

Intesztinális mikrobiom-elemzési munkafolyamatok optimalizációja

36. OTDK Biológia Szekció

Henkrich Miksa

Biológia BSc IV. évfolyam

Témavezetők:

Dr. Tombácz Dóra egyetemi docens

Gulyás Gábor PhD-hallgató

SZTE SZAOK Orvosi Biológiai Intézet

GeMiNI munkacsoport

Szeged

2022

Tartalomjegyzék

Tartalomjegyzék	2
Rövidítésjegyzék	3
1. Bevezető	4
1.1. A mikrobiom-kutatás jelentősége	4
1.2. A bél mikrobiom-kutatás jelenlegi kihívásai	5
1.2.1. Módszertani eredetű adattorzítások székletminta feldolgozása során	7
1.3. A bél mikrobiom-kutatás alkalmazott módszerei	8
1.3.1. DNS-izolálás székletmintából	8
1.3.2. Szekvenálási megközelítések	10
1.3.3. Bioinformatika	15
2. Célkitűzés	19
3. Anyagok és módszerek	20
3.1. Feldolgozott minták	20
3.1.1. Kutya székletminta	20
3.1.2. Kontroll mikroba mix	20
3.2. A mintafeldolgozás összefoglalása	21
3.3. DNS-izolálás és minőségellenőrzés	23
3.4. Könyvtárkészítés	23
3.5. Szekvenálás	24
3.5.1. Rövid leolvasású Illumina szekvenálás	24
3.5.2. Hosszú leolvasású ONT és PacBio szekvenálás	24
3.6. Bioinformatikai analízis	25
4. Eredmények és értékelésük	25
4.1. DNS-izolálás minősége	25
4.2. Szekvenálási könyvtárak minősége	26
4.3. Szekvenálási eredmények, a minta összetétele	27
4.3.1. Shotgun metagenomi szekvenálás	27
4.3.3. 16S rDNS hosszú leolvasású szekvenálás	28
4.3.4. 16S rDNS V régiók amplikon szekvenálása	31
4.3.6. A kontroll mikroba mix szekvenálása	33
4.3.7. Gram-pozitív kinyerés székletmintából	35
5. Összefoglalás	37
Köszönetnyilvánítás	38
Irodalomjegyzék	38
Ábrák forrása	40

Rövidítésjegyzék

16S: 16 Svedberg egység

ASV: amplitudon szekvencia variáns

DIN: DNS integritási érték

DNS/DNA: dezoxiribonukleinsav

gDNS: genomi DNS

EM: expectation-maximization algoritmus

NGS: újgenerációs szekvenálás

ONT: Oxford Nanopore Technologies

OTU: operációs taxonómiai egység

PacBio: Pacific Biosciences

PCR, qPCR: polimeráz láncreakció, kvantitatív ~

RNS: ribonukleinsav

rRNS: riboszomális RNS

rDNS: riboszomális RNS gén

tRNS: transzfer RNS

V1-V9: a 16S rRNS gén variábilis régiói

1. Bevezető

1.1. A mikrobiom-kutatás jelentősége

Az, hogy az egészséges bélrendszer a teljes szervezet egészsége szempontjából kulcsfontosságú, már az ókori orvoslás számára is ismert volt. A belek egészsége és a szervezet egészsége közötti kapcsolat megértésében azonban óriási előrelépést jelentett a 21. században elterjedt mikrobiom-tudomány.

Szervezetünk szuperorganizmusnak tekinthető, hiszen elképesztő mennyiségű mikrobával élünk együtt, melyek gyakorlatilag minden testfelületünkön megtalálhatók. Bár számos gomba, egysejtű eukarióta és vírus is él velünk, jelenleg a szimbióta mikroorganizmusok közül a prokariótáknak tulajdonítja a legnagyobb jelentőséget a tudomány. A velünk együtt élő bakteriális és archea sejtek számát $3,8 \cdot 10^{13}$ -ra becsülik - ez a szám a tulajdon emberi sejteink számához hasonló (Sender és mtsai. 2016). Ezek a mikrobák a gazdaszervezet élettani folyamatait jelentősen képesek befolyásolni, így különböző fiziológiás állapotok, vagy akár betegségek kialakulásában és fennállásában lehet szerepük.

A legtöbb szimbióta mikroba a bélcsatornában, elsősorban a vastagbélben él - itt akár 10^{11} sejt/ml koncentrációban (Sender és mtsai. 2016). A bélbaktériumoknak ismert biológiai szerepük van a bevitt tápanyagok feldolgozásában és biokonverziójában (Braune, Blaut 2016), a tápanyaggal bevitt energia elérhetőségének szabályozásában, bioaktív anyagok, pl. rövid szénláncú zsírsavak és vitaminok szintézisében (Indira és mtsai. 2019). A bélhám integritásának szabályozásával, illetve az intesztinális immunrendszer, a nyálkahártya-asszociált limfoid szövet folyamatos stimulálása révén a szervezet védelmi funkcióit befolyásolják (Chelakkot és mtsai. 2018, Venegas és mtsai. 2019), emellett az idegi és mikrobiális folyamatok szoros összefonódására is számos bizonyíték utal (Appleton 2018)

