportálni az érintett profilokat, és azokat be kell tölteni a Familias programba az egyszeri találatkeresés céljából. A kevert DNS-profilok találatkeresési projektjének elindításakor meg kellett teremteni annak a lehetőségét, hogy a Familias szoftver kezelni tudja a CODIS-profilokat. Mindezek mellett a statisztikai számításokhoz szükséges populációs allélgyakorisági adatbázist is létre kellett hozni és folyamatosan bővíteni a Familias program számára. Tekintettel arra, hogy a találatkeresést nagyszámú (több, mint 200 000) CODIS személyi DNS-profil között, tömegszűrősszerűen terveztem elvégezni, a valószínűsítési alapú találatkeresés ellenére számolni kellett a véletlen (azaz nem valós nyomhagyásból eredő) egyezések előfordulására is. Ezek hatékony kiszűrésére lehet felhasználni a DNAXs szakértői szoftver rendszert oly módon, hogy a Familias által mutatott kvalitatív egyezés a DNAXs-ben kvantitatív LR-statisztikával is vizsgálatszerűsítésre kerül. Amennyiben az allélintenzitások figyelembevételével kapott LR-érték is kellően magasnak bizonyul, illetve a számítás átmege a modellilleszkedési validációs teszten, akkor a beazonosított személy nyomhagyása a genetikai bizonyíték révén megerősíthető. Ehhez ugyanakkor a DNAXs szoftver használatát az ENFSI vonatkozó legjobb gyakorlata (URL4) szerint Intézetünkben verifikálni kellett, amely a találatkeresési tudományos projekttel párhuzamosan megvalósult (Füredi, 2023).

A kevert minták találatkeresési projektjében a vizsgálandó DNS-profilok körét két csoportra osztottam. Az egyik csoportba azok a profilok kerültek, amelyeket a bűnügyi DNS-profil nyilvántartásba korábban már feltöltöttek. Ennek a nyilvántartotti csoportnak a vizsgálata hármas célt szolgált. Az eddig ki nem mutatott egyezések megtalálásán túl kíváncsi voltam arra, hogy az LR-alapú találatkeresési módszerrel szintén kimutathatók-e azok az egyezések, amelyeket a CODIS ebben a csoportban már korábban beazonosított. Erre a CODIS-ban azért volt lehetőség, mivel ebben a DNS-

profil csoportban CODIS-találatkeresésre is alkalmas, egyszemélyi profilkomponenst tartalmazó, úgynevezett kijelölt keverékek (assigned mixture) is voltak. A Familias szoftver ugyanakkor az ilyen típusú LR-alapú keresetében nem tudja figyelembe venni az egyszemélyi DNS-profilkomponenseket, mint többletinformációt, mivel a CODIS-szoftverból a profilok ezzel a többletinformációval nem exportálhatók ki. Amennyiben ennek ellenére a Familias program a többletinformáció nélkül is képes megtalálni azt az egyezést, amit a CODIS korábban már kimutatott, akkor ez bizonyíthatja a kevert DNS-profilokban a magyarországi bűnügyi DNS-profil nyilvántartás terhelti DNS-profiljaival elvégzett LR-alapú, tömegszűrés szintű találatkeresési módszer alkalmasságát és létjogosultságát. Tekintettel arra, hogy a CODIS adatbázisában nemcsak a kevert DNS-profilok egy része, hanem a rossz minőségű egyszemélyi profilok sem vesznek részt a folyamatos és automatikus találatkeresésben, ezért célul tűztem ki azt is, hogy az LR-alapú módszerrel ezekre a profilokra is rákeressek. A rossz minőségű egyszemélyi DNS-profilokban az artefaktumok (az úgynevezett allélkiesés és -beesés) lehetősége olyan szintű, hogy ezeket a profilokat az ügyekben eljáró szakértők a CODIS típusú összehasonlításokra alkalmatlannak találták, és az ennek megfelelő CODIS-indexbe kezdeményezték a nyilvántartás kezelőjénél a profilok adatbázisba vételét. Ahogy azonban a függelékben szerepelt, a fél- és teljesen folyamatos statisztikai modellű találatkeresés figyelembe tudja venni az allélbe- és -kieséseket, kimutatva azokat a részleges egyezéseket is, ahol az összehasonlításra kerülő DNS-profilok az előbbieken említett okok miatt nem teljesen azonosak.

A kevert minták találatkeresési projektjében a vizsgálandó DNS-profilok másik csoportját értelemszerűen a még nem nyilvántartott, részben vagy teljes egészében ismeretlen személyi eredetű kevert DNS-profilok alkották. Ezek a DNS-profilok ugyan az ügyviteli és laboratóriumi információmenedzselő szoftverrendszerünk (LIMS) részét képezik 2007 óta, azonban ezek a profilok nem a CODIS-nak megfelelő (például kevert) indexekben vannak tárolva, hanem azokhoz az ügyekhez vannak csatolva, amelyek szakértése során meghatározták őket. Tekintettel arra, hogy ilyenformán a

kerestetésre alkalmas és indokolt, várhatóan nagy számú kevert DNS-profil LIMS-ből történő kinyerése rendkívül munka- és időigényes, illetve a 2019 előtt Intézetünkben vizsgált azon bűnügyeknek a jelentős része, ahol a legtöbb találat várható (vagyon elleni bűncselekmények), már valószínűleg elévült, ezért a vizsgálandó nem nyilvántartott kevert DNS-profilokat tartalmazó ügyek körét az Intézetünkbe 2019-től beérkezett ügyekre szűkítettem. Az volt a célkitűzésem, hogy éves bontásban (2019, 2020, 2021 és 2022) dolgozom fel a szóba jövő DNS-profilokat és értékelem ki a kapott találatokat a Familias és a DNAXs szoftver segítségével.

A vizsgált DNS-profilok

A vizsgált DNS-profilok darabszámra vonatkozó adatait az 1. számú táblázat mutatja be. A nyilvántartásban lévő személyi (terhelti) DNS-profilokat három időpontban (2022. 11. 29., 2023. 07. 06. és 2023. 08. 24.) bocsátotta a nyilvántartás a rendelkezésemre. Látható, hogy az idő előrehaladtával a nyilvántartásban tárolt személyi profilok száma növekszik. A teljesség igénye miatt az általam megkapott személyi DNS-profilok egy része több, esetenként különböző lokuszkészletű változatban is megtalálható volt a CODIS-ből kinyert adatsorban. Ezek olyan multiplikátumok, amelyek ugyanazoknak a személyeknek különböző időpontban és/vagy különböző kittel történő DNS-profil vizsgálati eredményei. A multiplikátumok egy-egy DNS-profil kivételével nem vesznek részt a CODIS automatikus találatkeresésében, ugyanakkor hasznosak lehetnek a kevert profilok tömegszűrés szintű kerestetésében. A táblázatban azonos színnel (kék, zöld vagy sárga) vannak jelölve azok a személyi és biológiai anyagmaradvány DNS-profilok, amelyeket egymással összehasonlítottam.

Időszak	Nyilvántartott DNS-profil (db)		Nem nyilvántartott kevert DNS-profil (db)	Személyi DNS-profil (db)
	Kevert	Rossz minőségű egyszemélyi		
- 2022.11.29.	5181	581		284 450
- 2023.07.06.				296 481
- 2023.08.24.				298 243
2019			528	
2020			669	

1. számú táblázat
A vizsgált DNS-profilok számadatai

A CODIS-adatbázis megfelelő indexeiből a DNS-profilok tömegével ki-exportálhatók úgynevezett xml fájlformátumba. Az xml fájl egy olyan szövegfájl, amelyben egymás után egyes sorok, illetve sortömbök kódolják a DNS-profilokra vonatkozó információkat. Egy DNS-profil CODIS xml fájlja a következőképpen néz ki (csak 1 db lokusz adatait tüntettem fel):

```
<?xml version="1.0"?>
<CODISImportFile xmlns="urn:CODISImportFile-schema">
  <HEADERVERSION>3.3</HEADERVERSION>
  <MESSAGEYPE>Import</MESSAGEYPE>
  <DESTINATIONORI>HUFSD1087</DESTINATIONORI>
  <SOURCELAB>HUFSD1087 </SOURCELAB>
  <SUBMITBYUSERID>[...]</SUBMITBYUSERID>
  <SUBMITDATETIME>2023-08-24T11:58:00</SUBMITDA-
TETIME>
  <SPECIMEN SOURCEID="Yes" CASEID="DN KÓDDAL" PAR-
TIAL="false">
    <SPECIMENID>XX/YYYYYYYY</SPECIMENID>
    <SPECIMENCATEGORY>Convicted Offender</SPECI-
MENCATEGORY>
    <SPECIMENCOMMENT>--Nyilvántartásba véve "DN" kód-
dal</SPECIMENCOMMENT>
```

```
<LOCUS PARTIALLOCUS="false" KIT="COfiler">  
  <LOCUSNAME>CSF1PO</LOCUSNAME>  
  <READINGBY>[...]</READINGBY>  
  <READINGDATETIME>2005-06-02T14:30:00</REA-  
DINGDATETIME>  
  <ALLELE>  
    <ALLELEVALUE>12</ALLELEVALUE>  
  </ALLELE>  
</LOCUS>  
[...]  
</SPECIMEN>  
</CODISImportFile>
```

A *CODIS xml* fájl tartalmának összetartozó részeit a fenti szövegben azonos színnel tüntettem fel, illetve a felhasználói értelmezés szempontjából legfontosabb elemek félkövér betűtípussal vannak jelezve, és a jelentésük a következő:

SOURCELAB: A nyilvántartás kezelőjének a kódja, ez aktuálisan a magyarországi igazságügyi DNS-profil nyilvántartásé

SPECIMEN: Egy DNS-profil leírásának kezdő- és végpontja

SPECIMENID: Az anonim DNS-profil azonosító kódja

SPECIMENCATEGORY: A DNS-profil *CODIS* indexneve (a példában elítélt)

LOCUS: Egy lokusz leírásának kezdő- és végpontja

LOCUSNAME: A lokusz tudományos neve

ALLELE: Egy allél leírásának kezdő- és végpontja

ALLELEVALUE: A lokusz alléljának a típusa

A teljes DNS-profil leírásánál a zölddel és kékkel jelölt elemek annyiszor ismétlődnek megfelelő tartalommal, ahány alléleredményt tartalmazó

lokuszról áll a DNS-profil. Ehhez hasonlóan több DNS-profil együttes leírásánál a sárga-zöld-kék blokk ismétlődik a konkrét DNS-profilok adataival. Amint a fenti példából látható, a CODIS xml fájl az allélintenzitási értékeket nem mutatja, mivel azokat a CODIS-adatbázis nem tárolja.

A bűnügyekben szakértői vizsgálat alá eső biológiai anyagmaradványok nem nyilvántartott, kevert DNS-profiljait az ügyviteli LIMS szoftverünk-ből nyertem ki. Mind a CODIS, mind a Familias szoftver a vizsgálati (referencia- és biológiai anyagmaradvány) minták összehasonlítását kizárólag egy-egy darab DNS-profil összehasonlításával végzi el, azaz a vizsgálati mintát csakis egyetlen profil képviselheti. Ugyanakkor minőségbiztosítási okokból, illetve az alacsony DNS-tartalmú, kevert és/vagy bomlott anyagmaradványok esetében gyakran és véletlenszerűen előforduló, korábban részletezett allélbeesések és -kiesések miatt a vizsgálati mintákból legalább két darab (ügynevezett replikált) DNS-profilvizsgálatot végzünk. Biológiai anyagmaradványok analízise során gyakran előfordul, hogy az egy mintából meghatározott DNS-profil replikátumok alléltípusaikban és allélintenzitásaikban részben eltérnek egymástól. Az ilyen esetekben a szakértői értékelés után a DNS-profiloknak az összes alléltípusa (ügynevezett kompozit DNS-profil), vagy csak a megfelelően reprodukálódott (általában az esetek több mint felében detektálásra kerülő) része (ügynevezett konszenzus DNS-profil) kerül a DNS-profil nyilvántartásba feltöltésre vagy a Familias szoftverben összehasonlításra (Bekaert, Van Geystelen, Vanderheyden, Larmuseau & Decorte, 2012).

E tanulmány írásának időpontjáig az Intézetünkbe 2019-ben és 2020-ban beérkezett ügyeket tudtam feldolgozni. Ehhez 706 és 870 db, 2019-ben, illetve 2020-ban Intézetünkbe beérkezett olyan ügy adatait tekintetem át egyenként, amelyekben kevert DNS-profil volt kimutatható. Az érintett ügyszámokat és DNS-profil kódokat Intézetünk informatikusa a LIMS-ünkből annak alapján tudta számomra kilistázni, hogy kimutatható-e az ügyben olyan DNS-profil, amelyben legalább négy lokuszon legalább 3 allél volt detektálható. A ritka ügynevezett trialléliás (Clayton, Guest, Urquhart, & Gill, 2004) és az allélbeesési esetek kivételével egy személy egy

autoszómás STR-markeren legfeljebb két alléllal – egy apai és egy anyai típussal – rendelkezhet, így egy profilban több lokuszon legalább három db allél jelenléte már a biológiai anyagmaradvány többszemélyi eredetét jelentheti. Az ügyekben előterjesztett genetikai szakértői vélemények és DNS-nyilvántartási találati tájékoztatások áttanulmányozásával tudtam azt megállapítani, hogy a kérdéses kevert DNS-profil korábban már személyazonosítva lett-e. A korábbi személyazonosítás [pontosabban a nyomhagyás feltevésének genetikai bizonyítékkal történő (statisztikai) alátámasztása] már megtörténhetett az eljáró hatóság által megnevezett, feltételezett nyomhagyó (POI-) személy DNS-profiljával történő összehasonlítással, vagy a kevert DNS-profil kijelölt egyszemélyi komponensének nyilvántartásba való feltöltése után kapott hidegtalálattal. Amint az 1. számú táblázatban látható, a két év ügyeiben 528 és 669 db, addig még be nem azonosított kevert DNS-profil találtam, néha ügyön belül többet is. Ezeket a profilekat xml fájlba ügyenként tudtam kiexportálni a LIMS szoftverünkből. Ez úgy volt lehetséges, hogy korábban kérésre Intézetünk informatikusa és egyben a LIMS-ünk programozója be tudta kódolni az ügyviteli szoftverünkbe ezt a DNS-profil export lehetőséget is. Az ügyenként történő profilkinyerés ugyanakkor azt is jelentette, hogy a két év ügyeiből összesen 696 db xml DNS-profil fájl keletkezett, melyeknek a Familias szoftverbe való betöltése gyakorlatilag kivitelezhetetlen lett volna, mivel a szoftver egyszerre csak 1 db fájl adatait tudja beolvasni. Szerencsére a Windows Parancssorában (Command Prompt) az xml fájlok a `copy /b *.xml final.xml` paranccsal binárisan egy fájlá összefűzhetők, amely a Familias szoftverrel már kényelmesen kezelhető.

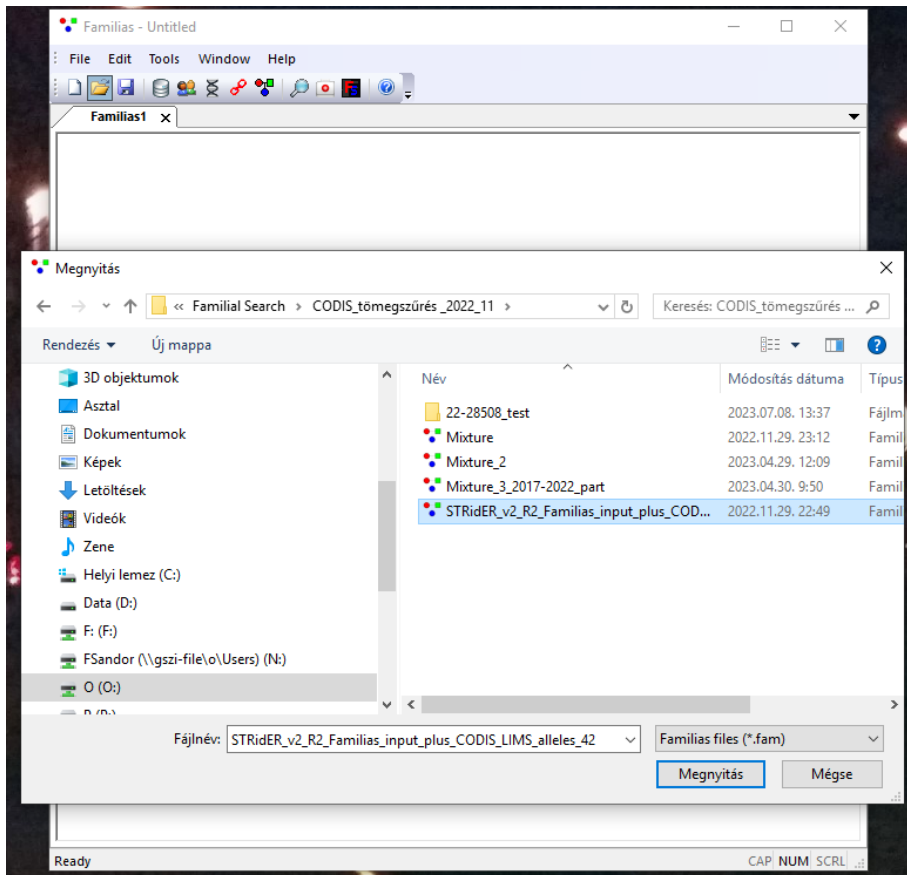
Találatkeresés a Familias v3.3.1 szoftverrel

A Familias program az igazságügyi autoszómás genetikai rokonságvizsgálatokban leggyakrabban használt ingyenes statisztikai szoftverek egyike. A szoftver tulajdonképpen egy több modulból álló programcsomag, melynek egyik része az úgynevezett rokonkeresési (Familial Searching) modul. A

rokonkeresési modul lehetőséget nyújt autoszómás genetikai profilok között meglévő közeli rokoni kapcsolatok (például szülő/gyermek, édestestvér, féltestvér) statisztikai (LR-) alapú, tömegszűrés szintű valószínűsítésére. A szoftver programozója a rokonkeresési modulban ezek mellett lehetővé tette azt is, hogy az összehasonlításra kerülő profilok közös személyi eredetét, a közvetlen egyezést (direct match) is ki lehessen mutatni fél-folytonos (kvalitatív) modellű statisztikai LR-számítás révén. A tömegszűrés szintű direct match vizsgálat csak egyszemélyi DNS-profilok, vagy egyszemélyi és többszemélyi profilok között lehetséges, tehát többszemélyi profilok így nem hasonlíthatók össze egymással. Az összehasonlításban van egy felső korlát az alléltípusok lokuszonkénti darabszámában, amely jelenleg nyolc, azaz négy személynél több személytől származó biológiai anyagmaradvány DNS-profiljával csak akkor végezhető elemzés a szoftverrel, ha a személyek alléljainak átfedése miatt a lokuszonkénti alléltípusok száma nem haladja meg a nyolcat. A Familias szoftver 2014 előtt nem tudta a CODIS xml fájlokat kezelni. Kérésemre a program fejlesztője, Daniel Kling lehetővé tette a DNS-profiloknak xml fájlból történő tömeges betöltését, amivel megvalósítható vált a CODIS adatbázisok DNS-profiljaival történő tömegszűrés rokonkeresési (Kling & Füredi, 2016) és direkt találatkeresési célokból is.

A Familias-nak e tanulmányban bemutatott találatkereséshez használt 3.3.1 verziójával való munka első lépéseként a programhoz létrehoztam és folyamatosan bővítettem a vizsgált lokuszok populációs allélgyakorisági adatbázisát. Ez az adatbázis a publikus európai STRidER 2.0 allélfrekvencia adatbázison (URL5) alapul, amelyet folyamatosan kiegészítettem az általunk megfigyelt új (azaz a STRidER adatbázisából hiányzó), ritka allélok adataival. A Familias-ba beimportált allélgyakorisági adatok külön projekt- (.fam kiterjesztésű) fájlba kimenthetők, és a konkrét találatkeresési vizsgálat előtt kényelmesen visszatölthetők a szoftverbe.