Alapvetően a fenti útvonalak valamelyikén keresztül hatva képes az intesztinális mikrobaközösség meghatározó szerepet vállalni a gazda egészségi állapotában, ezen útvonalak megértése pedig nem egyszerűen új diagnosztikai és terápiás innovációkhoz kövezi ki az utat, de a betegség és egészség, teljesítmény és állapot fogalmainak újraértelmezésére is lehetőséget ad. Egy egészségtelen, alacsony diverzitású, a környezeti behatásokra érzékenyebb, szisztémikus gyulladásra és krónikus betegségekre hajlamosító mikrobiom diagnózisa, majd egy, a teljes közösség stabilizálását célzó

terápia alkalmazása nagy prevenció értékkel bírhat, illetve már folyamatban lévő terápiák sikerét növelheti (Hitch és mtsai. 2022). Ha felismerjük, milyen konkrét mikrobiális mintázatok állhatnak egyes krónikus betegségek, fogyási nehézség, gyengülő sport-és mentális teljesítmény vagy akár pszichiátriai kórképek mögött, a mikrobaközösség precíziós módosítása is lehetségessé válhat, akár egy vagy néhány törzs feldúsítását vagy kiszorítását célozva (Sorbara, Pamer 2022). Így tehát a mikrobiomot és annak a szervezettel való kölcsönhatásait leíró alapkutatás, illetve a mikrobiomban rejlő diagnosztikai és terápiás lehetőségeket vizsgáló alkalmazott kutatás eredményeire a következő évtizedekben olyan egészségügyi innovációk épülhetnek, melyek a betegellátás új, prevencióközpontú fejezetét támogatják (Dietert RR, Dietert JM 2015). A mélyreható és korai diagnosztika, illetve a rendszerbiológiai elvek integrálása által nagyban csökkenthetőek az egészségügyi ellátórendszer kiadásai, a korán megkezdett, személyre szabott és precíziós terápiák segítségével pedig a kezelések lehetnek sokkal eredményesebbek (Weston, Hood 2004).

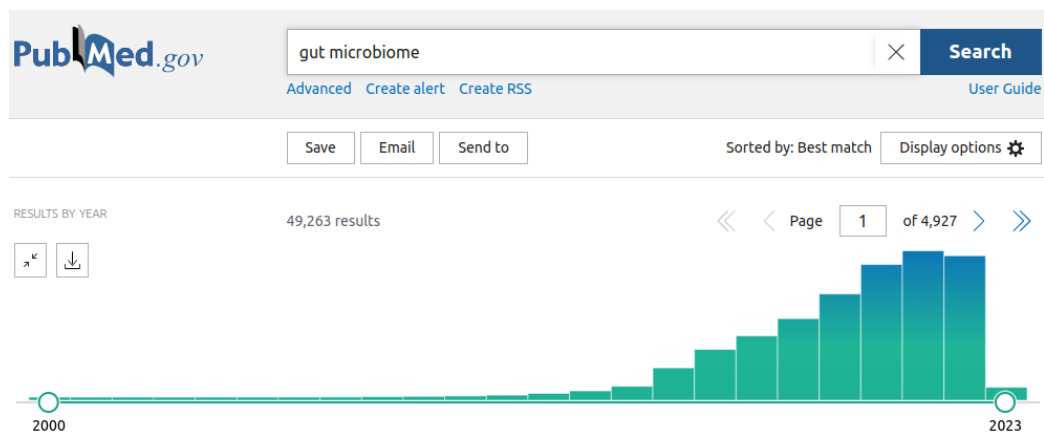
Emellett nem lehet elhanyagolni, hogy az emberi test mellett minden állati és növényi szervezet, talaj, víz és légtér rendelkezik a rá jellemző mikrobaközösséggel, beleértve a gazdasági jelentőséggel bíró haszonállat-és növényfajokat, mezőgazdasági művelés alatt álló talajokat stb. A mikrobiom-kutatás ismeretanyaga és módszertani tárháza így számtalan, gyakran nagy gazdasági jelentőségű területen alkalmazható a biotechnológiától a környezetvédelmen és a mezőgazdaságon át akár a járványügyig - például a szúnyogok által terjesztett humánpatogén vírusfertőzések biokontrolljának egy új módja a szúnyog mikrobiom dúsítása olyan baktériumokkal, amelyek megnehezítik a vírusok transzmisszióját (Huang és mtsai. 2020).

1.2. A bél mikrobiom-kutatás jelenlegi kihívásai

Ahhoz azonban, hogy a fenti alkalmazásokhoz eljusson a tudomány, még rengeteg munka áll a kutatók előtt. A modern mikrobiomika fiatal tudományág, csak az elmúlt években jött igazán lendületbe a molekuláris biológiai és szekvenálási technikák fejlődésének köszönhetően (NIH Human Microbiome Portfolio Analysis Team, 2019) (1. ábra).

A korábbi évek mikrobiom-kutatói a klasszikus mikrobiológia módszereivel, tenyésztésen és morfológiai jellemzők alapján történő osztályozással dolgoztak, ez azonban korlátozott lehetőségeket kínált, mivel a bél mikrobiom legtöbb tagja anaerob és nehezen tenyésztendő (Andrade és mtsai. 2020). Manapság elsősorban masszív paralell szekvenálást használunk a különféle mikrobiomok vizsgálatára; ezek a módszerek a mikrobákból izolált nukleinsavakkal

dolgoznak, és molekuláris biológiai markerek alapján végeznek taxonómiai és funkcionális osztályozást.