Füredi Sándor: A magyarországi bűnügyi DNS-profil nyilvántartás találatkeresési módszerének fejlesztése



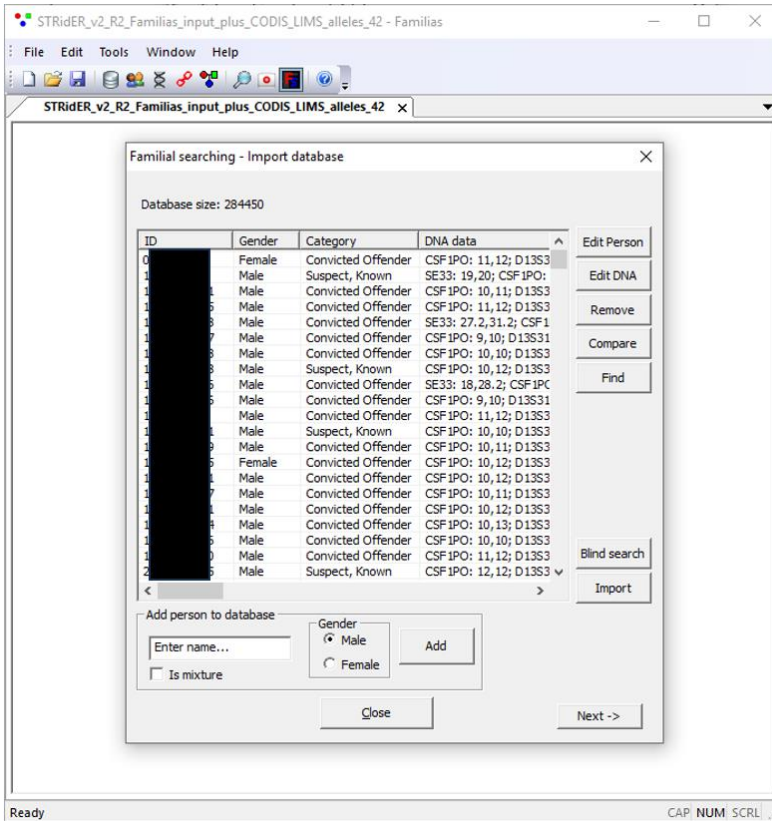
1. számú ábra

Az aktuális populációs allélgyakorisági adatbázis projekt fájl betöltése a Familias v3.3.1 szoftverbe a File/Open parancs segítségével

A találatkeresésben vizsgált személyi DNS-profilok a szoftver Tools / Familial searching menüpontjával vagy az **F** ikonnal megnyitható adatbázis ablak Import gombja segítségével tölthetők be (2. számú ábra).

Az ablak a nyilvántartott biológiai anyagmaradvány DNS-profilok kelesztéséhez használt, betöltött 284 450 db személyi DNS-profil első 21

tagját mutatja. Adatvédelmi okok miatt a személyi profilok nyilvántartási kódjai ki lettek takarva.

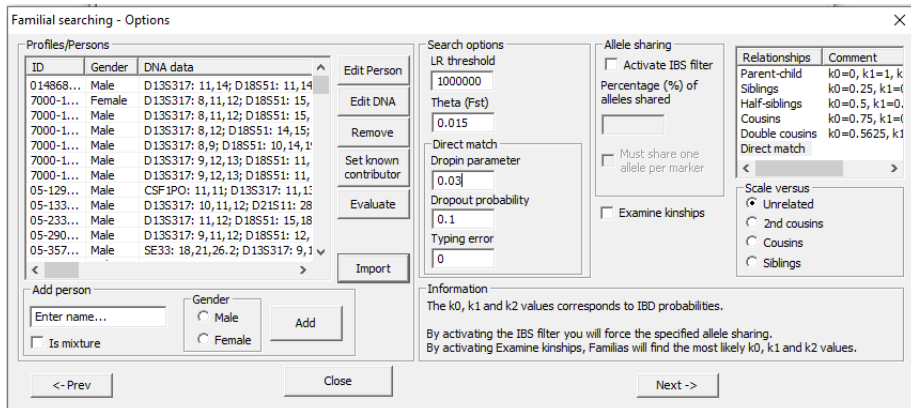


2. számú ábra

A Familias v3.3.1 szoftver Familial Searching moduljának személyi DNS-profilok betöltésére szolgáló Import database ablaka

A Next gomb megnyomásával megnyíló Options ablakban importálhatók be a biológiai anyagmaradványok DNS-profiljai az Import gomb segítségével.

Az ablak bal oldala a betöltött, nyilvántartott anyagmaradvány profilok első 12 tagját mutatja. Az ablak jobb oldalán lévő beállítások leírása a szövegben található.



3. számú ábra

A Familias v3.3.1 szoftver Familial Searching moduljának biológiai anyagmaradvány DNS-profilok importjára szolgáló Options ablaka

A 3. ábrán látható Options ablak jobb oldala találatkeresési beállításokra szolgál. A jobb felső panelben kiválasztható direkt találatok (Direct match) félfolytonos (kvalitatív) statisztikájú LR-számításához az allélbeesés valószínűségét a saját módszerverifikációnk során kapott maximum értékre, 0,03-ra (3%) állítottam, az allélkiesés valószínűségének pedig a nemzetközileg használt standard értéket, 0,1-et (10%) választottam. Annak érdekében, hogy minimalizáljam a nem valós (véletlen egyezésen vagy hasonlóságon alapuló) találatok számát, az LR-küszöbértéknek egy kellően magas számot, egymilliót használtam. A szoftver így csak azokat a DNS-profil párokat listázta ki, amelyek statisztikai összehasonlítása egymilliós, vagy annál nagyobb LR-értéket eredményezett. Az $LR = 1\,000\,000$ azt jelenti, hogy a biológiai anyagmaradvány DNS-profiljának kimutatási valószínűsége egymilliószer nagyobb akkor, ha az anyagmaradvány a POI (esetünkben az egyezést mutató nyilvántartott) személy DNS-ét tartalmazza ahhoz

képest, hogy vele rokonságban nem lévő ismeretlen személy a nyomhagyó. A közös leszármazási együtthatónak (Theta vagy Fst) a magyarországi népességre korábban meghatározott értéke (0,015) (URL6) szintén a nem valószínű találatok kiszűrését szolgálta, mivel a theta (Fst) csökkenti a kiszámított valószínűségi hányadost. A Next gomb megnyomásával nyílik meg a találatkereső Perform search ablak (4. számú ábra). Itt a Search gomb megnyomásával indítható a találatkeresés.

Az ablak a nyilvántartott biológiai anyagmaradvány DNS-profilokra kapott találatokat mutatja. A találati oszlopok jelentését a szövegben írtam le. Adatvédelmi okok miatt a személyi profilok nyilvántartási kódjai ki lettek takarva.

Database search - Perform search

Database size: 284450 Number of matches: 13

Profile/Trace	Candidate	Index	Gender	Relationship	LR	Exclusions	Overla
22-26059-002-01-H-LT/0-1 (mixture)	[REDACTED]	162364	Male	Direct-match	9.5940523e+019	0	
22-28696-002-05-H-LT/2-1 (mixture)	[REDACTED]	283506	Male	Direct-match	5.5464306e+018	0	
22-28503-001-01-H-LT/3-1 (mixture)	[REDACTED]	59081	Female	Direct-match	8.5933524e+016	0	
22-29166-001-01-H-LT/0-1	[REDACTED]	126858	Male	Direct-match	1.1810234e+016	6	
22-25468-001-04-H-LT/0-1 (mixture)	[REDACTED]	276635	Female	Direct-match	2.0112351e+015	0	
22-26988-002-01-H-LT/0-1 (mixture)	[REDACTED]	5896	Male	Direct-match	6.464439e+013	0	
22-26988-002-01-H-LT/0-1 (mixture)	[REDACTED]	128363	Male	Direct-match	4.344388e+012	0	
22-35822-002-01-H-LT/1-1 (mixture)	[REDACTED]	143574	Male	Direct-match	4.1015421e+012	0	
22-25468-001-02-H-LT/0-3 (mixture)	[REDACTED]	276635	Female	Direct-match	4.2497827e+011	0	
22-30251-002-02-H-LT/0-1 (mixture)	[REDACTED]	69885	Male	Direct-match	2.6857077e+009	1	
22-29155-001-01-H-LT/0-1 (mixture)	[REDACTED]	165508	Male	Direct-match	8.3449146e+008	2	
22-25468-001-04-H-LT/1-1 (mixture)	[REDACTED]	276635	Female	Direct-match	2.9950284e+008	0	
22-30306-003-01-H-LT/0-1 (mixture)	[REDACTED]	268313	Male	Direct-match	2.0318851e+008	0	

Search

Search

Sort

Subset

Display

Match

View match

Report match

Remove

Save summary

Export list

<- Previous Close

Füredi Sándor: A magyarországi bűnügyi DNS-profil nyilvántartás találatkeresési módszerének fejlesztése

Database search - Perform search

Database size: 284450 Number of matches: 13

	Exclusions	Overlapping markers	Shared alleles	IBS=2	IBS=1	IBS=0	Kinship	Kappa0	Kappa1	Kappa2
3e+019	0	21	66.7%	38.1%	57.1%	4.8%				
6e+018	0	21	78.6%	66.7%	23.8%	9.5%				
4e+016	0	22	52.3%	22.7%	59.1%	18.2%				
4e+016	6	17	82.4%	64.7%	35.3%	0.0%				
1e+015	0	21	69.0%	42.9%	52.4%	4.8%				
9e+013	0	21	42.9%	14.3%	57.1%	28.6%				
8e+012	0	21	54.8%	28.6%	52.4%	19.0%				
1e+012	0	15	66.7%	33.3%	66.7%	0.0%				
7e+011	0	20	70.0%	50.0%	40.0%	10.0%				
7e+009	1	20	67.5%	50.0%	35.0%	15.0%				
6e+008	2	21	61.9%	33.3%	57.1%	9.5%				
4e+008	0	11	72.7%	45.5%	54.5%	0.0%				
1e+008	0	15	46.7%	20.0%	53.3%	26.7%				

Search: Search, Sort, Subset, Display

Match: View match, Report match, Remove

Save summary, Export list

<- Previous Close

4. számú ábra

A Familias v3.3.1 szoftver Familias Searching találatkeresésre szolgáló Perform search ablaka két részletben

A 4. ábrán látható találati lista első, Profile/Trace oszlopa azoknak a biológiai anyagmaradvány DNS-profiloknak a kódjait mutatja, amelyekre LR-küszöbérték feletti egyezést (Relationship oszlopban Direct-match) adott a Candidate oszlopban nyilvántartási kóddal szereplő személy DNS-profilja. E személyek genetikai nemét a Gender oszlop jeleníti meg. A genetikai nem az autoszómás DNS-profil vizsgálatokhoz társított Amelogenin lokusz analízisen alapul (Akane et al., 1991). Az Amelogenin lokusz az ivari kromoszómákon helyezkedik el, és az Y-kromoszómás allélja hosszabb, mint az X-kromoszómás párja, amely méretkülönbség lehetővé teszi a két allél DNS-fragmensanalízissel történő elkülönítését, és így a mintaadó/nyomhagyó személy genetikai nemének meghatározását. A találatok aktuális LR-értékét az LR oszlop mutatja. Az Overlapping markers oszlop az összehasonlított két DNS-profil alléleredményt tartalmazó közös (átfedő) autoszómás lokuszainak számát jeleníti meg. Az Exclusions oszlopban lévő szám pedig annak az alléltípus számnak felel meg, amennyi

személyi minta allél hiányzik a biológiai anyagmaradvány eredményt mutató lokuszaiból. A találati ablakban szereplő többi adat jelen munka szempontjából nem bír jelentőséggel. Az ablakban szereplő találati lista az Export list gombbal szövegfájlba kiexportálható, amely aztán táblázatkezelő programok (például Excel) segítségével tovább értékelhető.

Az e tanulmányban leírt találatkereséseket három adagban végeztem el a nyilvántartott, valamint a 2019. és 2020. évi ügyek nem nyilvántartott biológiai anyagmaradvány DNS-profiljain a Familias szoftver segítségével. Tekintettel arra, hogy a legtöbb, nyilvántartott biológiai anyagmaradványt érintő DNS-profillal történő páros összehasonlítás mintegy 1,64 milliárd db statisztikai LR-számítást igényelt, és ehhez legerősebb asztali számítógépként egy Intel Core i5-10500 CPU @ 3.10 GHz központi proceszorú PC állt rendelkezésemre, így ez a számítássorozat több, mint egy hónapot vett igénybe. A Familias szoftver sajnos csak egy CPU magon tud számolni, és a kalkulációkhoz nem tudja igénybe venni a grafikus proceszort, ugyanakkor különállóan párhuzamosan több Familias is együtt tud futni az egyes CPU magokon, így – a rendelkezésre álló belső memória méretétől függően – több számítássorozat is elindítható ugyanazon a számítógépen.

A találatok statisztikai visszaellenőrzése a DNAXs v2.5.8 szoftverrel

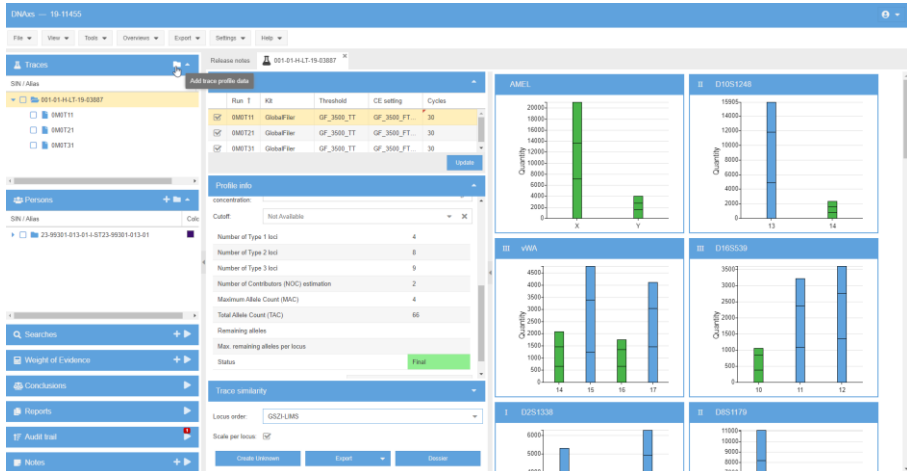
A teljesen folytonos (kvantitatív) statisztikai modellű DNS-profil kiértékelő DNAXs szakértői szoftverrendszer (Benschop, 2019) rutin szakértői gyakorlatban történő felhasználásának lehetőségeit a holland fejlesztőknek egy 2020-ben tartott webináriumban ismertem meg. A szoftver moduláris felépítésénél fogva a maga nemében egyedülálló igazságügyi szakértői szoftvercsomag (forensic expert system), mivel a laboratóriumi elemzés során meghatározott nyers DNS-profilok ügyekbe rendezve kezelhetők és értékelhetők a segítségével. Az értékelési lehetőségek és eszközök egyebek mellett magukban foglalják a DNS-profilok replikátumainak együttes ke-

zelésével történő profilvalidálást, a kompozit DNS-profil és annak konszenzus, illetve domináló (major) komponensének meghatározását, a profilok alléltípus alapján történő tömeges összehasonlítását, valamint a kvalitatív és kvantitatív statisztikájú személyazonosítási számításokat és azok eredményeinek modellilleszkedési teszten alapuló hitelesítését. Javaslatomra a DNAXs szoftvert Intézetünk beszerezte, és a program használatát bevezette a rutin szakértői gyakorlatunkban (Füredi, 2023).

Amint az a függelékben olvasható, a valószínűsítésen alapuló statisztikai genetikai személyazonosítás nem tudja kategorikusan kizárni a véletlenül alapuló DNS-profillegyezés lehetőségét. Ez a probléma, mint II. típusú (másodfajú) hiba tömegszűrés esetén fokozott mértékű lehet még félfolytonos statisztikai modellű LR-számításon alapuló találatkeresés esetén is. A találat hitelesítésének egy hatékony módszerét az allélintenzitások figyelembevételével történő statisztikai analízis jelentheti, mivel a kvalitatív számítással kapott LR-érték csak akkor marad megfelelően magas szinten, ha a biológiai anyagmaradvány DNS-profiljában az egyezést mutató feltételezett nyomhagyó személy alléljainak intenzitása összhangban van azon alléloknak egy személytől való származtathatóságával. A megfigyelt és várt allélintenzitások jelentős különbsége nemcsak az LR-értéket csökkenti, hanem a DNAXs modellilleszkedésre vonatkozó teszteredményét is negatívan befolyásolhatja.

Intézetünk informatikusa kérésre lehetővé tette a biológiai anyagmaradvány DNS-profiloknak a LIMS-ünkből DNAXs szoftverrel kompatibilis formában történő kiexportálását. A kinyert szöveges formátumú adatfájl a DNS-profilok mintakódján, lokusz- és alléltípus adatain kívül az allélok detektált jelerősségét is tartalmazza. A biológiai anyagmaradványok DNS-profiljai a DNAXs szoftver főablakának bal felső részén lévő Traces paneljében tölthetők be (5. számú ábra). A betöltött DNS-profilra vonatkozó szakmai adatok (például az elvégzett DNS-profilvizsgálat típusa, az anyagmaradvány feltételezett nyomhagyóinak száma) a főablak közepén jelenik meg, míg a DNS-profilnézetben jobb oldalon a DNS-profil lokuszainak alléloszlopdiagramjai láthatók. Minden oszlop egy-egy alléltípust képvisel,

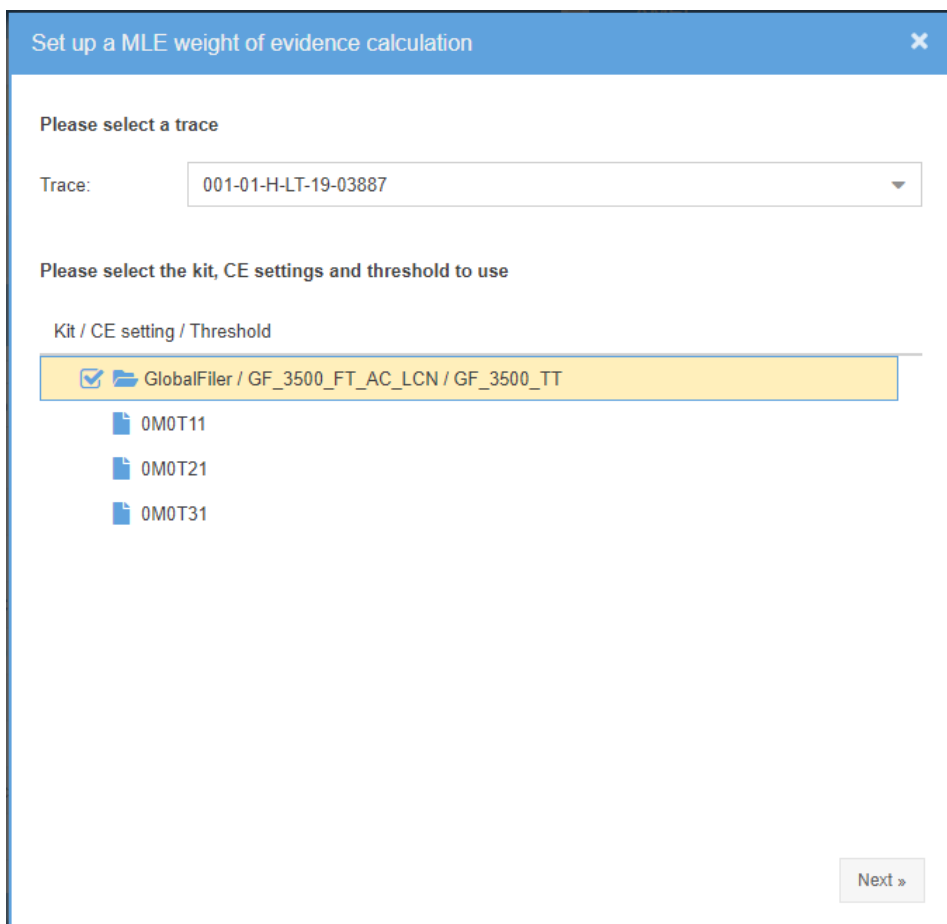
az oszlopok magassága pedig az allélok detektált intenzitását mutatja. Ha egy DNS-profil több replikátumból áll, akkor az azonos alléltípusú eredmények oszlopaikat egymásra helyezi a szoftver, így kumulálva azok allélintenzitásait, de közben vízszintes határoló vonallal jelezve az egyes replikátumok allélintenzitásainak nagyságát. A DNS-profil megfelelően replikálódott (konszenzus) allélkomponensét zölddel, a dominálón replikálódott (major) komponensét kézzel mutatja a szoftver. A nem megfelelően replikálódott allélok oszlopai piros színt kapnak. A személyi (egyszemélyi) DNS-profilok a DNAXs főablakának bal oldali középső Persons paneljében tölthetők be. Ez a panel – a Traces panelhez hasonlóan – egyelőre nem képes CODIS xml formátumú DNS-profilokat befogadni, ezért az Intézetünk informatikusa segítségével megvalósítottam a LIMS ügyviteli szoftverünkön keresztül a CODIS xml DNS-profiloknak a DNAXs-szel kompatibilis konverzióját.