1. ábra: a "gut microbiome" kulcsszavakra kiadott PubMed-cikkek száma 2000 és 2023 között. Jól kivehető, hogy a tudományterület exponenciális növekedésnek indult a 2010-es évek közepén

Számos protokoll létezik a mikrobiális szervezetek DNS-ének és RNS-ének nagy tisztaságú izolálására, szekvenálására, illetve a szekvenciák bioinformatikai elemzésére. Egy mikrobiom-minta DNS- és RNS-állományának elemzésével a taxonómiai összetételtől a transzkriptomikai és metabolikus folyamatokig számos adatra következtethetünk. Az adatokig elvezető munkafolyamat azonban számos molekuláris biológiai és bioinformatikai lépés sorozata, melyek megközelítése egyenként eltérő lehet.

Az eltérő megközelítések eltérő eredményekhez is vezethetnek, a mikrobiomikai eredmények alkalmazhatóságának pedig az a kulcsa, hogy olyan standard módszerekkel rendelkezünk, amelyek egy adott célkitűzést nagy megbízhatósággal teljesítenek (Fiedorová és mtsai. 2019). Diagnosztikai és terápiás megközelítéseket hiba lehet olyan munkafolyamatokra építeni, amelyek a nem megfelelően megválasztott, vagy nem kellően ismert módszerek révén torzítják az elvégzett vizsgálatok eredményét, hiszen egy bizonytalan lábakon álló terápia kisebb eséllyel lehet sikeres, ugyanakkor nagyobb kockázatot rejt magában. A terjedőben lévő új módszerek kezdeti sikerei különösen fontosak a kutatói és orvosi közösség, a további kutatások potenciális támogatói és a leendő páciensek bizalmának elnyerésében, ami biztosítja a dinamikus fejlődést.

Természetesen, mivel a mikrobiális metagenomi minták eltérő feldolgozási lehetőségei részben azt a célt szolgálják, hogy különböző jellegű adatokhoz juthassunk, nem elvárható egy olyan egységes és validált protokoll kidolgozása, amely mintavételezéstől diagnózisig lépésről lépésre követhetően

leírja egy minta feldolgozási útvonalát, és amely minden lehetséges helyzetben alkalmazható, kezdve egy elhízott páciens általános mikrobiom-összetételének vizsgálatától egy longitudinális egérkísérlet két modellállatának metatranszkriptomikai elemzéséig. Elengedhetetlen azonban, hogy az egyes módszerek buktatói, esetleges torzításai, adott helyzet függvényében változó előnyei és hátrányai alaposan fel legyenek térképezve és ezek az adatok elérhetőek legyenek a tudományos közösség számára. Ez az alapja annak, hogy egy adott kísérleti, klinikai vagy ipari munkafolyamat tervezésekor magas színvonalú szakmai döntés alapján lehessen megválasztani a módszereket, és a valósághoz a lehető legközelebb álló eredményeket tudjunk nyerni (Galloway-Peña, Hanson 2020).

1.2.1. Módszertani eredetű adattorzítások széketminta feldolgozása során

Kutatócsoportunk munkája a mikrobiom étrendi szénhidráttartalomtól való függése, elhízásban és annak társbetegségeiben betöltött szerepe körül összpontosul. Szeretnénk hozzájárulni annak megértéséhez, hogy a bélflóra hogyan változik a különböző élettani állapotok és étrendek függvényében, ezért egészséges elhízott, inzulinrezisztens és kettes típusú diabéteszes betegek széketében megtalálható baktériumokat, gombákat és vírusokat vizsgálunk genom és transzkriptom szintjén, mialatt a betegek egységesen alacsony szénhidráttartalmú, vagy személyre szabottan összeállított diéta-és mozgásprogramokat követnek. Ugyanakkor, mivel emberi páciensek esetén jelentős eltérések merülhetnek fel a genetikai háttér és az életmód terén, illetve a diéta szigorú tartása sem garantálható minden esetben, indokolt olyan modellorganizmusok használata, melyek tartási körülményeit nagyobb mértékben lehet kontrollálni a fenti változókra nézve. Laboratóriumunkban a humán mikrobiom jelenleg ismert legjobb modellszervezetével (Coelho és mtsai. 2018), fajtatiszta, ismert rokonsági viszonyú, különböző szénhidráttartalmú tápokon nevelt kutyákkal, illetve mintáikkal végzünk a humán csoporttal párhuzamos vizsgálatokat. Emellett anaerob tenyésztési körülmények között lezajló baktérium-baktérium kölcsönhatásokat is kutatunk transzkriptomikai, illetve epigenetikai alapú megközelítésekkel. Munkánk hosszútávú célja a mikrobiom működésének és a metabolikus szindrómával összefüggő betegségek patomechanizmusának jobb megértése, így potenciális diagnosztikai és terápiás alkalmazások megalapozása.