6. számú ábra

A DNAXs v2.5.8 szoftver DNS-profilnézeti főablaka

A DNAXs-ben az MLE módszerű kvantitatív statisztikai számítások a főablak bal oldalán lévő Weight of Evidence panel + gombjának megnyomásával és az MLE módszer kiválasztásával indíthatók el. A felugró ablak (7. számú ábra) Trace sorában választhatjuk ki a listából a vizsgálni kívánt biológiai anyagmaradványt. Az ablak alsó fele mutatja a kiválasztott DNS-profil laboratóriumi meghatározására és replikátumaira vonatkozó kódokat. A Next gombbal nyithatjuk meg a feltételezett nyomhagyó személy(ek) (Person) megadását és a statisztikai számítás hipotéziseinek felállítását szolgáló ablakot (8. számú ábra). Az ábrán bemutatott példában a kiválasztott személyt nyomhagyónak (DNS-donornak) tekintjük a H1 feltevésben, míg az ellenhipotézisben (H2) ő nem nyomhagyó. Az ablak alján beállíthatjuk az összes nyomhagyó személy feltételezett számát (Set number of contributors), ami ebben az esetben kettő volt mindkét hipotézisre.



7. számú ábra

A DNAXs v2.5.8 szoftver biológiai anyagmaradvány kiválasztására szolgáló ablaka az MLE módszerű kvantitatív statisztikai számításához

Set up a MLE weight of evidence calculation

Please add persons from the menu (if applicable)

Person: SIN Add person

Person ↑	Contributed to H1	Contributed to H2
23-99301-013-01-I-ST23-99301-013-01	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Hypothesis parameters

Approach: Set number of contributors

H1: 2 H2: 2

« Previous Advanced settings » Calculate

8. számú ábra

A DNAXs v2.5.8 szoftver MLE módszerű kvantitatív statisztikai számításába bevonandó POI személy(ek) megadására és a származásra vonatkozó hipotézisek felállítására szolgáló ablak

A statisztikai számítás elindítása előtt az Advanced settings gomb megnyomásával felugró ablakban (9. számú ábra) kiválaszthatjuk és beállíthatjuk a következő statisztikai modelleket és paramétereiket (az e tanulmányban leírt vizsgálatokhoz az ábrán látható beállítást használtam):

- degradációs modell be- és kikapcsolása;

- stutter (úgynevezett dadogási PCR-artefaktum) modellek be- és ki-kapcsolása;
- a populációs allélgyakorisági adatbázis kiválasztása;
- a közös leszármazási együttható értékének megadása;
- a ritka (a populációs adatbázisban nem szereplő) allélok előfordulási gyakoriságának megadása;
- az allélbeesés (Drop-in) valószínűsége;
- annak az esélye, hogy egy beesett allél intenzitása éppen megegyezik az alléldetektálási küszöbértékkel.

Set up a MLE weight of evidence calculation

Please set the weight of evidence parameters

Degradation:

Backward stutter:

Forward stutter:

Population: STRidER_europe_2019-08-02

Coancestry coefficient: 0.015

Rare allele frequency: 0.001

Drop-in probability: 0.03


Drop-in peak height: 0.01

« Previous Calculate

9. számú ábra

A DNAXs v2.5.8 szoftver MLE módszerű kvantitatív statisztikai számításához szükséges modellparaméterek beállítására szolgáló ablak

A statisztikai számítás a Calculate gombbal indítható el. A számítás időigénye átlagos személyi számítógépen néhány másodperctől néhány óráig terjedhet attól függően, hogy mennyire összetett a biológiai anyagmaradvány DNS-profilja, és hány db replikátummal kell a kalkulációkat elvégezni. Az átlagos számítógépek számítási korlátai miatt négynél több személytől származtatott DNS-minta teljeskörű statisztikai elemzése a DNAXs szoftverrel sem lehetséges.

A statisztikai számítások eredményei a korábban említett Weight of Evidence panel  gombjának megnyomásával jeleníthetők meg a megnyíló Weight of Evidence ablakban (10. számú ábra). Az ablakban sorokba rendezetten láthatók az egyes számításokra vonatkozó adatok és eredmények, köztük:

- a biológiai anyagmaradvány (Source) kódja;
- a felállított hipotézisekben (Hypothesis 1, 2) szereplő nyomhagyó személy(ek) kódja;
- a megállapított LR-érték;
- az alkalmazott statisztikai számítás módszere (Calculation, esetünkben MLE).



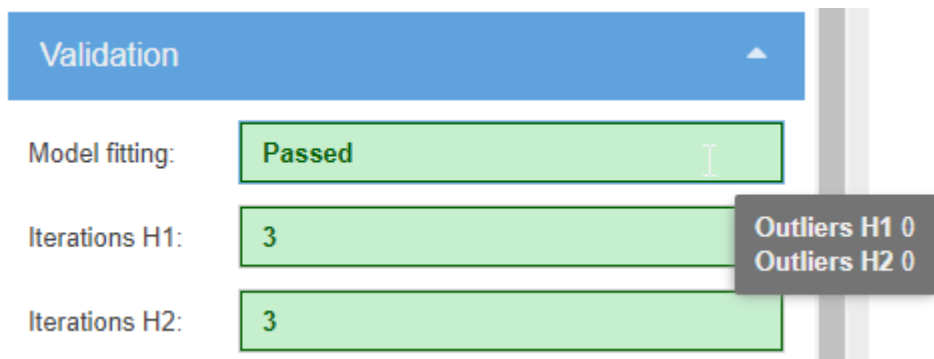
Status	ID	Source	Hypothesis 1	Hypothesis 2	Threshold 1	LR	Submitted by	Case number	Date	Run duration	Calculation	
✓	1520	001-01-H-LT-19-03887	23-99301-013-011-ST23-99301-013-011	1 unknown	2 unknown	GF_3500_TT	2.259e+19	root	19-11455	12/08/2023 @ 18:02	00:00:01	MLE

10. számú ábra

A DNAXs v2.5.8 szoftverrel végzett statisztikai számítások eredményeit összesítő Weight of Evidence ablak

Egy Weight of Evidence eredménySORRA történő dupla egérgattintással a DNAXs főablakának középső és jobb oldali részébe betöltődnek az adott statisztikai számításra vonatkozó részletes eredmények. Ez az ablak többek között az egyes lokuszokra külön számolva is megmutatja azok LR-értékeit, illetve a felállított két hipotézisre kiszámolt paraméterértékeket, köztük – kevert DNS-profilok esetében – az ismert és ismeretlen nyomhagyó személyek DNS-ének becsült relatív arányát a több személytől származó

biológiai anyagmaradványban. A találatkeresés eredményének statisztikai hitelesítése szempontjából ezen eredményablak Validation panelje a legfontosabb, amely megmutatja, hogy vajon az elvégezett statisztikai számítás átment-e a modellilleszkedési validációs teszten (11. számú ábra).

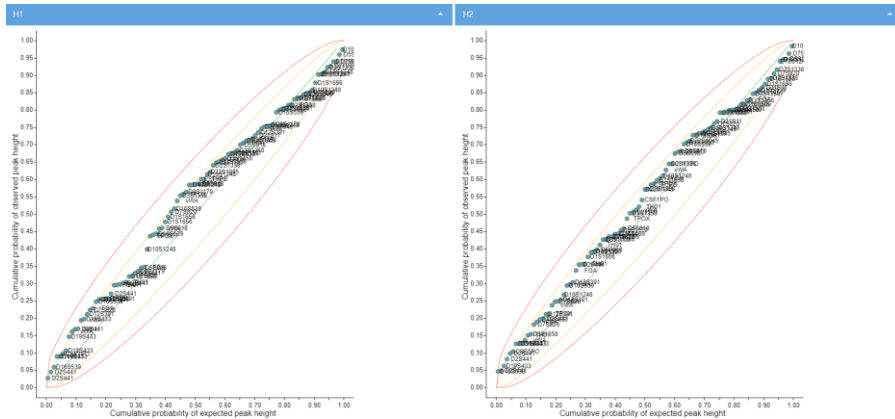


11. számú ábra

A DNAXs v2.5.8 szoftver MLE módszerű kvantitatív statisztikai számításához tartozó modellilleszkedési validációs teszteredményt bemutató panel

Amennyiben a Modell fitting vizsgálat Passed vagy Warning eredményt mutat, akkor a modellilleszkedés megfelelőnek tekinthető. Ezekben az esetekben a lokuszokon megfigyelt allélintenzitások (pontosabban azok kumulált valószínűségi értékei) – legfeljebb 2 lokusz kivételével – nem térnek el szignifikánsan a két hipotézishez megbecsült, elvárt intenzitásértékektől (12. számú ábra).

A kék pöttyök a vád (H1) és a védelem (H2) feltevéséhez kiszámolt allélintenzitási valószínűségek lokuszonként ábrázolva. A vízszintes tengely a várt, a függőleges tengely a megfigyelt allélintenzitásokra számolt valószínűségeket mutatja. A piros pöttyözött vonal által határolt terület a modellhipotézistől szignifikánsan nem eltérő megfigyelések régiója.



12. számú ábra

A DNAXs v2.5.8 szoftver MLE módszerű kvantitatív statisztikai számításához tartozó modellilleszkedési validációs teszteredmény grafikonjai

A Modell fitting vizsgálat Failed eredménye arra utal, hogy a felállított hipotézisünk (például a nyomhagyó személyére vagy a nyomhagyó személyek számára vonatkozóan) téves, vagy a DNS-profil rossz minőségénél fogva nem alkalmas kvantitatív modellű statisztikai elemzésre.

A kevert vagy rossz minőségű, egy személytől származó biológiai anyagmaradványok nyilvántartott DNS-profiljaira kapott találatok

2022. november 29-én a bűnügyi DNS-profil nyilvántartás CODIS adatbázisában szereplő összesen 5 762 db kevert vagy rossz minőségű, egy személytől származó biológiai anyagmaradvány DNS-profiljának a *Familias* szoftverrel történő kvalitatív LR-statisztikai alapú találatkeresésének eredményeképpen 1 024, illetve 67 db egyezést kaptam LR = 1 000 000 küszöbérték fölött 284 450 db személyi DNS-profillal történő összehasonlításban (2/a. számú táblázat).

a)

Nyilvántartott biológiai anyagmaradvány	Összes DNS-profil (db)	Találatot mutató DNS-profil (db)	%	Még kiadatlan találat (db)	Összes profilhoz képest (%)	Összes találathoz képest (%)
Kevert (többszemélyi)	5181	1024	20%	350	6.8%	34%
Rossz minőségű egyszemélyi	581	67	12%	28	4.8%	42%

b)

Nyilvántartott biológiai anyagmaradvány	Kiadatlan vagyon elleni találat (db)	Kiadatlan élet elleni találat (db)
Kevert (többszemélyi)	269	81
Rossz minőségű egyszemélyi	28	-

2. számú táblázat

5 762 db kevert vagy rossz minőségű, egy személytől származó biológiai anyagmaradvány DNS-profiljaira kapott találatok összesített adatai

Egyes anyagmaradvány DNS-profilok több referencia DNS-profillal is egyezést mutattak az LR-küszöbérték fölött, ezekben az esetekben csak a legmagasabb LR-értéket mutató (tudniillik a legvalószínűbb) találatot vettem figyelembe. Minden egyes találatot ügyenként kiértékeltem abból a szempontból, hogy a vizsgálatom időpontjáig az ügyben az érintett biológiai anyagmaradvány személyazonosítva lett-e az eljáró szakértő vagy a nyilvántartás részéről hideg- vagy melegtalálattal. A kiértékelés után azt kaptam, hogy a Familias szoftverrel találatot mutató kevert biológiai anyagmaradványok 66%-a, míg a rossz minőségű, egy személytől származó minták 58%-a korábban már beazonosításra került az általam is kimutatott személlyel. Az a tény, hogy ezt az ilyen számú és arányú azonosítási eredményt a korábbi eljáró szakértői, illetve CODIS találatkeresési módszertől függetlenül a Familias szoftver segítségével is el tudtam érni, igazolja ennek a kvalitatív LR-számításon alapuló statisztikai találatkeresési módszernek az érvényességét és hatékonyságát kevert vagy rossz minőségű biológiai anyagmaradványok esetében is.

A 2a táblázatban látható, hogy az esetek jelentős részében (34% és 42%) olyan találatokat sikerült kimutatnom, amelyek korábban még ismeretlenek

voltak. E biológiai anyagmaradványoknak a bűncselekmények típusával kiegészített részletes találati listáját is összeállítottam (nem közölt eredmény). A lista alapján több, mint háromszáz bűnügyben rögzített biológiai anyagmaradvány vétlen vagy elkövető nyomhagyójának személyazonosítására lenne lehetőség (2b táblázat). A beazonosított feltételezett nyomhagyók 89%-a férfi, 11%-a nő. Ez az arány hasonló, de a férfiak javára kicsit magasabb, mint a nyilvántartott férfiak részaránya, ami 83% a 2023. szeptember havi állapot szerint. A 2a táblázatban szereplő találati arányok egyben az is jelentik, hogy a CODIS találatkeresési hatékonyságát kevert minták esetében 52%-kal, míg rossz minőségű biológiai anyagmaradványok esetében 72%-kal tudtam növelni a Familias találatkeresési módszer segítségével. A kevert biológiai anyagmaradványokat érintő, még kiadatlan találatok 77%-a vagyon elleni, 23%-a élet elleni bűncselekményben volt kimutatható. A nyilvántartott rossz minőségű egyszemélyi DNS-profilok eddig ismeretlen találati kizárólag vagyon elleni ügyekben keletkeztek. Az összes, még ki nem adott találat bűncselekménytípusonkénti eloszlását a 3. számú táblázat tünteti fel.

Bűncselekmény típus	Arány
Lopás	66.4%
Rablás	9.5%
Jármű önkényes elvétele	3.7%
Testi sértés	2.9%
Szexuális bcs.	2.4%
Emberölés	2.1%
Kábítószerrel elkövetett bcs.	2.1%
Hamisítás	1.6%
Kifosztás	1.6%
Csalás	0.8%
Orgazdaság	0.8%
Rongálás	0.8%
Zsarolás	0.8%
Embercsempészés	0.5%
Garázdasság	0.5%
Közveszéllyel fenyegetés	0.5%
Robbanóanyaggal visszaélés	0.5%
Zaklatás	0.5%
Eltűnés	0.3%
Ittas járművezetés	0.3%
Kapcsolati erőszak	0.3%
Lőfegyverrel visszaélés	0.3%
Nem azonosítható	0.3%
Orvvadászat	0.3%
Segítségnyújtás elmulasztása	0.3%

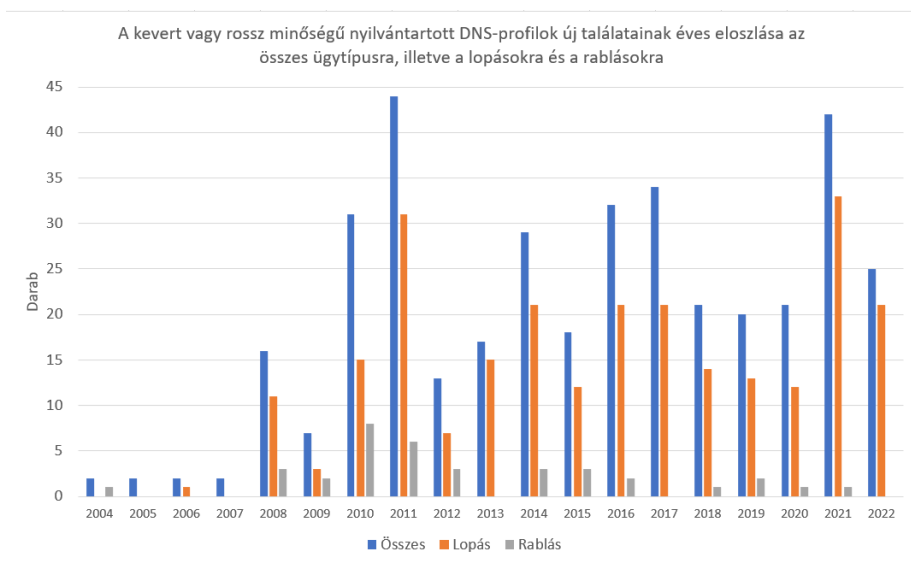
3. számú táblázat

A kevert vagy rossz minőségű egyszemélyi biológiai anyagmaradványok nyilvántartott DNS-profiljaira Familias szoftverrel kapott, eddig még nem ismert találatainak bűncselekménytípusonkénti eloszlása csökkenő sorrendben

A találati lista legelején toronymagasan a lopás áll, ami nem meglepő, mivel egyrészt a lopás bűnügyekből (mint ügynevezett tömegügyekből) kerül összességében a legtöbb biológiai anyagmaradvány DNS-vizsgálatra. Másrészt ezekben az ügyekben (főleg a sok érintkezési nyom miatt) ebben

az ügýtípusban a legnehezebb a nyomhagyó hagyományos nyomozati vagy szakértői módszerekkel történő beazonosítása. Örvedetes tény, hogy 2% feletti találati arányt kaptam több súlyosabb bűncselekménytípusban is, jelesül rablásban (9,5%), testi sértésben (2,9%), szexuális bűncselekményekben (2,4%), emberölésben és kábítószerrel kapcsolatos bűnygyekben (mindkettő 2,1%).

A 2022. évhez 2022. november 29-ig állt rendelkezésre adat.



13. számú ábra

A kevert vagy rossz minőségű egyszemélyi biológiai anyagmaradványok nyilvántartott DNS-profiljaira Familias szoftverrel kapott, eddig még nem ismert taláatainak éves eloszlása az összes ügýtípusra, illetve a lopásokra és rablásokra vonatkozóan

Az összes találatot, illetve a legtöbb találatot adó két ügýtípusra, a lopásra és rablásra kapott találatokat éves bontásban is kiértékeltem (13. számú ábra). A bünygi DNS-profil nyilvántartásunk szakrendszereinek beüzemelési évétől, 2004-től kezdve Intézetünkbe mindegyik évben beérkezett ügyekben ki tudtam mutatni eleddig ismeretlen találatokat, 2008-tól kezdve már nagyobb számban. Ennek valószínűsíthető oka a vizsgált és

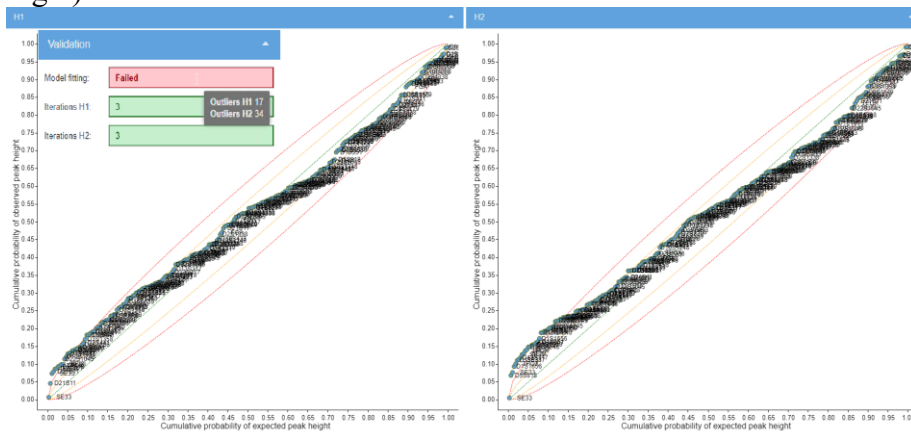
nyilvántartásba kerülő kevert vagy rossz minőségű DNS-profilok számának szignifikáns növekedése ettől az évtől kezdve. Míg az összes találatához képest a lopás ügyek találatai egymáshoz viszonyítva hasonló arányt mutatnak az egyes években, a rablások esetén a fluktuáció nagyobb. A 13. ábrán bemutatott találati statisztika alapján, ha és amennyiben lényegesen nem változik meg a közeljövőben a DNS-profil nyilvántartásba kerülő biológiai anyagmaradvány profilok mennyisége, minősége és ügytípusa, akkor az ebben a tanulmányban bemutatott új találatkereső módszer segítségével évente kb. 15-35 nyilvántartott ügyben (köztük kb. 10-20 lopásban és kb. 1-3 rablásban) várható olyan, a feltételezett nyomhagyót beazonosító egyezés, amely a hagyományos CODIS-kerestetéssel nem lenne kimutatható. Ez egyben azt is jelenti, hogy az új módszer alkalmazásával a nyilvántartásba kerülő DNS-profilokra kapott CODIS-találatok száma kb. 5-10%-kal lenne növelhető éves viszonylatban.

A 2019. és 2020. évi ügyek kevert biológiai anyagmaradványainak nem nyilvántartott és eddig be nem azonosított DNS-profiljaira kapott találatok

A 2019-ben és 2020-ban Intézetünkbe összesen beérkezett 3 142, illetve 3 181 db büntető-, körözési és közigazgatási ügy közül 706, illetve 870 db ügyben mutattunk ki kevert DNS-profilt (1. számú táblázat). Az ügyek áttekintése után 528, illetve 669 db olyan részben vagy teljes egészében ismeretlen személyi eredetű kevert profilt találtam, amelyek rossz minőségűknél és/vagy kevert jellegűknél fogva nem kerültek be a DNS-profi nyilvántartásba, viszont szakmailag és az ügyadatokból kiolvasva potenciálisan nyomozatilag relevánsak lehetnek a Familias szoftver alapú találatkeresésre. A CODIS személyi DNS-profilokkal összehasonlításban elvégezve a találatkeresést 101, illetve 167 db profilra kaptam kecsegtető biológiai anyagmaradvány/személy direkt találatot. A beazonosított feltételezett nyomhagyók 87%-a férfi, 13%-a nő, ami nagyon hasonló a nyilvántar-

tott DNS-profilokra kapott férfi/női találati arányhoz (89 és 11%). A központi nyomozó szervtől kapott visszajelzés alapján a 2019. évi ügyek 101 db találatából 48 db DNS-profilra érintő 45 db ügyben még nem történt vádemelés, elévülés vagy megszüntetés 2023 augusztusáig. Ezekben az ügyekben az egyezéseket kvantitatív LR-statisztikai vizsgálatnak is alávettem a DNAXs szoftver segítségével. Az esetek döntő többségében (41 db találat, 85%) az egyezést a DNAXs elemzés meg tudta erősíteni. Két esetben (4%) a DNAXs statisztika bizonytalan eredményt mutatott az alacsonyabb LR-érték vagy a nem teljesen megfelelő modellilleszkedés miatt (14. számú ábra), és ennek okán a nyomhagyás megbízható megerősítéséhez kiegészítő DNS-vizsgálatra lenne szükség a vizsgálati minták Intézetünkben tárolt maradékan vagy DNS-preparátumán.

Mindkét hipotézis megbukott a teszten, mivel mindkét esetben több, mint 3 lokusz (outliers) allélintenzitási valószínűségi értékei kívül estek a még elfogadható szignifikancia-területen (piros pöttyözött vonallal határolt régió).



14. számú ábra

A Familias szoftverrel kapott egyik találati eredmény kvantitatív statisztikájú DNAXs-vizsgálattal történő modellilleszkedési validációs tesztjének negatív eredménye

5 db találatot (10%) DNAXs számítással nem lehetett visszaigazolni a kapott nagyon alacsony LR-értékek miatt. A tévesnek tekinthető Familias

alapú azonosítások nem meglepő módon kizárólag a találatkeresésre kerülő lehető legbonyolultabb, 4 személytől származtatható DNS-profilokat érintették. A megerősített és laboratóriumi megerősítésre váró találatokat mutató biológiai anyagmaradványok nyomhagyóinak becsült átlagos száma 2,7; a beazonosított személyek DNS-ének becsült átlagos részaránya a keveredésekben 50% (4. számú táblázat kék színű sorai). Ezek a számok arra utalnak, hogy az általam bemutatott új találatkereső módszer bevetése leginkább legfeljebb 3 személytől származó, és az azonosítandó személy DNS-ét nem alacsony részarányú komponensként tartalmazó kevert biológiai anyagmaradványok esetében lehet nagyon hatékony.

Füredi Sándor: A magyarországi bűnügyi DNS-profil nyilvántartás találatkeresési módszerének fejlesztése

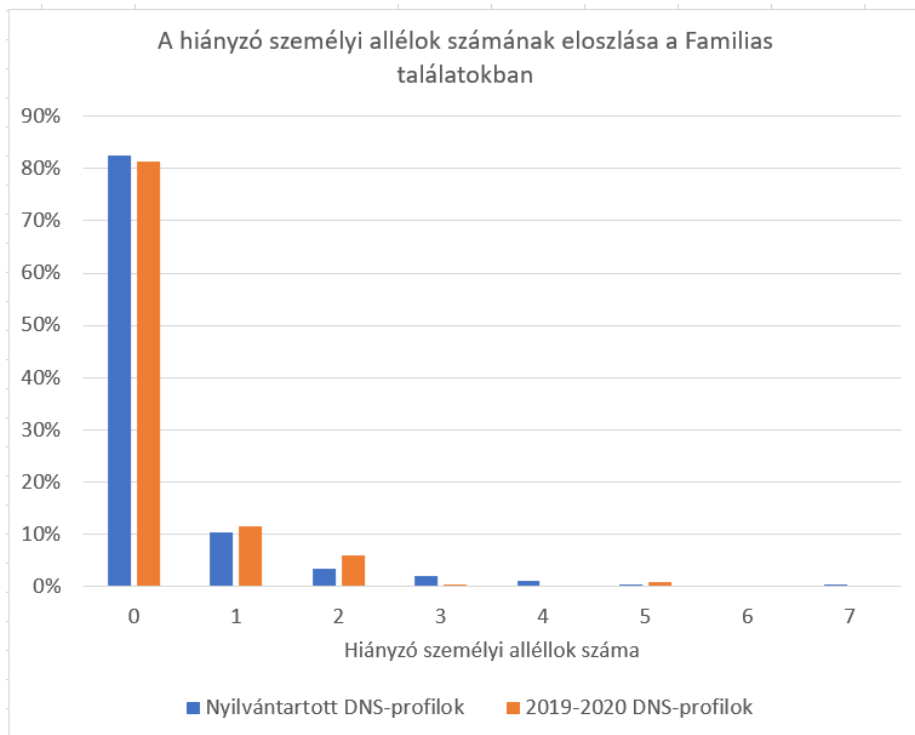
	2019 (db)	2020 (db)	Összesen (db)	Összesen (%)
Beérkezett ügyek száma	3142	3181	6323	
Kevert DNS-profilú ügyek száma	706	870	1576	24.9%
Azonosítatlan, nem nyilvántartott, kevert DNS-profilok száma	528	669	1197	18.9%
<i>Familias</i> találat	101	167	268	4.2%
<i>Familias</i> találat lezáratlan ügyekben	48			
<i>DNAxs</i> -szel megerősített találat lezáratlan ügyekben	41			85.4%
<i>DNAxs</i> -szel megerősített, de tovább vizsgálandó találat lezáratlan ügyekben	2			4.2%
<i>DNAxs</i> -szel nem visszaigazolható találat lezáratlan ügyekben	5			10.4%
A megerősített és laboratóriumi megerősítésre váró találatokat mutató biológiai anyagmaradványok nyomhagyóinak becslött átlagos száma	2.7			
Beazonosított személyek DNS-ének becslött átlagos részaránya a keveredésekben	0.5			
lopás (Familias találat)	66	96	162	60.4%
emberölés (Familias találat)	5	13	18	6.7%
kábítószer-kereskedelem (Familias találat)	7	5	12	4.5%
rablás (Familias találat)	3	9	12	4.5%
testi sértés (Familias találat)	5	6	11	4.1%
csalás (Familias találat)	4	6	10	3.7%
pénzhamisítás (Familias találat)	0	10	10	3.7%
szexuális erőszak (Familias találat)	1	4	5	1.9%
emberölés kísérlet (Familias találat)	0	4	4	1.5%
kábítószer birtoklás (Familias találat)	2	2	4	1.5%
kifosztás (Familias találat)	0	4	4	1.5%
garázdaság (Familias találat)	0	3	3	1.1%
jármű önkényes elvétele (Familias találat)	2	1	3	1.1%
lőfegyverrel vagy lőszerral visszaélés (Familias találat)	2	1	3	1.1%
rongálás (Familias találat)	2	1	3	1.1%
személyi szabadság megsértése (Familias találat)	0	2	2	0.7%
általános közigazgatási rendtartás (Familias találat)	1	0	1	0.4%
magánlaksértés (Familias találat)	1	0	1	0.4%

4. számú táblázat

A 2019. és 2020. évben Intézetünkbe beérkező büntető-, körözési és közigazgatási ügyek több személytől származó biológiai anyagmaradványainak nem nyilvántartott és eddig be nem azonosított DNS-profiljaira kapott találatok adatai

A kevert DNS-profilokban a Familias szoftverrel történő tömegszűrés-szintű találatkeresés módszertanának optimalizációja és a találati eredmények pontosabb értelmezhetősége érdekében megvizsgáltam, hogy kimutatható-e összefüggés az egyezést mutató DNS-profilok közös (átfedő) lokuszainak száma, illetve a biológiai anyagmaradvány DNS-profiljából esetlegesen hiányzó személyi allélok száma és a kapott LR-értékek között.

Elméleti megfontolások alapján minél több lokuszon minél kevesebb allélhiánnyal mutatható ki egyezés, annál nagyobb lehet az LR értéke. Természetesen ezt jelentősen befolyásolhatja az egyező alléloknek a népességben való előfordulási gyakorisága, hiszen gyakoribb allélokra alacsonyabb LR-érték várható a statisztikai számítások eredményeképpen. A Familias szoftverrel találatokat mutató biológiai anyagmaradványok DNS-profiljaiból esetlegesen hiányzó személyi allélok eloszlása nagyon hasonló a nyilvántartott, illetve a 2019. és 2020. évi ügyek nem nyilvántartott, korábban beazonosítatlan DNS-profiljai között (15. számú ábra). Mindkét sorozatban a találatok döntő többségében (83, illetve 81%) az egyezést mutató személyi DNS-profil összes allélja megtalálható volt a biológiai anyagmaradvány DNS-profiljában. Ugyanakkor az esetek nem elhanyagolható részében (17, illetve 19%) legalább egy személyi allél hiányzott az anyagmaradvány profiljából. A hiányzó allélok száma folyamatosan csökkenő gyakorisággal egytől hétig terjedt. Ha és amennyiben a CODIS-adatbázis képes is lenne ezeket a találatokat kimutatni, akkor is a detektált allélhiányok egyben a DNS-profil nyilvántartási törvények által megszabott azonosság hiányát is jelentenék a jogszabályok szigorúbb értelmezése szerint, megnehezítve ezen részleges egyezéseknek a büntetőügyek nyomozása és bizonyítása során történő hasznosítását.

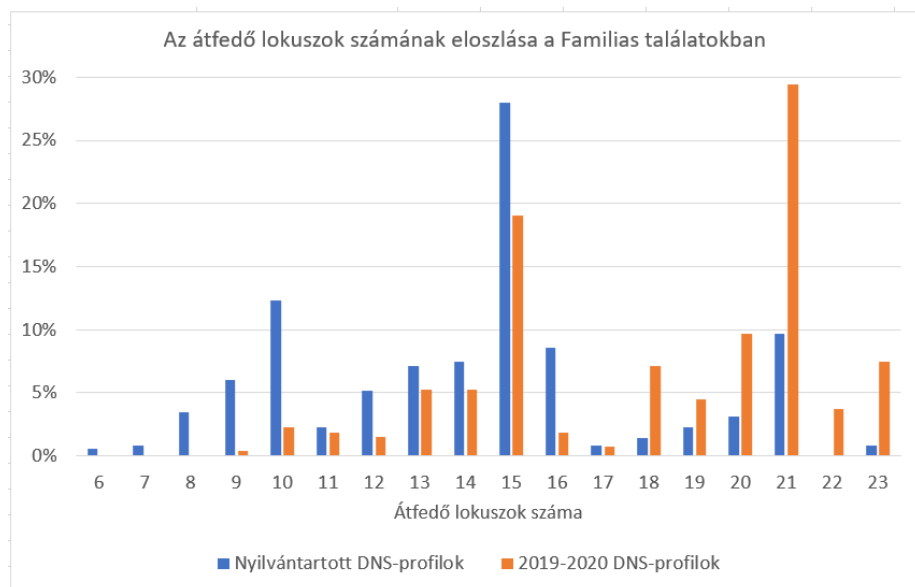


15. számú ábra

A Familias szoftverrel találatokat mutató biológiai anyagmaradványok DNS-profiljaiból esetlegesen hiányzó személyi allélok eloszlása a nyilvántartott, illetve a 2019. és 2020. évi ügyek nem nyilvántartott, korábban beazonosítatlan DNS-profiljai között

A Familias-vizsgálatban találatot adó DNS-profil párok közös (átfedő) lokuszainak száma részben eltér a nyilvántartott, illetve a 2019. és 2020. évi ügyek nem nyilvántartott, korábban beazonosítatlan DNS-profiljai között (16. számú ábra). 6-23 lokuszátfedéssel kaptam egyezéseket, melyben a nyilvántartott profilok találatainál a 6-15 lokuszos átfedések (73%-ban), míg a 2019-2020. évi ügyek profiljainál a 15-23 lokuszos átfedések (84%-ban) domináltak. Ez az eltérés egyértelműen arra vezethető vissza, hogy a 2019. év előtti ügyekben a DNS-profilok meghatározása kevesebb lokuszra terjedt ki. A 16. számú ábra adatsorában 3 kiugró lokuszátfedési érték (10,

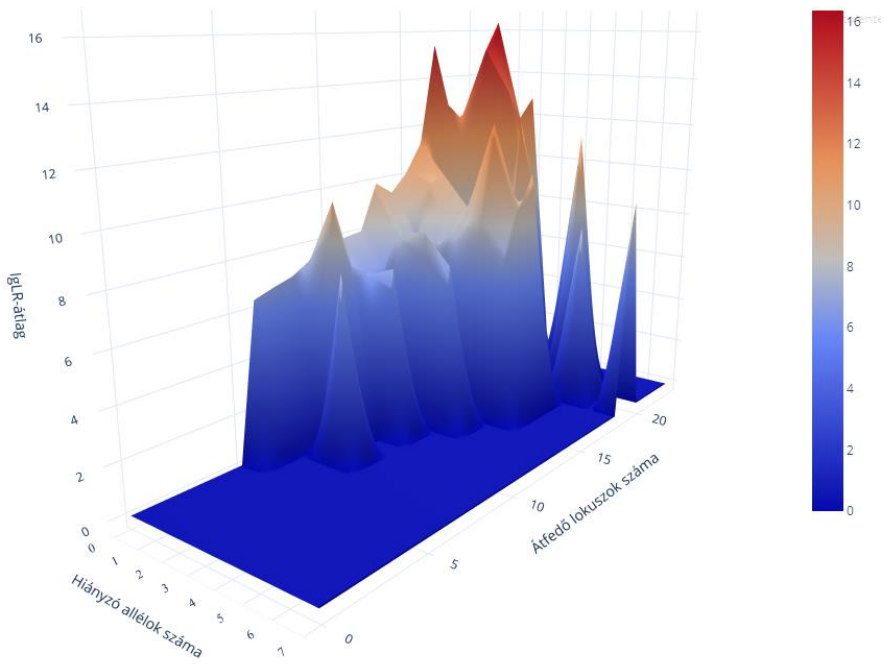
15, 21) figyelhető meg, ezek a 2004-2020. között a biológiai anyagmaradványok DNS-profil vizsgálatához legtöbbször használt, időben egymást felváltó vegyszerkitek (SGM Plus, NGM SElect, Globalfiler) által elemezhető STR-lokuszek számainak felelnek meg.



16. számú ábra

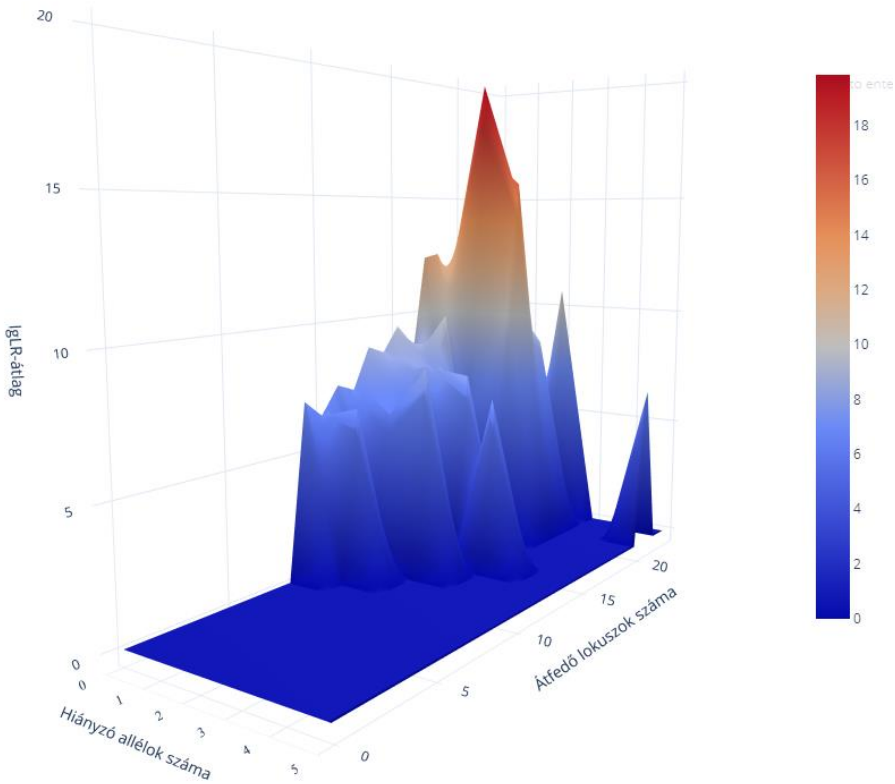
A Familias szoftverrel találatokat mutató DNS-profil párok közös (átfedő) lokuszszámának eloszlása a nyilvántartott, illetve a 2019. és 2020. évi ügyek nem nyilvántartott, korábban beazonosítatlan DNS-profiljai között

A Familias szoftverrel egyezést mutató DNS-profilok közös (átfedő) lokuszainak száma, illetve a biológiai anyagmaradvány DNS-profiljából esetlegesen hiányzó személyi allélok száma és a kapott LR-értékek közötti összefüggés vizsgálata során, ha egy átfedő lokuszszám/hiányzó személyi allélszám párhoz több LR-érték társult, akkor ahhoz a párhoz a hozzájuk tartozó LR-értékek tízes alapú logaritmusának átlagát rendeltem, mint arra a párra jellemző LR-számot. A logaritmizálásnak az az oka, hogy a nagyobb LR-értékek esetében azok nagyságrendje számít igazából a személyi származtatás megbízhatóságának statisztikai számításon alapuló szakértői értékelésénél. Az LR-értékek nagyságrendjét pedig azok logaritmusai fejezik ki. A 17. és 18. számú ábra 3D grafikonjai egymáshoz hasonló trendeket mutatnak a nyilvántartott és a 2019–2020. évi nem nyilvántartott, korábban beazonosítatlan kevert DNS-profilok Familias találataihoz tartozó lgLR-átlagok eloszlásában. A biológiai anyagmaradványok kevert DNS-profiljaiból hiányzó személyi allélok számának csökkenésével és a találatot mutató DNS-profilokban a közös (átfedő) lokuszok számának növekedésével a lgLR-átlagok – néhány kilengéstől eltekintve – növekszenek. Ezen felül minél több az átfedő marker, annál nagyobb az esélye annak, hogy több személyi allél hiányában is meghaladja a vizsgált DNS-profil párra kapott valószínűségi hányados a felállított egymilliós LR-küszöbértéket. A vizsgálat sorozatomban megfigyelt találatok lgLR-értékek átlagai tehát megfeleltek az elméletileg vártaknak, ami egyben azt is jelenti, hogy a DNS-vizsgálatok során az analizált lokuszok számának emelése előnyös a valós találatok megbízható azonosításában, ugyanakkor ezekben az esetekben a személyi allélok hiánya is gyakrabban és nagyobb mértékben fordulhat elő.