A kutatás egyik legfontosabb eleme a vizsgált minták taxonómiai összetételének megállapítása (Banerjee és mtsai. 2018). A különböző paraméterekkel rendelkező páciensek mintáinak összetételét tudva összefüggések ismerhetők fel egyes állapotok és egyes mikrobiom-mintázatok között, ezekre pedig további következtetések és kísérletek alapozhatók a mintákban talált baktériumtörzsek már megismert jellemzőit, anyagcseréjét, biológiai szerepét figyelembe véve. Már

a kutatás korai fázisában rá kellett jönnünk viszont, hogy nincsenek egységes és kielégítő szakmai ajánlások azokkal a munkafolyamatokkal kapcsolatban, amelyekre a kutatásunk épül. A minták feldolgozása során minden lépésnél kulcsfontosságú, hogy milyen módszert (melyik DNS-izoláló kité, szekvenálási megközelítést, célzott amplikon szekvenálás esetén melyik riboszomális V régiót, adatfeldolgozás során melyik bioinformatikai programot stb.) választanak a kutatók, mert, például, a helytelenül megválasztott izoláló kit a Gram-pozitív és Gram-negatív fajok sejtfalát eltérő hatékonysággal lizálhatja (de Boer és mtsai. 2010), a nem megfelelően megválasztott taxonómiai marker génrégió pedig vezethet például ahhoz, hogy rokon, de nem azonos fajokat nem sikerül megkülönböztetni (Chakravorty és mtsai. 2017). Bár ilyen és hasonló problémák a legtöbb különböző módszernél fennállhatnak, a jelenlegi ismereteink szerint bizonytalan, hogy melyek azok a megközelítések, amelyek csak a lehető legkisebb torzítást adják a minták taxonösszetételében.

Ezzel a problémával szembesülve megterveztünk egy kísérletsorozatot, amelyben a rendelkezésünkre álló legfontosabb DNS-izoláló protokollokat, szekvenálási megközelítéseket és bioinformatikai eszközöket futtattuk le ugyanazon a kutya székletmintán.

1.3. A bél mikrobiom-kutatás alkalmazott módszerei

1.3.1. DNS-izolálás székletmintából

A mikrobiom-vizsgálati céllal gyűjtött székletminta összetételének csupán töredékét teszik ki a keresett nukleinsavak. Ezek elválasztása a minta és a bakteriális sejtek egyéb összetevőitől bonyolult feladat, melynek során több lépésben, fizikai és kémiai módszerekkel tisztítható a vizsgálni kívánt DNS. A bonyolult izolációs folyamatok elvégzésére ma elsősorban kereskedelmi forgalomban kapható kitéket használunk, melyek tartalmazzák a folyamat elvégzéséhez szükséges eszközöket és reagenseket, valamint az azok használatára vonatkozó protokollokat.

A különböző cégek által fejlesztett kiték megközelítése eltérő lehet, ez pedig befolyásolhatja a kinyerendő DNS mennyiségét és minőségét. Mivel a székletminta egy komplex biológiai mátrix, számos tényező van potenciális hatással a kivont DNS további analízise során elérhető eredményekre. A széklet jelentős mennyiségben tartalmaz a polimeráz láncreakciót gátló ún. PCR-inhibitorokat, például epesavas sókat, hemoglobin-bomlástermékeket és komplex poliszacharidokat (Monteiro és mtsai. 1997, Flekna és mtsai. 2007, Nechvatal és mtsai. 2008). Mivel a mikrobiom-vizsgálat DNS-izolálást követő lépése gyakran, így jelen kutatás során is -

elsősorban - molekuláris markergének PCR-amplifikációja, a reakciót negatív irányban befolyásoló inhibitorokat a mintából el kell távolítani.

A székletmintában fellelhető baktériumok változatos sejtfalszerkezettel rendelkeznek, a Gram-pozitív fajok sejtfalának lízise nagyobb kihívást jelent. Amennyiben egy extrakciós protokoll a Gram-pozitív és Gram-negatív sejtek feltárását eltérő hatékonysággal végzi, az jelentősen torzítja a kapott adatokat (de Boer és mtsai. 2010, Santiago és mtsai 2014).

A kinyert DNS tisztasága és taxonómiai eredete mellett a mennyisége és az integritása is befolyásolja a downstream elemzést. Az alacsony kinyerés a hamis negatív eredmények arányát növelheti (Salonen és mtsai. 2010), azaz csökkenti a ritka taxonok detektálásának az esélyét, valamint a PCR-amplifikációs hibák esélyét is növeli (Akbari és mtsai. 2005). Egy ideális kit tehát képes nagy mennyiségű és integritású, inhibitoroktól mentes és a minta összetételét jól reprezentáló DNS-minta izolálására.

Kutatásaink során négy eltérő DNS-izoláló kitet teszteltük le, a Qiagen, az Invitrogen, a Macherey-Nagel és a ZymoResearch gyártók termékeit. A protokollokban közös, hogy a -80°C-on tárolt mintákat olvasztást követően sejtlízis, fehérjebontás és centrifugálás módszerekkel kezelik, feltárva a sejteket, és elválasztva a DNS-t a sejtalkotóktól. Az egyes lépések, reagensek, sejtfeltárási és DNS-elválasztási módszerek terén azonban a kitek különböznek egymástól.