17. számú ábra

Nyilvántartott, korábban beazonosítatlan kevert DNS-profilok Familias találatainak lgLR-átlagai az átfedő lokuszok és a hiányzó allélok számának függvényében



18. számú ábra

2019–2020. évi nem nyilvántartott, beazonosítatlan kevert DNS-profilok Familias találatainak lgLR-átlagai az átfedő lokuszok és a hiányzó allélok számának függvényében

A Familias szoftverrel a 2019. és 2020. évi ügyekben kapott találatokat érintő ügytípusok eloszlása hasonló összképet mutatott a nyilvántartott DNS-profilokra korábban megállapított egyezések ügytípusaihoz képest (4. számú táblázat zöld sorai és 3. számú táblázat). A lopás ügyek itt is hasonló arányban (60%) domináltak, és a 2% vagy afeletti részarányú ügytípusok között ebben a vizsgálatsorozatban is megtalálható volt az emberölés, a kábítószerrel elkövetett bűncselekmény, a rablás, a testi sértés és a

szexuális bűncselekmény. Ezek az eredmények azt mutatják, hogy a nyilvántartásba be nem kerülő beazonosítatlan kevert DNS-profilok esetében is a lopás ügyekben lehetne a legtöbb személyazonosítást elérni az új találatkereső módszer segítségével, és mindemellett évente várhatóan 20–30-cal több, súlyosabb bűncselekményben (köztük emberölésben) tudnánk beazonosítani a vétlen vagy elkövetői ismeretlen nyomhagyót a módszer CODIS-kerestetést kiegészítő alkalmazásával. A 2020. és 2021. évek ügyeinek találati adatait és a Familias-alapú találatkeresésnek a 2019. évi ügyekben kapott 10%-os potenciális tévesztését (mint II. típusú, másodfajú hibát) alapul véve évente átlagosan kb. 110-120 olyan ügyben tudnánk sikeres személyazonosítást végrehajtani a nem nyilvántartott többszemélyi eredetű DNS-profilokra vonatkozóan, amely profilokon a jelenlegi személyazonosító módszereink nem vezetnének eredményre. A 110-120 db ügy az Intézetünkbe beérkező, összesen 3 000 körüli büntető-, körözési és közigazgatási ügyhöz képest nem tűnik soknak. Ugyanakkor, ha a 110-120 db ügyhöz hozzáadjuk azt az évi mintegy 15-35 db azonosítást, amit a nyilvántartott DNS-profilok várható találataira becsültünk meg korábban, és figyelembe vesszük azt, hogy évente jelenleg átlagosan 400 körüli egyezést mutat ki a CODIS-adatbázis, akkor a 125-155 darabos találati növekmény 31-39%-os emelkedést jelent az ismeretlen biológiai anyagmaradványok DNS-profil nyilvántartás segítségével történő személyazonosításában. Ez véleményem szerint számottevő fejlesztési eredmény a magyarországi bűnügyi DNS-szakértői vizsgálatok területén.

További feladatok, tervek és célok

Az igazságügyi genetikai vizsgálatok technológiájának fejlődése folyamatos szakmai, jogi és etikai kihívások elé állítja az igazságszolgáltatás szereplőit (Nogel, 2019; Pádár et al., 2020; Smith & Singh, 2024), amely aspektusokat a DNS-vizsgálatok eredményeinek interpretációja, a találatkeresés és azok fejlesztése során is figyelembe kell venni.

A nem nyilvántartott ismeretlen személyazonosságú kevert DNS-profilok közül eddig a 2019. és 2020. évben Intézetünkbe beérkezett ügyekre vonatkozóan tudtam a Familias szoftver alapú találatkeresést elvégezni. A találatkeresést a 2020 után beérkezett ügyekben is folytatni szeretném. Egy adott év DNS-profiljait utólag a következő év második felétől érdemes vizsgálni, mert akkora már nagy valószínűséggel befejeződtek a szakértői munkák az előző év ügyeiben. A központi nyomozóhatósággal való együttműködésben a még aktuális ügyekben el szeretném végezni a találatok kvantitatív modellű statisztikai számításra alapuló megerősítését a DNAXs szoftver segítségével.

A tanulmányomban bemutatott LR-alapú találatkeresési módszernek a folyamatban lévő ügyekben történő, rutinszerű végrehajthatósága szempontjából az egyik legfontosabb feladat a DNAXs és a CODIS szoftver közötti adatkapcsolat létrehozása. Ezt a DNAXs-t kifejlesztő holland Nemzeti Forenzikus Intézet (NFI) olyan informatikai környezetben valósította meg, ahol a DNAXs szakértői ügyprojektjeinek adatbázisa, a statisztikai számításokat végző szoftvermodulok, valamint a találatkereső (SmartRank vagy ProbeRank) modul (Benschop et al., 2017; Hoogenboom, Sijen & Benschop, 2023) szerveren, szervercsoporton (cluster) vagy tárolókon (containers) fut. Sajnos Intézetünk jelenlegi szerverparkja nem teszi lehetővé a számításokat végző DNAXs szoftvermodul (DNAStatistX) szervercsoporton történő futtatását, illetve egyetlen szerver nem tudná egyszerre kiszolgálni az összes aktuális ügyünkben keletkező számítási igényeket. Ez jelenleg csakis a szakértők személyi számítógépein, önálló (standalone, ti. nem szerver) DNAXs üzemmódban valósítható meg. Ugyanakkor kérdéssemre az NFI specialistái lehetségesnek tartják kizárólagosan a kereső (SmartRank vagy ProbeRank) modul szerver üzemmódba állítását, és így a CODIS-ba történő központosított bekérdezés lehetőségét. Ennek a lehetőségnek a megvalósítása az Intézetünkben informatikai fejlesztést igényelne, amelynek az esetleges hardver, szoftver és humán erőforrás szükséglete még meghatározásra vár. A fejlesztés megvalósulása esetén a nyil-

vántartási szakértők a CODIS-ban képesek lennének a DNAXs program segítségével a rossz minőségű és/vagy kevert DNS-profilokon történő egyszeri találatkeresésre a szakértői vélemények kiadása előtt. Jelenleg is hasonlóan történik ez az egyszeri kerestetés az egyszemélyi DNS-profilokkal a LIMS-ünk és a CODIS között.

További szakmai feladatot jelent a nyomhordozókon lévő háttérszennyeződés és a rendszeres használók DNS-ének alaposabb feltérképezése annak érdekében, hogy a nyilvántartásban történő találatkereséssel beazonosított személy nyomhagyásának módja és így annak a bűncselekményben betöltött – véltlen vagy elkövetői – cselekvési szerepe pontosabban meghatározható legyen (Gill et al., 2020; Reither et al., 2023). Ugyanez érvényes a biológiai anyagmaradványok közvetítő személyen vagy tárgyon keresztül történő indirekt átkerülésére is (Sessa et al., 2023).

A jogszabályi háttérrel illetően indokolt és szükséges lenne a magyarországi nyilvántartási törvényeknek olyan szövegmódosítása, amely a találatok kimutatásának és az érintett hatóságokkal való közlésének szakmai feltételét a jelenlegi „azonosság” szó helyett például a „teljes vagy részleges egyezés” kifejezéssel szabná meg a törvények szövegében. Ez egyben szakmai harmonizációt is jelentene az ENFSI által az ENFSI-tagországoknak javasolt ajánlásával (URL2), amely egyezésként (angolul „*match*”), és nem azonosságként (angolul „*identity*”) definiálja a nyilvántartási találatot.

Köszönetnyilvánítás

Szeretném megköszönni Kormos Zoltánnak, az NSZKK Genetikai Szakértői Intézet (GSZI) DNS-Nyilvántartó Osztály vezetőjének a nyilvántartási adatok rendelkezésemre bocsátását. Köszönettel tartozom Horváth Zoltánnak, a GSZI informatikusának a LIMS-ünk, a Familias és a DNAXs szoftver közötti adatkapcsolat megteremtéséért, illetve a találatokat érintő ügyadatoknak a LIMS-ünkből történő lekérdezéséért. Az ORFK részéről hálás vagyok dr. Petőfi Attila r. vezérőrnagy úrnak és Kompoltiné Berényi Anita

asszonynak a találatot mutató ügyek státusának megállapításáért. Végül, de nem utolsósorban szeretném megköszönni Lontai Mártonnak, az NSZKK főigazgatójának és Dobos Ágotának, a GSZI volt igazgatójának a folyamatos támogatását ebben a találatkereső projektben is.

Felhasznált irodalom

Akane, A., Shiono, H., Matsubara, K., Nakahori, Y., Seki, S., Nagafuchi, S., Yamada, M., Nakagome, Y.: Sex identification of forensic specimens by polymerase chain reaction (PCR): two alternative methods. *Forensic Sci Int* 1991 /1. 81–8. o.

Forrás: [https://doi.org/10.1016/0379-0738\(91\)90174-H](https://doi.org/10.1016/0379-0738(91)90174-H)

Balding, DJ., Nichols, RA.: DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic Sci Int* 1994/2-3. 125–40. o.

Forrás: [https://doi.org/10.1016/0379-0738\(94\)90222-4](https://doi.org/10.1016/0379-0738(94)90222-4)

Bekaert, B., Van Geystelen, A., Vanderheyden, N., Larmuseau, MH., Decorte, R.: Automating a combined composite-consensus method to generate DNA profiles from low and high template mixture samples. *Forensic Sci Int Genet*, 2012/5. 588–593. o.

Forrás: <https://doi.org/10.1016/j.fsigen.2012.02.001>

Benschop, CC., Sijen, T. (). LoCIM-tool: An expert's assistant for inferring the major contributor's alleles in mixed consensus DNA profiles. *Forensic Sci Int Genet*. 2014. 154–165. o.

Forrás: <https://doi.org/10.1016/j.fsigen.2014.03.012>

Benschop, CCG., van de Merwe, L., de Jong, J., Vanvooren, V., Kempenaers, M., van der Beek, CPK., Barni, F., Reyes, EL., Moulin, L., Pene, L., Haned, H., Sijen, T.: Validation of SmartRank: A likelihood ratio software for searching national DNA databases with complex DNA profiles, *Forensic Science International: Genetics*. 2017. 145–153. o.

Forrás: <https://doi.org/10.1016/j.fsigen.2017.04.008>

Benschop, CCG., Hoogenboom, J., Hovers, P., Slagter, M., Kruise, D., Parag, R., Steensma, K., Slooten, K., Nagel, JHA., Dieltjes, P., van Marion, V., van Paassen, H., de Jong, J., Creeten, C., Sijen, T., Kneppers, ALJ.: DNAXs/DNAStatistX: Development and validation of a software suite for the data management and probabilistic interpretation of DNA profiles. *Forensic Sci Int Genet.* 2019, 81–89. o.

Forrás: <https://doi.org/10.1016/j.fsigen.2019.06.015>

Benschop, C. (2022). Likelihood Ratio calculations using EuroForMix or DNAStatistX. CEPOL kurzus előadás. Riga. 2022. október.

Bleka, Ø., Benschop, CCG., Storvik, G., Gill, P.: A comparative study of qualitative and quantitative models used to interpret complex STR DNA profiles. *Forensic Sci Int Genet.* 2016. 85–96. o.

Forrás: <https://doi.org/10.1016/j.fsigen.2016.07.016>

Butler, JM.: Differential Extraction. In: Butler, JM.: *Advanced Topic in Forensic DNA Typing: Methodology.* Elsevier. 2011. 37–40. o.

Forrás: <https://doi.org/10.1016/C2011-0-04189-3>

Clayton, TM., Guest, JL., Urquhart, AJ., Gill, PD.: A genetic basis for anomalous band patterns encountered during DNA STR profiling. *J Forensic Sci* 2004/6. 1207–1214. PMID: 15568691.

Gill, P., Hicks, T., Butler, JM., Connolly, E., Gusmão, L., Kokshoorn, B., Morling, N., van Oorschot, RAH., Parson, W., Prinz, M., Schneider, PM., Sijen, T., Taylor, D.: DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence - Guidelines highlighting the importance of propositions. Part II: Evaluation of biological traces considering activity level propositions. *Forensic Sci Int Genet.* 2020. 102186.

Forrás: <https://doi.org/10.1016/j.fsigen.2019.102186>

Foreman, LA., Smith, AFM., Evett, IW.: Bayesian Analysis of DNA Profiling Data in Forensic Identification Applications. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 1997/3, 429–469. o.

Forrás: <http://www.jstor.org/stable/2983264>

Füredi, S.: Az igazságügyi genetikai vizsgálatok statisztikai interpretációja. In: Tóth, É., Belovics, E. (szerk.): A büntetőeljárás segédtudományai I. Pázmány Press. Budapest, 2015. 363–384. o.

Füredi, S.: Kvantitatív statisztikai modellű személyazonosító szoftverek verifikálása a magyarországi bűnügyi DNS-vizsgálatokhoz. Poszter prezentáció a XXII. „Genetikai Műhelyek Magyarországon” c. minikonferencián. Szeged, 2023. szeptember 15.

Hoogenboom, J., Sijen, T., Benschop, C.: ProbRank: An efficient DNA database search method for complex mixtures per a quantitative likelihood ratio model. *Forensic Sci Int Genet.* 2023. 102884.

Forrás: [https://doi.org/10.1016/0379-0738\(91\)90174-h](https://doi.org/10.1016/0379-0738(91)90174-h)

Kling, D., Füredi, S.: The successful use of familial searching in six Hungarian high profile cases by applying a new module in Familias 3. *Forensic Sci Int Genet.* 2016. 24–32. o.

Forrás: <https://doi.org/10.1016/j.fsigen.2016.05.012>

Nogel, M. (2019) A bűnügyi célú DNS-vizsgálatok szabályozása Magyarországon: múlt, jelen és jövő. *Magyar Jog.* 2019.11.29. 689–698. o.

Pádár, Zs., Kovács, G., Nogel, M., Czebe, A., Zenke, P., Kozma, Zs.: Genetika és bűnüldözés – Az igazságügyi célú DNS-vizsgálatok első negyedszázada Magyarországon I. *Belügyi Szemle* 2019/12. 7–34. o.

Forrás: <https://doi.org/10.38146/BSZ.2019.12.1>

Pádár, Zs., Kovács, G., Nogel, M., Czebe, A., Zenke, P., Kozma, Zs.: Genetika és bűnüldözés – Az igazságügyi célú DNS-vizsgálatok első negyedszázada Magyarországon II. *Belügyi Szemle* 2020/1. 9–32. o.

Forrás: <https://doi.org/10.38146/BSZ.2020.1.1>

Reither, JB., Taylor, D., Szkuta, B., van Oorschot RAH.: Exploring how the LR of a POI in a target sample is impacted by awareness of the profile of the background derived from an area adjacent to the target sample. *Forensic Sci Int Genet.* 2023. 102868.

Forrás: <https://doi.org/10.1016/j.fsigen.2023.102868>

Saiki, RK., Scharf, S., Faloona, F., Mullis, KB., Horn, GT., Erlich, HA., Arnheim, N.: Enzymatic Amplification of β -Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science*. 1985. 1350-1354. o.

Forrás: <https://www.science.org/doi/10.1126/science.2999980>

Sessa, F., Pomara, C., Esposito, M., Grassi, P., Cocimano, G., Salerno, M.: Indirect DNA Transfer and Forensic Implications: A Literature Review. *Genes (Basel)* 2023/12. 2153. o.

Forrás: <https://doi.org/10.3390/genes14122153>

Smith, J., Singh, M.: Forensic DNA Profiling: Legal and Ethical Considerations. *Journal of Scientific Research and Reports*. 2024. 141–144. o.

Forrás: <https://doi.org/10.9734/jsrr/2024/v30i51929>

Taylor, D., Bright, J-A., Scandrett, L., Abarno, D., Lee, S-I., Wivell, R., Kelly, H., Buckleton, J.: Validation of a top-down DNA profile analysis for database searching using a fully continuous probabilistic genotyping model, *Forensic Science International: Genetics*. 2021. 102479.

Forrás: <https://doi.org/10.1016/j.fsigen.2021.102479>

Interneten elérhető hivatkozások

URL1: Frequently Asked Questions on CODIS and NDIS

Forrás: <https://www.fbi.gov/how-we-can-help-you/dna-fingerprint-act-of-2005-expungement-policy/codis-and-ndis-fact-sheet>

URL2: ENFSI Board (2023). Guideline for DNA Database Management. Review and Recommendations

Forrás: <https://enfsi.eu/wp-content/uploads/2023/10/ENFSI-GUIDELINE-FOR-DNA-DATABASE-MANAGEMENT-REVIEW-AND-RECOMMENDATIONS.pdf>

Letöltés ideje: 2024.06.18.

Füredi Sándor: A magyarországi bűnügyi DNS-profil nyilvántartás találatkeresési módszerének fejlesztése

URL3: Familias.no

Forrás: <https://familias.no/>

Letöltés ideje: 2024.06.18.

URL4: Best Practice Manual for the internal validation of probabilistic software to undertake DNA mixture interpretation

Forrás: <https://enfsi.eu/wp-content/uploads/2024/04/Best-Practice-Manual-for-the-internal-validation-of-probabilistic-software-to-undertake-DNA-mixture-interpretation-v1.docx-1-1.pdf>

Letöltés ideje: 2024.06.18.

URL5: STRidER. STRs for identity ENFSI Reference database, v3/R2. Frequencies

<https://strider.online/frequencies>

Letöltés ideje: 2024.06.18.

URL6: Füredi, S.: Humán polimorf mikroszatellita (short tandem repeat) lokuszok igazságügyi genetikai vizsgálata magyar populációkban. Doktori értekezés (2003)

Forrás: <https://docplayer.hu/3560577-Human-polimorf-mikroszatellita-short-tandem-repeat-lokuszok-igazsagugyi-genetikai-vizgalata-magyar-populaciokban.html>

Letöltés ideje: 2024.06.18.

URL7: Harvey, A., Howell, E. (2022). How many stars are in the universe?

Forrás: <https://www.space.com/26078-how-many-stars-are-there.html>

Letöltés ideje: 2024.06.18.

Felhasznált jogszabályok

2009. évi XLVII. törvény a bűnügyi nyilvántartási rendszerről, az Európai Unió tagállamainak bíróságai által magyar állampolgárokkal szemben

hozott ítéletek nyilvántartásáról, valamint a bűnügyi és rendészeti biometrikus adatok nyilvántartásáról

2013. évi LXXXVIII. törvény a körözési nyilvántartási rendszerről és a személyek, dolgok felkutatásáról és azonosításáról

Függelék

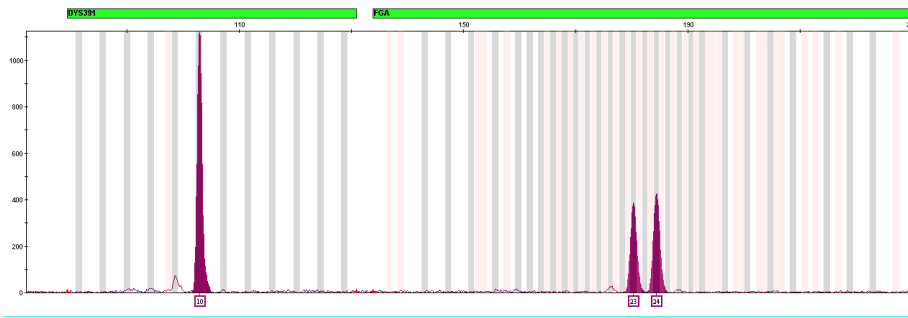
Az igazságügyi DNS-profil vizsgálatok alapjai

Az igazságügyi genetikai vizsgálatok fő célja a bűncselekményekkel kapcsolatba hozható biológiai anyagmaradványok, ismeretlen személyazonosságú élő vagy holt személyek, illetve a genetikai rokonok azonosítása. Mindezen vizsgálatok jelenleg az úgynevezett Short Tandem Repeat genetikai markerek (lokuszok) (STR-ek, rövid tandem ismétlődések, mikroszatelliták) fluoreszcens PCR DNS-fragmensanalízisen alapulnak. Az igazságügyi DNS-elemzések során az élőlénynek (leggyakrabban az embernek) nem a teljes örökítő anyagát (genomját) vizsgáljuk, hanem – vizsgálati-típustól függően – általában csak kb. 20-30 db kitüntetett helyét, amelyeket összefoglalóan markereknek vagy lokuszoknak hívunk. A markerek egy csoportját képezik az STR-lokuszek, amelyekben a 2–7 nukleotidból¹ álló rövid DNS-szakaszok egymás után, tandem módon többször (általában 5–50-szer) ismétlődnek. Az ismétlődések száma sok STR-markernél a magasabbrendű élőlények egyedei, így az emberek között is változatosságot (polimorfizmust) mutat, amelynek köszönhetően a polimorf STR-ek lokuszonként – a markerek típusától függően – általában 5–100 féle típusban (általában) fejeződnek ki az emberi népességben. Az STR-allélokot azzal a számmal szokták jelölni, ahány db ismétlődő szakaszt tartalmaznak. Az STR-markereket általában a kromoszómájukon elfoglalt egykópiás lokuszorszámuk szerint nevezik el (például D21S11: a 21. kromoszóma 11. egykópiás lokusza), vagy arról a génről kapják a nevüket, amelyikben vagy amelyik közelében helyezkednek el (például vWA: von Willebrand Factor A típus). A polimorf STR-lokuszek az összes emberi kromoszómán megtalálhatók, így a testi (autoszómás) és a nemi (X- és Y-) kromoszómákon. Tekintettel arra, hogy az emberek a testi kromoszómájukat és a nők az X-kromoszómájukat mindkét szülőtől kapják, ezért – néhány ritka kivételtől

¹ nukleotid: a DNS és RNS örökítőanyagot alkotó nukleotidlánc szerkezeti egysége

eltekintve – minden ember az autoszómás, és minden nő az X-kromoszómás STR-markerein két-két alléllal rendelkezik (szakszóval diploid allélok). A férfiak a szüleiktől csak egy X- és egy Y-kromoszómát örökölnek, így lokuszonként csak egy X- vagy egy Y-STR allélt hordoznak (haploid allélok).

Az emberi sejtek DNS-tartalma tömegét tekintve rendkívül alacsony, kb. 6 pg (6×10^{-12} gramm). Ennek folytán a genetikai állomány bűnügyi mintákon hatékonyan, közvetlenül nem analizálható. Az 1985-ben felfedezett, később Nobel-díjjal jutalmazott polimeráz láncreakció (PCR) (Saiki et al., 1985) alapuló sokszorozási (amplifikációs) eljárással azonban a vizsgálni kívánt DNS-szakaszok mennyisége több milliószorosára növelhető, így az elemzésre kiválasztott markerek alléljai már egyszerűbb detektálási eljárással is láthatóvá tehetők. A PCR-technika másik óriási előnye, hogy faj- (például humán-) és markerspecifikus vizsgálatot tesz lehetővé, azaz az elemzést nem zavarja a többi DNS-szakasz és esetleg más fajoktól (például mikrobáktól) származó genetikai anyag jelenléte. A fluoreszcens PCR-amplifikáció során az STR-allélok színjelölést kapnak egy-egy rájuk helyezett fluoreszkáló festék révén, melyek lézersugárral megvilágítva az adott festékre jellemző hullámhosszúságú fényt bocsátanak ki. A különböző méretű (hosszúságú) allélok kapilláriscsőben lévő gélbe vezetett áram segítségével, elektroforézissel szétválaszthatók, mivel a kisebb allélok az elektromos térben gyorsabban mozognak a nagyobb társaikhoz képest. A fix helyzetű lézersugáron áthaladó allél fluoreszcens fényt bocsát ki, amely a genetikai analizáló műszer görbéjén (az elektroferogramon) egy csúcsként jelenik meg. Az allélcsúcs magassága egyenesen arányos a mintában lévő azon eredeti (úgynevezett templát) DNS mennyiségével, amelyről a detektált allél felsokszorozva lemásolódott. Ha a személy a vizsgált diploid lokuszon a szüleitől ugyanolyan típusú allélt kapott, akkor az adott marker-nél csak egy darab dupla magasságú csúcs figyelhető meg (homozigóta genotípus), míg, ha az egyén két különböző méretű allélt örökölt, akkor két darab, kb. fele magasságú csúcs detektálható (heterozigóta genotípus) (F1. számú ábra).

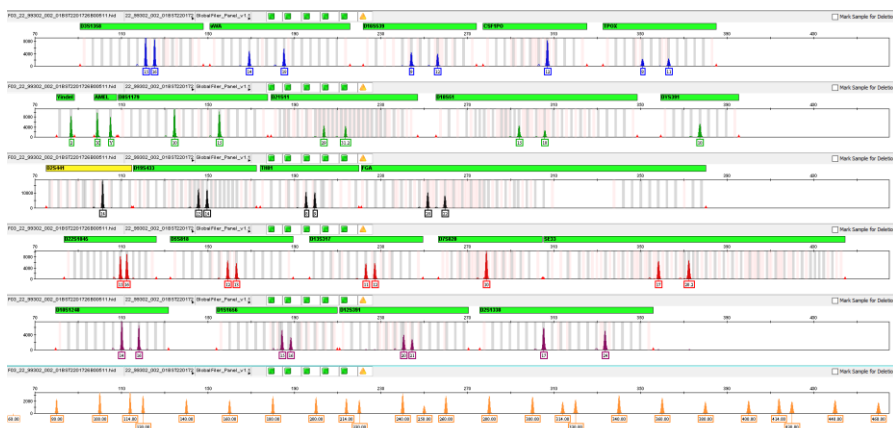


F1. számú ábra

Homozigóta (bal oldal) és heterozigóta (jobb oldal) STR-lokuszt elektroferogramja

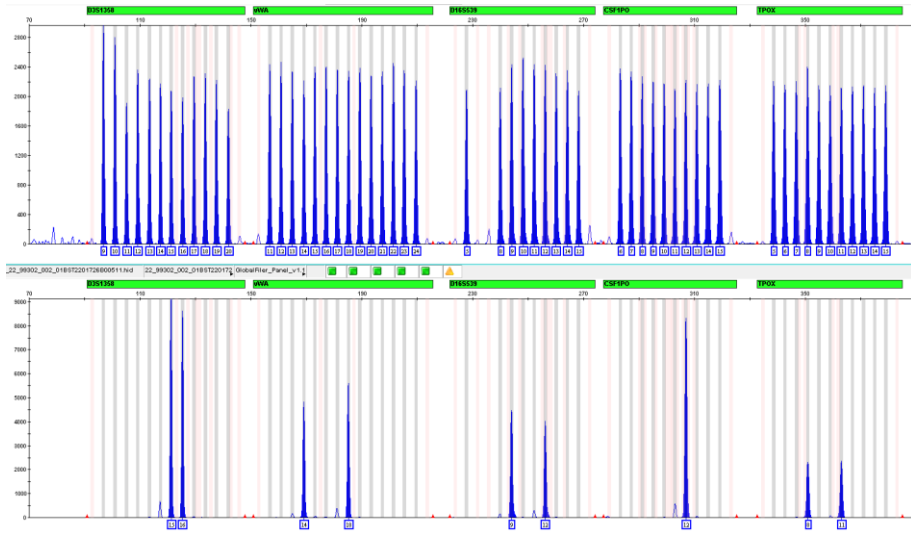
Az STR-lokusztok hagyományos PCR-rel vizsgálható mérete általában 80–500 nukleotid között mozog. Ahhoz, hogy egyetlen vizsgálattal 5 db STR-nél többet lehessen egyszerre elemezni, az szükséges, hogy az egymással méretükben átfedő markerek különböző fluoreszcens festékekkel legyenek megjelölve. Hatféle festék használatával 20–30 db lokuszt már egyszerre analizálható válik (F2. számú ábra).

Az első öt sor az ötféle fluoreszcens festékekkel jelölt lokusztok alléljait mutatja, a hatodik sor pedig az úgynevezett belső méretstandard csúcsait jelzi.



F2. számú ábra
Egy személy DNS-profiljának elektroferogramja

A DNS-fragmensanalízis során a vizsgálati minták tipizálása (ti. DNS-profiljuk meghatározása) egy úgynevezett belső méretstandard és egy úgynevezett külső típusstandard (alléllétra) segítségével történik. A külön fluoreszcens színnel jelölt, ismert méretű DNS-fragmensekből álló belső standardot (lásd F2. számú ábra 6. sora) hozzákeverik a minta DNS-éhez, mely az elektroforézis során együtt futva a minta DNS-ével kalibrációs méretgörbét szolgáltat a minta alléljainak nukleotid (bázis) egységben történő méretmeghatározásához. A külső típusstandard (alléllétra) a vizsgált lokuszokon a népességben leggyakrabban előforduló, ismert típusú allélok köztéje. A belső méretstandarddal összekeverve, de a mintáktól külön kerül vizsgálatra. Az ismeretlen DNS-profilú minta alléljait aszerint tipizálják, hogy azok az alléllétra mely csúcaival kerülnek azonos pozícióba (F3. számú ábra). A tipizálás automatizálásához már természetesen erre a célra kifejlesztett szoftverek állnak rendelkezésre.



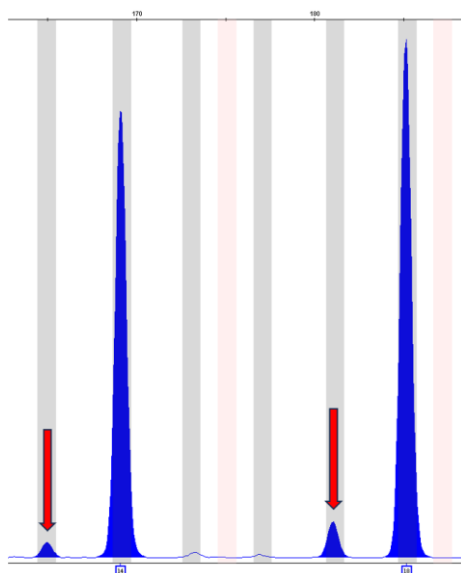
F3. számú ábra

Egy alléltre (felső sor) és egy személy DNS-profiljának (alsó sor) elektroferogramja

Egy lokuszon kimutatott alléltípusok összessége a genotípus, a személyi mintából vagy a biológiai anyagmaradványból megállapított genotípusok összessége pedig a DNS-profil. A DNS-profil, attól függően, hogy testi vagy nemi kromoszómás STR-ek alléljait tartalmazza, lehet autoszómás, X- vagy Y-kromoszómás. Tekintettel arra, hogy az Y-kromoszóma kizárólag apáról fiúra öröklődik, így az Y-kromoszómás DNS-profil mindig csak egyetlenegy szülő, az apa alléljait hordozza, ezért az Y-STR profilt Y-haplotípusnak is hívják. A DNS-profil vizsgálatok standardizálása, a költség-hatékony és magas minőségű laboratóriumi analízis, valamint az eredmények nemzetközi szintű összehasonlíthatósága érdekében gyártói vegyszer-csomagok (kitek) kerültek kifejlesztésre és a hazai szakértői vizsgálatokban is bevezetésre az STR-elemzések céljából. A gyártók által a kiteknek adott fantázianevekkel (például Globalfiler, PowerPlex Fusion 6C) az ügyszakértői munka során, a nemzeti és nemzetközi adatcserében vagy a szakirodalomban egyszerűen definiálható és meghivatkozható azoknak a DNS-

profiloknak a lokuszkészlete, amely profilok az adott vegyszercsomag segítségével lettek meghatározva.

Az STR-elektroferogram általában nemcsak allélcsúcsokból áll. Az STR PCR-amplifikáció inherens sajátossága az úgynevezett dadogó (stutter) műtermékek képződése. Ezek a PCR során döntően a templát, ritkábban a képződő DNS-szálak leggyakrabban egy, ritkábban több ismétlődő egységnyi kihurkolódásával keletkeznek, és az elektroferogramon legintenzívebben a szülőalléljuknál egy egységnyivel kisebb méretben, a szülőallél előtt, annál alacsonyabb csúcsként helyezkednek el (F4. számú ábra).



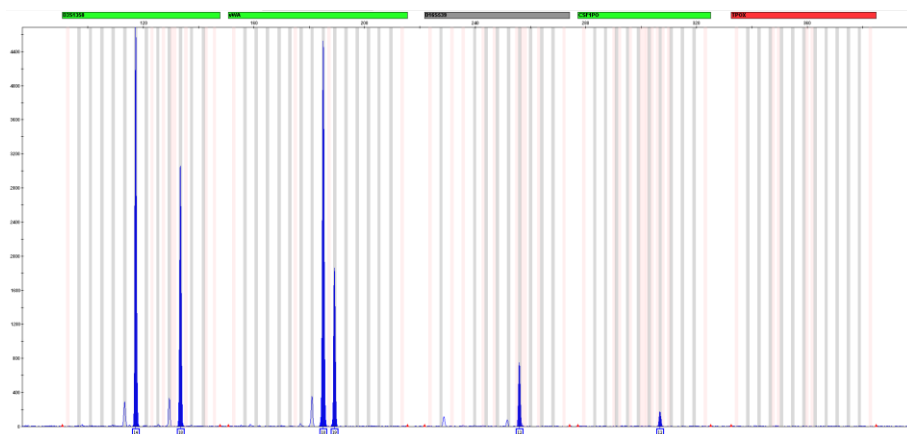
F4. számú ábra

STR-allélok hátsó (backward, N-1) stutter csúcsai (piros nyilakkal jelölve)

A DNS-profilok egyezése vagy nagyfokú hasonlósága/átfedése alátámasztja a biológiai minták (például a személyi referencia minta és a helyszíni biológiai anyagmaradvány) közös személyi eredetét, míg a profilok között megfigyelt eltérés kizárhatja azt. A rendkívül kis mennyiségű, bomlott és/vagy többszemélyi eredetű biológiai anyagmaradványokból nem

mindig lehet pontosan kimutatni a nyomhagyó személy(ek) DNS-profilját. Ha a kiindulási DNS mennyisége vagy bomlottsági foka egy kritikus tartományba kerül, akkor a PCR-során egyes allélok véletlenszerűen nem szorozódnak fel a már éppen detektálható szintig. Ezt a jelenséget allélkiesésnek (allelic dropout) hívjuk. Sok esetben detektálható mértékű amplifikáció ugyan történik, de csak csökkent mennyiségben, így a vártnál alacsonyabb allélcúcsok keletkeznek. Ez az elektroferogramon kialakíthatja a bomlott DNS úgynevezett degradációs DNS-profil képét, ahol is a nagyobb méretű lokuszok a kisebbekhez képest tipikusan alacsonyabb allélcúcsintenzitással vagy allélkieséssel mutatkoznak (F5. számú ábra).

Az utolsó lokuszon nem volt kimutatható allél a DNS bomlott állapota miatt.



F5. számú ábra

Bomlott holttestminta degradált DNS-profilú elektroferogramja

Tekintettel arra, hogy a jelenleg alkalmazott módszereinkkel a DNS-profil kimutathatóságának optimális DNS-templát mennyiségi tartománya 0,1-1 ng ($0,1-1 \times 10^{-9}$ gramm), ezért a többszemélyi eredetű (úgynevezett kevert) minták esetében annak a nyomhagyó személynek a genetikai profilja sokszor kimutathatatlanává válik, akinek a DNS-e nem éri el a mintában

lévő összes DNS egytizednyi mennyiségét. Ez a kimutathatatlanság az alacsony DNS-tartalmú vagy bomlott egyszemélyi mintákéhoz hasonlóan részleges is lehet, azaz az úgynevezett minor nyomhagyónak nem mindegyik allélja tűnhet el. Az alacsony DNS-tartalmú biológiai anyagmaradványok esetében még egy jelenség, az úgynevezett allélbeesés (allelic drop-in) befolyásolhatja a mintából meghatározható DNS-profilt. Háttér- vagy felülszennyeződés révén a vizsgálni kívánt maradványba minimális mértékben bele- vagy rákerülhetnek indifferens személyek alléljai is, mely allélok kis számuk, véletlenszerű megjelenésük és gyenge intenzitásuk miatt személyazonosításra alkalmatlanok. A nagyobb mértékű felülszennyezés már személyazonosítható minőségű DNS-profil eredményezhet, amely jelenséget az allélbeeséstől megkülönböztetve kontaminációnak hívunk.

A magyarországi igazságügyi DNS-profil nyilvántartások

Öt évvel a világ első bűnügyi DNS-profil nyilvántartásának, a brit DNS-adatbázisnak a felállítása után, 2000-ben Magyarországon is hatályba lépett a kriminalisztikai célú DNS-profil nyilvántartást szabályozó törvény. Mindez a többi európai országhoz képest viszonylag gyorsan, csupán 8 évvel azután történt, hogy Magyarországon először a Nemzeti Szakértői és Kutató Központ (NSZKK) jogelődjében, a Bűnügyi Szakértői és Kutatóintézetben (BSZKI) az igazságügyi DNS-vizsgálatokat bevezették (Pádár és mtsai, 2019). A hazai igazságügyi DNS-profil nyilvántartásokat jelenleg alapvetően két törvény szabályozza, melyek meghatározzák gyakorlatilag az összes olyan DNS-minta nyilvántartását és kerestetését, melyek a forenzikus személyazonosítás céljából jelentőséggel bírnak közigazgatási, körözési és büntetőügyekben. A magyarországi DNS-profil nyilvántartások egymástól elkülönítve és anonimizálva az alábbi személyek és biológiai anyagmaradványok DNS-profiljait tartalmazzák:

- bűncselekményekkel kapcsolatba hozható, ismeretlen eredetű nyomok és biológiai anyagmaradványok (például helyszínről, elkövetési eszközről, ruházatról);
- terhelt (gyanúsított, vádlott, elítélt) személyek;
- ismeretlen személyazonosságú holttestek, holttestmaradványok;
- eltűnt, körözött személyek és hozzátartozóik;
- büntetőeljárással összefüggésben véten nyomszennyezés kockázatát hordozó tevékenységet végző személyek.

A nyilvántartási törvények meghatározzák a DNS-profilok adatbázisba helyezhetőségének és kerestethetőségének kritériumait, a nyilvántarthatóság idejét és a nyilvántartásokból való törlés feltételeit. A terhelt személyek vonatkozásában például számos kivételtől eltekintve csak ötévi vagy ennél súlyosabb szabadságvesztéssel büntetendő szándékos bűncselekmény elkövetésével meggyanúsított személyek DNS-profiljai nyilvántarthatóak, vagy másik példaként a bűncselekmények ismeretlen eredetű DNS-profiljai nem hasonlíthatók össze az eltűnt személyek hozzátartozóiéval. A törvények ugyanakkor lehetővé teszik, hogy a nyilvántartásban kezelt adatok (ti. a DNS-profilok is) statisztikai, valamint tudományos kutatási célra – személyazonosításra alkalmatlan módon – átadhatók és felhasználhatók.