A **Qiagen QIAmp Fast DNA Stool Mini Kit (Qiagen, Németország)** egy komponense, az InhibitEx puffer alkalmas arra, hogy a mintákból eltávolítsa a PCR-inhibitorokat. A DNS tisztítására a Qiagen kit Spin Column oszlopot használ, amely a centrifugálás során köti a nukleinsavakat, miközben a szennyeződések átáramolnak rajta. A folyamat végén a DNS az oszlopról eluálható.

Az **Invitrogen PureLink™ Microbiome DNA Purification Kit (Thermo Fisher Scientific, USA)** a sejtfeltárás során gyöngyöket alkalmaz, melyek a fizikai sejtfeltárást nagymértékben segítik. A pufferes lízis ugyanis nem feltétlenül képes ugyanolyan hatékonysággal feltárni a Gram-pozitív sejtek sejtfalát, mint a Gram-negatívokét (de Boer és mtsai. 2010). A DNS tisztítása során az Invitrogen kit is alkalmaz Spin Column oszlopot.

A **Macherey-Nagel NucleoSpin DNA Stool Mini Kit (Macherey-Nagel, Németország)** szintén alkalmaz gyöngyöket a fizikai sejtfeltárási megkönnyítésére. A lízist követően a minta Inhibitor Removal Column-ra kerül, amely a Qiagen InhibitEx pufferhez hasonlóan a PCR-inhibitorok

eltávolítására szolgál. A DNS tisztítása itt is oszlopos megoldással történik, a NucleoSpin DNA Stool Column felhasználásával, majd a DNS elúciójával az oszlopról.

A **ZymoResearch Quick-DNA™ HMW MagBead Kit (Zymo Research, USA)** a mintát a gyártó által fejlesztett DNA/RNA Shield-ben szuszpendálja, amely a nukleinsavak integritását biztosító folyadék. A protokoll kezdeti centrifugálási lépései során keletkező pelletből többszöri reszuszpendálás és centrifugálás során igyekszünk kinyerni a lehető legtöbb DNS-t, a Gram-pozitív sejtek feltárását pedig lizozim enzimmal segítjük. A DNS-t MagBinding Beads mágneses gyöngyökkel kötjük, amelyek a hozzájuk kötött DNS-sel együtt mágnes segítségével könnyen eltávolíthatók az oldatból. Az izolálás végén a DNS-t a mágneses gyöngyökről eluáljuk.

1.3.2. Szekvenálási megközelítések

1.3.2.1. 16S rRNS V régió amplikonok

A prokarióták molekuláris biológia alapú taxonómiai azonosításának fő módszere a 16S rRNS gén szekvenciáján alapuló azonosítás. A körülbelül 1600 bázispár hosszú gént evolúciósan konzervált és hipervariábilis, közeli rokon fajok esetén is eltérő szekvenciájú szakaszok építik fel (2. ábra). Az azonosítási módszer alapja, hogy a konzervatív régiókat univerzális PCR-primerekkel megcélózva a köztük lévő, tetszőleges számú variábilis (V) régiók felszorzozhatók, megszekvenálhatók, majd a szekvenciák egy adatbázis alapján rendelhetőek adott fajokhoz.



2. ábra: a 16s rRNS V régiói. **Konzervatív régiók,** variábilis régiók

Ugyanakkor az összesen 9, V1-V9-ként jelölt variábilis régió eltérő helyen található és eltérő funkcióval bír a riboszomális kis alegység RNS-komponensében, ezért a mutációs gyakoriságuk, variabilitásuk sem teljesen azonos (Bukin és mtsai. 2019, Schuwirth és mtsai. 2005, Schlutzen és mtsai. 2000). A kulcsfontosságú biológiai szerepet betöltő, a tRNS-ekkel és a 23S rRNS-sel interaktáló V4, V5 és V6 régiók a szerepükből adódóan kevésbé mutálódhatnak. Ezzel szemben a strukturális szerepű V2 és V8 régiók mutációs hajlandósága igen magas, mivel ezeknek a szekvenciája nincs közvetlen összefüggésben biológiai feladatukkal. A ma még kevésbé ismert funkciójú V3 és V7 régiók közepes mutációs hajlandóságot mutatnak.

Mivel a különböző V régiók nem ugyanazzal az evolúciós sebességgel halmoznak fel mutációkat, kérdés, hogy a taxonómiai azonosításra mindegyikük ugyanannyira alkalmas-e. Ezzel a

problémával több tanulmány is foglalkozott (Chakravorty és mtsai. 2007, Bukin és mtsai. 2019, Zhang és mtsai. 2018, Kerrigan és mtsai. 2019), a tanulmányok eredményei pedig arra mutatnak, hogy az azonosításra választott V régió befolyásolja a vizsgált minták taxonösszetételéről kapott eredményeket. Chakravorty és mtsai. (2007) eredményei alapján régiófüggő, hogy a taxonómiai azonosítás milyen mélységig sikeres. Minél konzervatívabb egy V régió (V4, V5, V6), annál hasonlóbb a közeli rokon fajok esetében, ezért inkább a magasabb, törzs és osztály szintű azonosításra alkalmas. A nagyobb mutációs hajlandóságú (V2, V8) régiók szekvenciája alapján biztosabban különböztethetők meg a közeli rokon fajok. Egyértelmű tudományos konszenzus nincs arra vonatkozóan, hogy melyik V régió a legideálisabb egy baktériumközösség taxonómiai feltérképezésére, bár Zhang és mtsai. 2018-ban hangsúlyozták, hogy a közeljövőben a teljes 16S rRNS gén hosszú leolvasású szekvenálása kell, hogy a preferált módszer legyen. Ugyanakkor - jelenleg - a hosszú leolvasású szekvenálások költsége lényegesen nagyobb, és a módszerek pontossága is kevésbé megbízható (Amarasinghe és mtsai. 2020). Ennek megfelelően jelen kutatásban - a hosszú leolvasások mellett természetesen - több V régió amplikont vizsgáltunk, melyek jellemzően 2 variábilis régiót tartalmaznak. Rövid leolvasású Illumina szekvenálással a 16S rRNS gén V1-V2, V1-V3, V3-V4 (illetve korábbi munkák során a V4) régiók amplikonjait szekvenáltuk, a teljes génről készült amplikonok leolvasását Oxford Nanopore Technologies (ONT) MinION és Pacific Biosciences (PacBio) Sequel IIe long-read platformokkal végeztük.

A V1-V2 amplikonok körülbelül 300 bázispár hosszúak, és tartalmazzák a V1, illetve a magas variabilitású V2 régiókat.

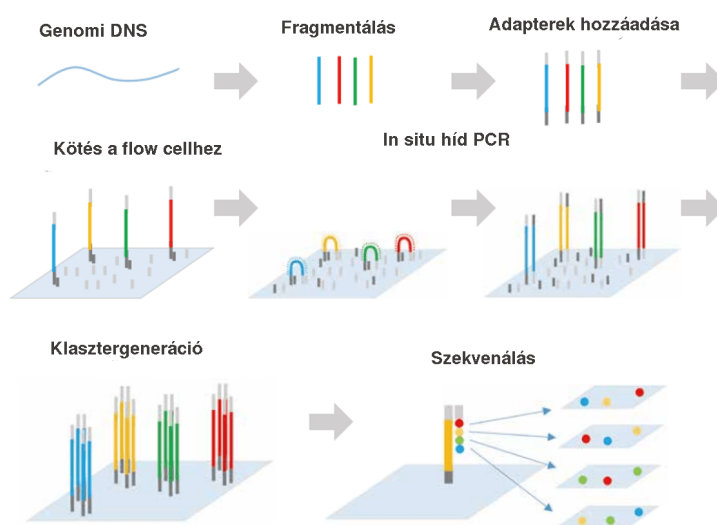
A V1-V3 amplikon hossza 600 bázispár körüli, és 3 variábilis régiót tartalmaz. Ezek közül a V2 variabilitása kimondottan nagy, a V3-é közepes, ami elméletileg ideális jelöltté teszi a nagyobb felbontású taxonómiai azonosítási alkalmazásokra. Ugyanakkor a V1-V3-ról fontos kiemelni, hogy a hossza az Illumina szekvenátorok leolvasási kapacitásának felső határán mozog; munkánk során jellemzően azt tapasztaltuk, hogy a paired-end szekvenálás és az ezt követő bioinformatikai szűrés során a régió középső, alapvetően a V2 régió terminális és az azt követő konzervatív régió iniciális szakaszára kiterjedő része elvész vagy alacsony minőségben kerül leolvasásra. Bár a proximális V1 és V3 szakaszok épsége megfelelő alapot biztosíthat a taxonómiai azonosításra, a V2 szekvencia technikai eredetű minőségcsökkenése befolyásolhatja az eredményeket.

A 400 bázispár körüli, egy magasabb és egy közepes konzerváltságú V régiót tartalmazó V3-V4 amplikon feltehetően megbízhatóan szerepel a magasabb rendszertani kategóriák azonosításában, de az alacsonyabb taxonómiai szintekre nem ideális. Korábbi, TDK-pályamunkaként is publikált

kutatásaink szerint, melyben a V1-V3 és V4 régiókat hasonlítottuk össze a jelenlegitől eltérő kísérleti felállásban, azt találtuk, hogy a V4 régió ampliconokkal több törzset és osztályt, ám kevesebb családot és nemzetséget tudunk azonosítani, mint V1-V3 szekvenálással (Henkrich, Járay, SZTE TTIK 2022). Az Illumina által 16S amplicon szekvenálásra hivatalosan is ajánlott génrégió a V3-V4 szakasz, így számos publikáció alapul erre a megközelítésre. Ugyanakkor egy japán populáción végzett 2021-es vizsgálat (Kameoka és mtsai. 2021), amely a V1-V2 és a V3-V4 régiók alapján megállapított mikrobiom-összetétel adatokat hasonlította össze, azt találta, hogy a qPCR-esszék alapján valószínűleg megállapított adatokhoz képest a V3-V4 amplicon szekvenálás hajlamos a *Bifidobacterium*-abundanciák alul- és az *Akkermansia*-adatok túlbecslésére. A diabétesz-asszociált *Akkermansia* (Hasani és mtsai. 2021) pontos felmérése ilyen irányultságú kutatásainkban különösen fontos.