A magyarországi bűnügyi DNS-profilnyilvántartás az NSZKK (korábban BSZKI) kezelésében 2004 óta működik az FBI által kifejlesztett Combined DNA Index System (CODIS) adatbázis szoftver (URL1) segítségével. A szoftver adatbázisa külön regiszterekben, úgynevezett indexekben tárolja az eltérő DNS-mintakategóriák (például biológiai anyagmaradvány, terhelti) profiljait, lehetővé téve azoknak egymástól független kezelését. A 2023. szeptember 13-i állapot szerint 254 635 db személyi és 8 209 db anyagmaradvány DNS-profil található meg az adatbázisban és vesz részt a találatkeresésben. A magyar nyilvántartásokba összesen legalább 25 ezer új személyi profil és mintegy 1 400 új anyagmaradvány profil kerül évente rögzítésre. A nyilvántartott DNS-profilok lokuszkészlete egymással lehet azonos vagy nagyrészt átfedő. Az elmúlt évtizedekben a DNS-technológia

fejlődésével a megbízhatóbb személyazonosítás érdekében a markerkészlet 9 db autoszómás STR-lokuszról 23 db markerre bővült. A profilok egy kisebb hányada nem tartalmazza a vizsgálat idején elérhető kit segítségével meghatározható összes STR alléladatát, mivel a csekély mennyiségű vagy rossz minőségű biológiai anyagmaradványok esetében a korábban kifejtett allélkiesés jelensége miatt egyes lokuszokon eredménytelen volt a DNS-analízis. A véletlen egyezések lehetőségének minimalizálása érdekében a nemzetközi gyakorlatot követve a magyar adatbázisban olyan DNS-profilok nem kerülhetnek összehasonlításra, amelyek között nincs legalább 6 db átfedő, alléleredményt tartalmazó STR-lokusz.

A nyilvántartásokba napi rendszerességgel kerülnek be és törölődnek onnan DNS-profilok. Ugyancsak naponta kerülnek összehasonlításra az új DNS-profilok az adatbázisban már bent lévőkkel. Ugyan a CODIS-szoftver képes nemcsak autoszómás, hanem X- és Y-kromoszómás DNS-profilok, valamint az úgynevezett mitokondriális haplotípusok (mitotípus)² tárolására és összehasonlítására is, a magyarországi adatbázis csak autoszómás profilokon végez összehasonlításokat, mivel kizárólag a testi kromoszómás DNS-profilok alkalmasak konkrét személy önálló, egyértelmű beazonosításra. Ennek az az oka, hogy a nem autoszómás DNS-ek között egyáltalán nincs, vagy nincs teljeskörű DNS-átrendeződés (úgynevezett rekombináció) az ivarsejtképződés során, így az utód csak az egyik vagy másik szülője DNS-ét kapja meg, gyakorlatilag változatlan formában. Ennélfogva az egyik szülő és gyermeke azonos DNS-típussal rendelkezik (kivéve a nők X-kromoszómái esetében).

A CODIS-ban a DNS-profilok összehasonlítása kizárólag az alléltípusok számszerű páros összevetéséből áll. A hatályos nyilvántartási törvények a nyilvántartási találat kritériumát a DNS-profilok azonosságaként írják le. Az „azonosság” szűkebb (szigorúbb) szakmai és jogi értelmezése

² Mitokondriális haplotípus (mitotípus): a sejtek energiatermelő és -raktározó, sejttagon kívüli sejt szervecskéiben, a mitokondriumokban lévő DNS-en meghatározott DNS-típusok együttese

szerint kizárólag teljes azonosság fogadható el találatként, azaz az összehasonlítható két DNS-profil között – pontosabban azok közös (átfedő) STR-készletében – nem lehet egyetlen eltérés vagy különbség sem. Tehát így például a 14/15/16 alléltípusú, többszemélyi eredetű (kevert) minta nem azonos a 14/15 alléltípusú személlyel, ezért nem fogadható el találatként. A CODIS-szoftver ugyanakkor igény esetén lehetőséget biztosít egymástól részben különböző DNS-profilok közötti találatkeresésre is. A szoftver mérsékelt és alacsony szigorúságú (moderate/low stringency) találatkeresési beállításai megengedik az egyezést mutató allélokon kívül egyéb allélok jelenlétét vagy egyes keresett allélok hiányát. A magyar igazságügyi DNS-profil nyilvántartások és azok működési kereteit meghatározó jogszabályok végrehajtási módja ugyan csak egy személytől származó DNS-profilok vagy profilkomponensek összehasonlítását teszi lehetővé, azonban a csökkentett szigorúságú találatkeresés révén olyan többszemélyi profilok is összehasonlításra kerülhetnek, amelyekben a későbbiekben tárgyalt szakértői értékelési módszerekkel egy egyetlen személytől származó DNS-profilkomponens volt elkülöníthető. Ezeket az általában biológiai anyagmaradványokból meghatározott speciális profilokat a CODIS-adatbázisunk egy elkülönített, úgynevezett kijelölt kevert minta (assigned mixture) nevű indexben tárolja.

Attól függően, hogy a CODIS mely indexei között mutatják ki a DNS-profilok azonosságát (egyezését), személy/személy, személy/anyagmaradvány vagy anyagmaradvány/anyagmaradvány találatokról beszélhetünk. Személy/személy találat előfordulhat akkor, ha ugyanattól a személytől – például hamis személyazonosító dokumentuma miatt – többször vesz mintát a hatóság, illetve az egypetéjű ikrek esetében, mivel a DNS-nyilvántartással kompatibilis genetikai markerekkel nem mutatható ki eltérés egypetéjű ikerpárok DNS-profiljai között. A személy/anyagmaradvány találat leggyakoribb esetei, amikor a bűncselekmény helyszínén, az elkövetési eszközön vagy az elkövetéskor viselt ruházaton rögzített minta DNS-profilja egyezést mutat egy terhelt személy profiljával, beazonosítva ezzel a

nyomhagyót, mint lehetséges elkövetőt. Amennyiben a nyilvántartási találatig az eljáró hatóságnak nem volt információja a beazonosított személyről vagy annak nyomhagyásáról, akkor a találatot az angol szakkifejezésből magyar nyelvbe tükörfordítással átültetve hidegtalálatnak (cold match/hit) nevezzük. A hidegtalálat kifejezés érzékletesen szemlélteti ezt az esetet, amikor is az eljáró hatóság kezében a DNS-bizonyítékon kívül nincs más adat az illető személy nyomhagyására vonatkozóan, így azt a kérdéses cselekmények és körülmények megnyugtató tisztázása érdekében lehetőség szerint egyéb nyomozati vagy bizonyítási eszközökkel meg kell erősíteni. Személy/anyagmaradvány találatnak egy másik típusa az ismeretlen holttest DNS-profiljának egyezése egy eltűnt személy profiljával, vagy nagyfokú hasonlósága egy eltűnt személy rokonával. Ez utóbbi eset nem klaszszikus, egy az egyes (ügynevezett direkt) találatnak, hanem ügynevezett rokonsági találatnak minősül. Az anyagmaradvány/anyagmaradvány találat pedig rámutat a biológiai anyagmaradványok közös, bár még ismeretlen személyi eredetére. A nyomhagyó ismeretlen személye ellenére ez a találat típus is hasznos információt szolgáltat a nyomozó hatóságok részére, mivel összeköti egymással a különböző bűncselekményeket, azok helyszíneit vagy az ismeretlen személyazonosságú holttestrészeket. Köszönhetően annak is, hogy Magyarország lakosságának kb. 2%-a van nyilvántartva a DNS-adatbázisban, a magyar igazságügyi DNS-profil nyilvántartások jelenleg naponta átlagosan egy-két ügy sikeres megoldásához járulnak hozzá a találatokkal.

Az igazságügyi genetikai szakértői vizsgálatok eredményeinek interpretációja

Az igazságügyi genetikai szakértői munka nem ér véget a DNS-profilok meghatározásával. Már maga a profilmeghatározás, különösen az ügy szempontjából releváns feltételezett nyomhagyó [továbbiakban POI- (person of interest)] személy alléljainak az azonosítása is sokszor nehézségbe ütközik a biológiai anyagmaradvány (például érintési nyomok) alacsony

DNS-tartalma, többszemélyi (kevert) eredete, bomlottsága vagy háttér- és felülszennyezettsége miatt. Mindezek a körülmények a POI egyes alléljainak véletlenszerű, részleges vagy teljes kimutathatatlanságát (allél- és lokuskiesés), és/vagy konkrét nyomhagyó személyekhez nem köthető idegen (sporadikus) allélok véletlenszerű megjelenését (allélbeesés) okozhatják. Az allélkiesés és -beesés statisztikai alapon kezelhető véletlenszerűsége miatt a minták szakértői elemzése során ismétlődő (replikált) DNS-profil vizsgálatokat végzünk annak érdekében, hogy a nyomhagyó személyek alléljait minél nagyobb eséllyel meghatározhassuk, és azokat elkülöníthessük a sporadikus alléloktól. Ehhez már rendelkezésünkre áll olyan szakértői szoftverrendszer, amely a szakértők által könnyen áttekinthető és kezelhető formában tudja megjeleníteni az ismétlések DNS-profil eredményeit (Benschop et al., 2019).

Az igazságügyi genetikus szakértő fő feladatát, az ismeretlen biológiai anyagmaradvány, ismeretlen személyazonosságú élő vagy holt személy, valamint a vérrokonság azonosítását a DNS-profilok összehasonlításával végzi el. Itt most egyelőre „csak” a személyi vagy anyagmaradvány minta forrásának, származásának DNS-szintű azonosításáról van szó. Egyszemélyi eredetű DNS-profilok összehasonlító – bár néha kihívásokkal is terhelt – vizsgálata sokkal egyszerűbb mind az ügyön belüli összehasonlítások, mind a nyilvántartásokban való kerestetés során. Emiatt a szakértői elemzések során törekszünk arra, hogy a biológiai anyagmaradványok adta lehetőségek mentén a kevert DNS-profilok egyszemélyi eredetű komponenseit meg tudjuk állapítani. Ehhez sejtdifferenciált DNS-tisztítási technológia (Butler, 2011), illetve az egyes DNS-tulajdonságoknak a biológiai anyagmaradványban lévő mennyiségi viszonyait kifejező allélintenzitás elemzési lehetőség áll a szakértők rendelkezésére (Benschop & Sijen, 2014). Az egyre növekvő arányban vizsgálatra kerülő kevert anyagmaradványok egy jelentős részében sajnos azonban nincs mód a nyomhagyók DNS-profiljainak egyértelmű szétválasztására. Modern statisztikai módszerek (Bleka, Benschop, Storvik & Gill, 2016) alkalmazásával a vizsgálati minta többségét adó (úgynevezett domináló) személyek DNS-profiljának

elválasztása egyre inkább megbízható, ugyanakkor a kevesebb DNS-t hátra hagyó egyének (akik között sokszor szerepel a POI) profiljának valószínűsítése továbbra is kihívást jelent a szakma számára.

Biológiai anyagmaradványok származásának megállapításánál, illetve rokonsági vizsgálatok során az anyagmaradványok vagy a vizsgált személyek azonos (közös) eredetének meghatározását végezzük el. Az azonos vagy nagyon hasonló tulajdonságok meglepte alátámasztja a származást. Ez nemcsak biológiai, hanem egyéb nyomokra is érvényes (például ujjnyom, ruhaszövet elemi szálai, üvegmaradvány, cipőtalplenyomat). Ennek alapján az ebben a fejezetben leírt úgynevezett Bayes-elvű statisztikai interpretáció nemcsak genetikai, hanem más igazságügyi szakterületeken is alkalmazható. Amennyiben a vizsgált DNS-profilok között az eltérés megfelelő mértékű, akkor kategorikusan kizárható a két DNS-minta közös eredete. A teljes egyezés vagy a nagyfokú hasonlóság felveti azt a kérdést, hogy a származás megállapítható-e kategorikusan, vagy számolni kell-e a véletlen egyezés lehetőségével is? Kategorikus véleményadás során a szakértői bizonytalanság nullának tekinthető, azaz a szakértői állítással ellentétes eset előfordulása kizárt. Fontos azonban megjegyezni, hogy tudományos ismervek alapján a származás nem bizonyítható kategorikusan. A szakértő a tudományos vizsgálati eredmények és tapasztalatai alapján mentálisan meggyőződik a származás kategorikus voltáról. A valószínűsítő véleményadás során ezzel ellentétben a szakértő tudományos, statisztikai módszerekkel megbecsüli a származás bizonytalanságát, amely mindig nagyobb nullánál. Az ártatlanság vélelme alapján a szakértő ezt a becslést úgy végzi el, hogy a perbe vont félre nézve a legkevésbé terhelő, de valószínű szakértői valószínűsítést adja. Ezt a világszerte megkövetelt és alkalmazott megközelítést konzervatív becslésnek hívjuk. Azt is fontos hangsúlyozni, hogy a valószínűsítő véleményadás is tartalmazhat szubjektív elemeket, ugyanis a szakértés a gépesítés nagyfokú fejlettsége mellett sem függetleníthető az emberi döntésektől (Füredi, 2015).

A mindennapi életben is használt valószínűség fogalma többféleképpen is definiálható. A valószínűség kifejezhető egy esemény bekövetkezésének

számával a lehetőségek (például próbák) számához viszonyítva. A valószínűség ugyanakkor csak egy adott körülmény- és feltételrendszerre lehet igaz. Ha a körülmények változnak, akkor a valószínűség számértékében is változhat. Minden valószínűség egyben feltételes valószínűség is, melynek matematikai leírása a következő: $Pr(H|E)$, ahol Pr a valószínűség (probability), H a vizsgált feltevés (hipotézis) és E a feltétel [itt most, mint szakértői esemény vagy bizonyíték (event, evidence)]. A feltételes valószínűség tehát E feltétel (bizonyíték) fennállása esetén H feltevés bekövetkeztének a valószínűsége. A DNS-profil nem tekinthető egyedinek, azaz nem 100%-ig biztos, hogy a Földön két személy nem rendelkezhet ugyanazzal a DNS-profillal. A DNS-profil egyediségének kérdését alapvetően négy, egymással részben összefüggő tényező befolyásolja. Az egyik, hogy a Föld teljes lakosságát (illetve állati, növényi populációit) nem tudjuk maradéktalanul megtipizálni. Ha erre képesek is lennénk, akkor ezt mindig folytatni kéne az újszülöttekkel. A másik egyediséget zavaró jellegzetesség, hogy az élőlények egy fajba tartozó egyedei szaporodási közösségükben közelebbi-távolabbi rokonságban állnak egymással, azaz genetikai állományuk egy része közös. Ennek szélsőséges esetét az egypetűjű ikrek képviselik, akik a jelenleg rutinszerűen vizsgált markereken azonos tulajdonságokkal rendelkeznek. A harmadik nehezítő tényező az, hogy az egyes lokuszok változottságának (polimorfizmusának) mértéke populációnként/földrajzi területenként változhat. Tehát hiába találjuk egyedinek az egyik népességben a kérdéses DNS-típust, ha a másik populációban esetleg épp az a leggyakoribb. Az egyediség kérdéskörében az utolsó alproblémát az elsősorban a bűncselekmények során előforduló kevert biológiai anyagmaradványok jelentik. A kevert anyagmaradványok egynél több személy vagy egyed sejtjeit és így DNS-ét tartalmazzák, tehát a vizsgálati minta egyszemélyi eredetéről eleve nem is beszélhetünk. Mindezen megfontolások alapján az igazságügyi genetikus szakértők – eltérően például a daktiloszkópus szakértők által alkalmazott gyakorlattól – nem kategorikus, hanem valószínűsítő véleményt adnak a személyazonosítás vagy származás-megállapítás

során. Két biológiai anyag közös eredete vagy a vérrokonság akkor valószínűsíthető, ha a DNS-profilok egyezést vagy nagyfokú hasonlóságot mutatnak (Füredi, 2015).

A származás, leszármazás statisztikai interpretációjára, a véletlen egyezés valószínűségének megbecslésére a szakemberek több, egymásra épülő módszert dolgoztak ki az elmúlt évtizedek során. Két alapmódszer alkalmazásával – a DNS-profil előfordulási gyakoriságának, illetve véletlen egyezési valószínűségének megbecslésével – eljuthatunk a jelenleg világszerte alkalmazott statisztikai eljáráshoz, a Bayes-féle hipotézisteszteléshez (Foreman, Smith & Evett, 1997). Az igazságügyi eljárásokban a DNS-profilok egyezése csak akkor bír megfelelő bizonyító erővel, ha a megállapított DNS-profil olyan ritka a népességben, hogy elhanyagolható az esélye annak, hogy a kérdéses személyen kívül más is rendelkezhet ugyanazzal a DNS-profillal. A DNS-vizsgálati eredmények bizonyító erejének megbecsléséhez ismernünk kellene tehát a DNS-profil előfordulási gyakoriságát a népességben belül. A biológiai nyom DNS-profiljának feltételes előfordulási valószínűségét a következőképpen jelölhetjük: $P = Pr(G_c/H_d)$, ahol G_c a biológiai nyom DNS-profilja és H_d a védelem azon feltételezése, hogy a biológiai nyomot nem a terhelt hagyta hátra. A modern genetikai vizsgálórendszerek nagy bizonyító ereje a DNS-profilok rendkívül ritka voltából fakad, amely azonban gondot is okoz a valószínűség megbecslésénél. A legtöbb DNS-profil ugyanis olyan ritka, hogy egyszer sem fordul elő az egyes népségeket reprezentáló populációs adatbázisokban. Szerencsére bizonyos feltételek teljesülése esetén a DNS-profil gyakorisága kifejezhető a profilt alkotó allélok előfordulási gyakoriságának segítségével. A főbb feltételek közé tartozik az, hogy testi kromoszómás lokuszokat érintsen a vizsgálat (ilyenek a DNS-profil nyilvántartások markerei is), a használt allélgyakorisági adatbázis legyen kellően reprezentatív, valamint a vizsgált népesség nem térhet el lényegesen a genetikai egyensúly állapotától. E feltételek az emberi népességben általában teljesülnek az igazságügyi eljárásokban vizsgált testi kromoszómás markereken, így megvalósul az allélok egymástól független öröklődése a lokuszokon belül és a lokuszok között.

A független öröklődés pedig lehetővé teszi azt, hogy a DNS-profil előfordulási gyakorisága megbecsülhető az allélgyakoriságok egyszerű szorzataként (Füredi, 2015).

Talán meglepő módon a származás bizonyításához azonban nem pontosan a DNS-profil előfordulási gyakoriságára van szükségünk, hiszen a profilt a biológiai anyagmaradványból – és teljes egyezés esetén a POI-ból – már kimutattuk. Igazából arra vagyunk kíváncsiak, hogy mi annak a valószínűsége, hogy a biológiai anyagmaradvány DNS-profilját a kérdéses személyen (például terhelt) kívül még egyszer megfigyeljük a népességben. Ennek az úgynevezett DNS-profil egyezési valószínűségnek a jelölése: $P_m = Pr(G_c/G_s, H_d)$, ahol az előfordulási gyakorisághoz (P) képest a feltételrendszerben megjelenik egy új elem, az egyezést mutató személy (például gyanúsított) DNS-profilja (G_s). Ha a védelem feltevésében a terhelt és a tényleges tettes között még távoli rokonság sem áll fenn, akkor a terhelt DNS-profiljának ismerete nem befolyásolja a profilegyezési valószínűséget, így az ekvivalens a DNS-profil előfordulási gyakoriságával ($P_m = P$). A védelem feltevésében azonban a terhelt közeli vagy távoli rokoni kapcsolatban is lehet a tényleges nyomhagyóval. Ez az úgynevezett rokonsági/alpopulációs hatás, mint eltérő körülményrendszer módosítja a DNS-profil egyezési valószínűség értékét. Kriminálisztikai kutatások kimutatták, hogy a terhelt és a tényleges tettes (amennyiben nem egy személy) igen gyakran ugyanabba az alpopulációba (szűkebb közösségbe) tartozik. Az alpopulációban a biológiai anyagmaradvány lokuszainak alléltípusa gyakoribb lehet, mint a teljes népességben, mivel az alpopulációk tagjai között nagyobb a genetikai kapcsolat (korreláció) a közös ősök, a nem véletlenszerű párválasztás és szélsőséges esetben a beltenyésztés miatt. Az alpopulációs hatás mértéke a népességben populáció-genetikai vizsgálatokkal megállapítható, és az úgynevezett közös leszármazási együtthatóval (F_{ST} vagy θ) jellemezhető. A θ -érték önmagában általában csak kis mértékben növeli a DNS-profil egyezési valószínűséget, azonban közeli vérrokonság esetén a rokonsági hatás már nagyfokú. A θ -korrekció már régóta részét

képezi a rutin statisztikai genetikai alapú személyazonosításnak (Balding & Nichols, 1994; Füredi, 2015).