1.3.2.2. Szekvenálási platformok és leolvasási hossz

Jelen projektben a V1-V2, V1-V3, V3-V4 régiókat tartalmazó ampliconokat rövid leolvasású Illumina szekvenálással, a V1-V9 régiókat tartalmazó teljes gén hosszú leolvasását pedig ONT MinIon és PacBio Sequel IIe platformon párhuzamosan végeztük.



3. ábra: az Illumina rövid leolvasású szekvenátorok működési elve

A **rövid leolvasású szekvenálás**, így az Illumina szekvenátorok (Ravi, Walton, Khosroheidari 2018) működési elve szerint a szekvenálni kívánt nukleinsav-szállal komplementer szál szintézise zajlik fluoreszcensen jelölt nukleotidokkal, egy akrilamiddal burkolt üveg flow cell felületéhez rögzítetten (3. ábra). Az egyes nukleotidok beépülését színes fényjelenség kíséri, melyet a szekvenátor kamerája detektál, és mivel a flow cell egy adott pontján nem csupán egy DNS-szál,

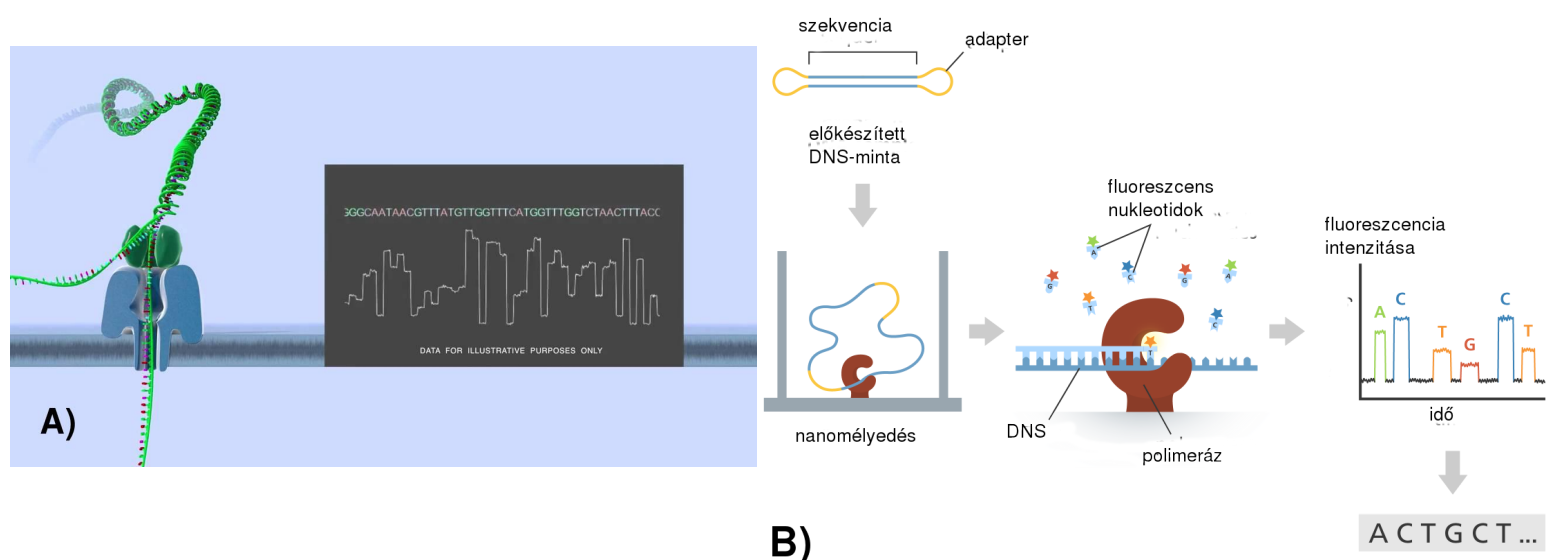
hanem egy kiindulási molekulából in situ PCR által létrehozott klaszter található, a fényjelenség amplifikáltan jelentkezik, csökkentve a hibás beépülések detektálásának esélyét. Az Illumina szekvenátorok nagy pontosságú, ugyanakkor rövid, két oldalról történő (paired-end) szekvenálás esetén is legfeljebb 600 bázispár körüli szekvenciák leolvasására képesek. Így a 16S rRNS gén teljes hosszában nem, csak egyes V régiók szintjén olvasható le ezzel a megközelítéssel. A jelenlegi ismereteink szerint azonban, mivel a különböző 16S génrégiók eltérő taxonómiai besorolást eredményezhetnek, ideális lenne hosszú leolvasással, a teljes génrégió vizsgálatával végezni az azonosítást (Zhang és mtsai. 2018).

A 16S V régió amplikon szekvenálás mellett ugyanakkor **shotgun metagenomi szekvenálás** is végezhető rövid leolvasású megközelítéssel, melynek során a célzott amplikonok helyett a teljes izolált DNS-pool fragmentálása és a fragmentek szekvenálása zajlik. Ez a megközelítés számos taxonómiai azonosításra alkalmas génhez és génszakaszhoz enged hozzáférést, emellett a metagenom funkcionális annotációját is lehetővé teszi. Mivel azonban a teljes metagenomot vizsgáló shotgun során leolvasott egyedi szekvenciák száma nagyságrendekkel nagyobb, mint amplikon szekvenálás esetén, így a módszer erőforrásigénye az anyagiak és a számítási kapacitás terén is jelentősebb.