Mind a vád vagy felperes, mind a védelem vagy alperes állítását figyelembe vevő statisztikai interpretációs módszer a Bayes-féle hipotézistesztelés, amely a legsokrétűbb és így világszerte a legelfogadottabb eljárás az igazságügyi személyazonosítási és származás-megállapítási gyakorlatban. A Bayes-féle hipotézistesztelés során legalább (de nem feltétlenül csak) két, egymást kölcsönösen kizáró feltevést vetünk statisztikai vizsgálat alá. A vád hipotézise (H_p) például szólhat így: „*A helyszínen talált vérnyom a terhelttől származott*”, míg a védelem (H_d) mondhatja azt, hogy „*A vérnyom a terhelttel rokonságban nem lévő ismeretlen személytől származott*”. E feltevéseket magában foglaló, esély formájú Bayes-képlet leírása a következő:

$$\frac{Pr(H_p|E,I)}{Pr(H_d|E,I)} = \frac{Pr(E|H_p,I)}{Pr(E|H_d,I)} \times \frac{Pr(H_p|I)}{Pr(H_d|I)}$$

$PoPR$ LR $PrPR$

A három esély (odds) tagból álló Bayes-képlet a valószínűségszámítás alaptörvényeiből levezethető. A képlet első tagja az úgynevezett *a posteriori* (következtetett) valószínűségi hányados ($PoPR$). Ez tulajdonképpen nem más, mint a DNS- és egyéb bizonyítékok/információk alapján a két hipotézis valószínűségének aránya. Az eljáró hatóság ezt a hányadost szeretné megállapítani, mivel ha $PoPR > 1$, akkor az terhelő lehet a vádlottra nézve, míg az egynél kisebb, különösen a nullát megközelítő $PoPR$ érték felmentést eredményezhet. Tekintettel azonban arra, hogy a genetikus szakértő általában nem ismeri vagy nem mérlegelheti az egyéb bizonyítékokat és információkat, így a gyakorlatban a szakértő *a posteriori* valószínűséget általában nem véleményezi (erre kivételt az apasági/rokonsági teszteknel találunk). A Bayes-képlet második tagja a valószínűségi (likelihood) hányados (LR), amely a két hipotézis teljesülésének feltételezése

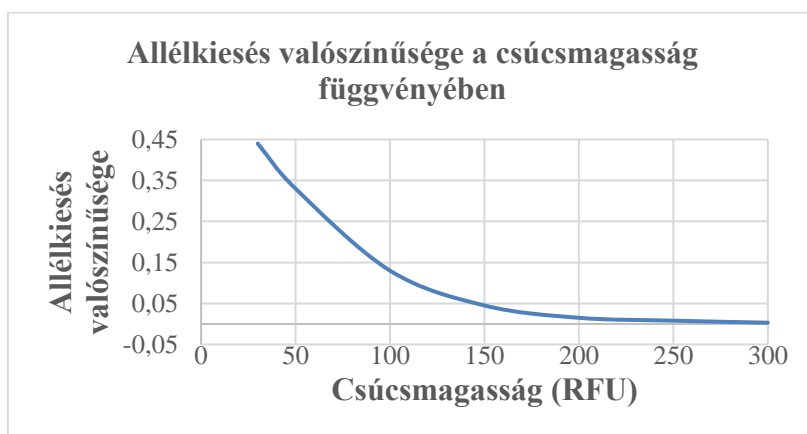
esetén a DNS-vizsgálati eredmény (E) valószínűségeinek aránya. Itt tehát nem a feltevések, hanem a genetikai eredmények megfigyelési/kimutatási valószínűségéről beszélünk. A valószínűségi hányados kiszámításánál a feltevéseket igaznak fogadjuk el, és azt vizsgáljuk, hogy ezekben az esetekben milyen valószínűséggel várható a meghatározott DNS-profil megjelenése a biológiai anyagmaradványban. Tekintettel arra, hogy az LR-érték kalkulációjához nincs feltétlen szükség egyéb, DNS-vizsgálatot nem érintő bizonyítékok és információk mérlegelésére, ezért mód nyílik az elfogulatlan értékelésre, ami miatt is a valószínűségi hányados számítás a nemzetközi és hazai gyakorlatban a statisztikai interpretáció alappillére. A Bayes-képlet utolsó tagja az *a priori* (előzetes) valószínűségi hányados ($PrPR$). A $PrPR$ az egyéb (előzetes) bizonyítékok és információk alapján a két hipotézis valószínűségének aránya, amely általában nem szakértői mérlegelés tárgya, és a szakértő sokszor nem is ismeri. Az előzetes valószínűségek ismerete nélkül tehát a posterior valószínűségek nem határozhatók meg (Füredi, 2015).

A valószínűségi hányados az egy személytől eredeztethető vizsgálati mintákra a következőképpen számítható ki, ha a módosító tényezők (például laboratóriumi hiba, bomlásból vagy alacsony DNS-tartalomtól fakadó alléltulajdonság-vesztés) kizárhatók:

$$LR = \frac{Pr(E|H_p, I)}{Pr(E|H_d, I)} = \frac{I}{Pr(G_c|G_s, H_d, I)} \approx \frac{I}{Pr(G_c|H_d, I)}$$

Ennek alapján a valószínűségi hányados egyszemélyi mintákra a profi-legezési valószínűség reciprokával, illetve, ha a rokonsági/alpopulációs hatás elhanyagolható, akkor a profilgyakoriság reciprokával egyezik meg (Füredi, 2015). Az igazságügyi DNS-technológia rendkívül gyors fejlődésének köszönhetően az STR-lokuszok szimultán vizsgálatával ma már olyan magas LR-érték nyerhető, amely már a világegyetemben levő csillagok becsült számát (10^{24}) (URL7) is nagyságrendekkel meghaladhatja.

Az alacsony DNS-tartalmú [low copy number (LCN), low level (LL), low template (LT)] vagy bomlott biológiai anyagmaradványok PCR-vizsgálatával sok esetben nem lehet tökéletesen meghatározni a nyomhagyó személy DNS-profilját. LCN-mintákban a PCR során egyes allélok sokszorozódása véletlenszerűen (sztochasztikusan) nem éri el az analitikai detektálhatóság szintjét, míg más, a nyomhagyó személy(ek)től nem származó (szennyező) allélok felsokszorozott mennyisége meghaladja a kimutathatósági küszöbértéket. Az előbbi esetet allélkiesésnek (D), az utóbbit allélbeesésnek (C) hívjuk. E két, igen gyakran együtt jelentkező technológiai jelenség hatására a detektált DNS-profil általában kis mértékben, de eltérhet a nyomhagyó személy(ek) genetikai profiljától. Elméleti megfontolások és gyakorlati tapasztalatok alapján megállapítható, hogy az allélkiesés valószínűsége [$Pr(D)$] és az allélbeesés valószínűsége [$Pr(C)$] fordítottan arányos a kimutatható, illetve a beeső allélok detektálási intenzitásával (RFU) (F6. számú ábra). Az allélkiesés és -beesés mértéke a vizsgálati mintában lévő DNS mennyiségének és minőségének, illetve a PCR-amplifikálás és DNS-fragmensanalízis körülményeinek a függvénye.



F6. számú ábra

Az allélkiesés valószínűsége a kimutatott (maradó) allélok csúcsmagasságának függvényében
RFU: relatív fluoreszcencia egység

Az allélkiesés és -beesés valószínűsége a Bayes-elvű statisztikai genetikai számításokba beépíthető. Például a vád hipotézisében egy A genotípusú biológiai anyagmaradványnak egy A/B genotípusú személytől való származása csak akkor lehetséges, ha az alacsony DNS-tartalmú mintában:

- az A allél nem tűnik el, melynek valószínűsége $1 - Pr(D) = Pr(\bar{D})$;
- a B allél kiesett, melynek valószínűsége $Pr(D)$;
- egyéb allél nem jelenik meg, melynek valószínűsége $1 - Pr(C) = Pr(\bar{C})$.

A vád állítására a genetikai bizonyíték valószínűsége a fenti valószínűségek szorzata: $Pr(E | H_p) = Pr(\bar{D})Pr(D)Pr(\bar{C})$. Hasonló okfejtéssel kapható meg a védelem feltevésére is a kevert DNS-profil valószínűsége, illetve a valószínűségi hányados (Füredi, 2015).

Az igazságügyi DNS-vizsgálatok érzékenységének növekedésével egyre több biológiai (főleg érintési, érintkezési) nyomról állapítható meg, hogy nem egy, hanem több személytől származott. Klasszikus biológiai anyagmaradványok és testvadások (vér, nyál, ondó, hüvelyvadásék) személyek közötti keveredése is gyakran előfordul például szexuális bűncselekmények esetén. A biológiai anyagok keveredése különböző lehet a személyek száma, illetve az egyes személyi komponensek abszolút és relatív mennyisége tekintetében. A különböző típusú kevert anyagmaradványok eltérő értékelési módszert kívánnak meg vagy tesznek lehetővé az interpretációt nehezítő körülmények és/vagy az eldöntendő kérdés függvényében. A legegyszerűbb értékelés során nem végzünk statisztikai számítást, csupán csak azt vizsgáljuk, hogy a kérdéses személy DNS-profilja kimutatható volt-e a nyomból (lehet – nem lehet kategorizálás). Abban az esetben, ha a kevert DNS-profilban az intenzívebb allélcúcsok alapján beazonosítható egy egyszemélyi domináló (major) komponens, akkor az egysze-

mélyi major DNS-profilra az egyezési valószínűség és a valószínűségi hányados a korábban megismert módszer szerint kiszámítható. A kevert biológiai nyomok teljes körű, de egyben a legösszetettebb statisztikai értékelését a Bayes-féle valószínűségi hányadosnak az egész kevert DNS-profilra (nemcsak annak major komponensére) történő meghatározása jelenti (Füredi, 2015).

A Bayes-elvű hipotézisvizsgálás nagyszerűsége abból is fakad, hogy rugalmasan lehet alkalmazni speciális problémák megoldására is, mint amilyen helyzetet a kevert biológiai anyagmaradványok jelentenek. Az alapelv itt is ugyanaz, mint az egyszemélyi DNS-profiloknál: az egyes hipotézisekhez tartozó genetikai bizonyíték valószínűségét kell megbecsülni. Kevert mintáknál – az egyszemélyi eredetű biológiai anyagmaradványokhoz viszonyítva – azonban gyakrabban lehet valóságalapja kettőnél több feltevésnek is. Például egy csoportosan elkövetett szexuális bűncselekmény bizonyítása során a vád és a védelem részéről az alábbi hipotézisek is elhangozhatnak:

A kevert DNS-profilú hüvelytörlet a sértett és

- H_p : az egyik vádlott biológiai anyagát tartalmazza (vád),
- H_{d1} : egy ismeretlen férfi biológiai anyagát tartalmazza (védelem1),
- H_{d2} : két ismeretlen férfi biológiai anyagát tartalmazza (védelem2),
- H_{d3} : a gyanúsított bátyjának biológiai anyagát tartalmazza (védelem3).

Kevert minták esetén a valószínűségi hányados kiszámítása úgy történik, hogy mind a vád, mind a védelem feltevésére külön-külön meghatározzuk azon genotípusok kombinációit és a hozzájuk tartozó valószínűségeket, amelyek pontosan kiadják az anyagmaradvány kevert DNS-profilját. Ezt egy egyszerű példával szemléltetve, ha a vizsgálati minta genotípusa $A/B/C$, a sértett A/B és a terhelt B/C , akkor

- $Pr(E/H_p) = 1$, mivel a sértett és a terhelt genotípusa együtt megegyezik az anyagmaradvány genotípusával;
- $Pr(E/H_d)$ kiszámításához fel kell írunk az összes lehetséges genotípus-kombinációt például arra a hipotézisre, hogy a minta a sértettől és egy ismeretlen személytől származott: $(A/B)_{\text{sértett}}$ és $(C/C)_{\text{ismeretlen}}$ vagy $(A/B)_{\text{sértett}}$ és $(A/C)_{\text{ismeretlen}}$ vagy $(A/B)_{\text{sértett}}$ és $(B/C)_{\text{ismeretlen}}$. Ennek megfelelően a valószínűség: $1 \times Pr(C/C)_{\text{ismeretlen}} + 1 \times Pr(A/C)_{\text{ismeretlen}} + 1 \times Pr(B/C)_{\text{ismeretlen}}$.

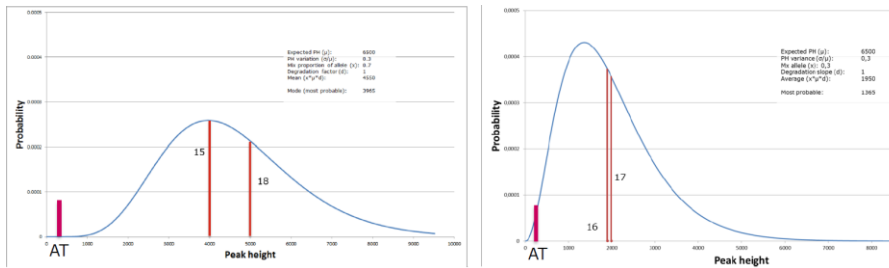
A valószínűségi hányados tehát erre a hipotézis-párra:

$$LR = \frac{1}{Pr(C/C)_{\text{ismeretlen}} + Pr(A/C)_{\text{ismeretlen}} + Pr(B/C)_{\text{ismeretlen}}}$$

A vád valószínűsége (az LR számlálója) azonban rögtön megváltozik akkor, ha a vád – ellentétben a védelemmel – azt állítja, hogy volt még egy ismeretlen elkövető is, akinek a DNS-profilja elfedve ugyan, de megtalálható a minta kevert DNS-profiljában. Ez az átfedés csak a következő ismeretlen elkövetői genotípus-kombinációk esetében valósulhat meg: $(A/A)_{\text{ismeretlen}}$ vagy $(A/B)_{\text{ismeretlen}}$ vagy $(A/C)_{\text{ismeretlen}}$ vagy $(B/B)_{\text{ismeretlen}}$ vagy $(B/C)_{\text{ismeretlen}}$ vagy $(C/C)_{\text{ismeretlen}}$. A vád valószínűsége ebben az esetben a $Pr(A/A)_{\text{ismeretlen}} + Pr(A/B)_{\text{ismeretlen}} + Pr(A/C)_{\text{ismeretlen}} + Pr(B/B)_{\text{ismeretlen}} + Pr(B/C)_{\text{ismeretlen}} + Pr(C/C)_{\text{ismeretlen}}$ összeg lesz, amely mindig kisebb 1-nél, ha az emberi népességben a vizsgált lokuszon A, B és C allél mellett más allél(ok) is előfordul(nak). Ennélfogva az így kiszámolt valószínűségi hányados kisebb lesz a fenti képlettel kalkulált értékhez képest, azaz a genetikai bizonyíték kevésbé lesz terhelő a vádlottra nézve. Amennyiben az eljáró hatóság a vád/felperes és/vagy a védelem/alperes oldaláról egynél több magyarázatot is valószínűnek tart a biológiai anyagmaradvány keletkezésére, akkor ennek megfelelően a szakértőnek több LR-értéket kell kiszámi-

tania. Az eljáró hatóság az ártatlanság vélelme alapján általában a gyanúsítotttra nézve a legkevésbé terhelő (legkisebb) LR-értéket építi be a bizonyítékok láncolatába. A hipotézisek felállításához az alpopulációs elmélet szerint definiálni kell az ismert nyomhagyó személyeket, az ismert, biológiai nyomot nem hagyó személyeket, valamint az ismeretlen nyomhagyó személyek számát. Ez utóbbi sokszor vita tárgya lehet. Ha nincs a felek között megegyezés, akkor szakértői gyakorlat szerint általában 2–4 (ismert + ismeretlen) személyre kell végigszámolni a valószínűségi hányadosokat. További statisztikai értékelési nehézséget jelent az, ha közeli rokon személyek vagy különböző populációcsoportba tartozó emberek biológiai anyaga van a vizsgálati mintában, és/vagy a kevert DNS-profil nem teljes (részleges) a biológiai anyagmaradvány alacsony DNS-tartalma vagy bomlott állapota miatt (Füredi, 2015).

Az előbbieken leírt LR-kalkulációkat félfolytonos statisztikai modellű (kvalitatív) valószínűsítési (probabilisztikus) számításoknak hívjuk, mivel az alléltípusok mellett – ahogy azt már láttuk – képesek statisztikailag kezelni a bűnügyi anyagmaradványoknál gyakran előforduló egyes jelenségeket (például allélkiesés és allélbeesés), azonban az allélok detektált intenzitását, mint hasznosítható információt nem veszik figyelembe. Az allélintenzitások többszemélyi eredetű (kevert) biológiai anyagmaradványoknál bírnak nagy jelentőséggel, ahol is segíthetnek az egymáshoz képest több vagy kevesebb DNS-t hátra hagyó személyek megbízhatóbb azonosításában (Bleka, Benschop, Storvik & Gill, 2016). A fluoreszcens PCR-rel felsokszorozott és kapilláris elektroforézis (CE) fragmensanalízissel elválasztott heterozigóta STR-allélok intenzitása úgynevezett gamma-eloszlással modellezhető. A PCR-amplifikáció sztochasztikája miatt még egysezemélyi mintáknál is gyakran előfordul, hogy a heterozigóta STR-lokuszt két allélja nem egyforma erősségű. Optimális vizsgálati körülmények között ez az eltérés kis mértékű és kisebb gyakoriságú. Ilyenkor a vizsgált lokuszt alléljainak várható csúcsintenzitásai keskeny csúcsú valószínűségi sűrűségfüggvényt mutatnak (F7. számú ábra jobb oldali grafikonja).



F7. számú ábra

Egy STR-lokuszt két heterozigóta genotípusának allélcúcs-intenzitási gyakoriságai sűrűségfüggvényekkel ábrázolva

Alacsony DNS-tartalmú, bomlott vagy PCR-gátló (inhibitor) anyagokat tartalmazó biológiai anyagmaradványok esetén az intenzitásokban mérhető eltérés nagyobb mértékű és gyakoribb lehet, így ilyenkor ez a sűrűségfüggvény-csúcs laposabb (F7. számú ábra bal oldali grafikonja). A teljesen folytonos statisztikai modellű (kvantitatív) szoftverek a nyomhagyásra felállított hipotéziseknek megfelelően, az előbbieken tárgyalt, vagy ahhoz hasonló intenzitáseloszlási függvények segítségével – kevert minták esetében is – kiszámítják a vizsgált feltevéshez leginkább illeszkedő DNS-profilok valószínűségeit a DNS-profil alkotó allél-, stutter- és allélbeesési csúcsok intenzitásának, illetve a DNS-profilok keveredési arányának és degradációs fokának megbecslésével. A szoftverek az illesztést általában a Maximum Likelihood Estimation (MLE) módszerével végzik, amikor is a biológiai anyagmaradványt hátrahagyó személy(ek) DNS-profiljának ismeretlen, meghatározandó paramétereit (a nyomhagyó személyek alléljainak intenzitását és azok varianciáját, keveredési arányát, stutereződését és degradációját) sorozatosan mindaddig változtatgatják, amíg az adott hipotézisre számolt DNS-profil valószínűség maximális értéket vesz fel. Amelyik hipotézisnek (leg)magasabb az így meghatározott DNS-profil valószínűsége, azt a feltevést támasztja alá a kvantitatív statisztikájú genetikai vizsgálat. A kvantitatív statisztikájú szoftverek az egyes hipotézisekhez illesz-

tett allélintenzitásoknak a megfigyelt értékekhez való illeszkedési pontosságát is meg tudják határozni úgynevezett modellilleszkedési validációs tesztek segítségével, és jelzik, ha a pontatlanság már nem elfogadható mértékű (Benschop et al., 2019).