A **hosszú leolvasású (long-read) szekvenálás** lehetővé teszi akár több tízezer nukleotid, így természetesen a teljes 16S rRNS génszakasz célzott és egyben történő leolvasását.

Az ONT MinION hosszú leolvasású szekvenátorának flow celljében nanopórusos membrán található: a pórusok olyan fehérjestruktúrák, melyeken a szekvenált DNS-molekula - egyszálú alakban - képes áthaladni (Jain és mtsai. 2016) (4.A. ábra). Az egyes bázisok áthaladása a membránon feszültségváltozást generál a membrán két oldalán, amelynek mértéke nukleinsav-bázisonként eltér, és a feszültségváltozások mérésével meghatározható a membránon áthaladó nukleinsav-szál szekvenciája. A nanopore szekvenálás során klasztergeneráció nem történik, így legfeljebb annyi leolvasás keletkezik, ahány nukleinsav-molekula a flow cellre felkerült. Bár a módszer által generált leolvasások hossza az egyéb megközelítésekhez képest drámaian nagyobb, és a metilációs mintázat detektálására is alkalmas, a megközelítés a short-read szekvenálásnál kevésbé pontos (Zhang és mtsai. 2020).

Az Illumina szekvenáláshoz hasonlóan szintézisalapú megközelítést alkalmaznak a Pacific Biosciences szekvenátorok, melyek képesek az akár 25 kilobázis hosszú leolvasások generálására (Rhoads, Au 2015) (4.B. ábra).



4. ábra: a hosszú leolvasású szekvenátorok működési elve. A: ONT nanopore szekvenálás, B: PacBio SMRT szekvenálás

A PacBio szekvenátorok flow cellje, a SMRT Cell mélyedéseiben (zero-mode waveguides, ZMW) rögzített DNS-polimeráz cirkularizált egyszálú DNS-molekulákhoz szintetizál komplementer szálát fluoreszcensen jelölt nukleotidokból. Egy templátról több ciklusban írható át komplementer szál, így számos ciklus lefutását követően nagy bizonyossággal állapítható meg a vizsgált nukleinsav szekvenciája. A PacBio Sequel IIe emellett a DNS-bázisok epigenetikai módosításainak, pl. metilációk detektálására is alkalmas. Ugyanakkor a magas költség, illetve az Illuminánál továbbra is alacsonyabb pontosság (Zhang és mtsai. 2020) nem feltétlenül teszik a legideálisabb szekvenátorrá. Egy kutatás, mely a PacBio és Illumina szekvenátorokat kombináltan alkalmazta, a módszert kimondottan hatékony megközelítésnek ítélte metagenomi adatok kinyerésében (Xu és mtsai. 2022).

Összefoglalva tehát, a 16S rRNS gént több szekvenálási megközelítéssel is vizsgálhatjuk. A gén egyes hipervariábilis szakaszairól Illumina rövid leolvasású amplikon szekvenálással nyerhetünk információt. Itt a szekvenciák nagy pontossággal - platform-függően, de jellemzően jóval 1%-os hibaarány alatt - és költséghatékonyan olvashatóak le, mivel azonban nem a teljes génszakasszal dolgozunk, az ezek alapján végzett taxonómiai azonosítás kevésbé megbízható, a módszerek a pontos fajszintű azonosításra nem ideálisak. A teljes 16S rRNS gén szekvenciája alapján megbízhatóbb azonosítást végezhetünk, és a szakértői vélemények is ennek a módszernek az alkalmazását vetítik elő és hangsúlyozzák a közeljövőre nézve. A long-read alapú módszerek képesek a teljes gén egyben történő leolvasására, így a megbízhatóbb azonosításra, ugyanakkor akár

10-15%-os is lehet a hibaarány (Zhang és mtsai. 2020; bár megjegyzendő, hogy a PacBio ún. circular consensus, azaz egy cirkuláris könyvtárról többször leolvasott és egymásra illesztett szekvenciái végül az Illuminával közel megegyező minőséget eredményeznek), és a módszerek költsége is magasabb. Shotgun metagenomi szekvenálással számos különböző taxonómiai azonosításra használható genomi régióhoz, továbbá vírusok, gombák elemzéséhez férhetünk hozzá, a módszer viszont anyagi és adatfeldolgozási ráfordítás szempontjából is megterhelőbb a 16S amplikon szekvenálásnál. A felsorolt, egymással több paraméter tekintetében is eltérő szekvenálási platformok mikrobiom-összetételi vizsgálatokra való alkalmasságát, az egyes alkalmazások helyzettől és kutatási céloktól függő indokoltságát ezek összehasonlítása segíthet jobban megérteni.

1.3.3. Bioinformatika

A szekvenálások eredményeként kinyert szekvenciák további elemzésre szorulnak ahhoz, hogy biológiailag releváns adatokként lehessen értelmezni őket. Egyes esetekben, pl. az ONT szekvenálások esetén el kell végezni a bázishívást, azaz a mért leolvasási adatok szöveges fastq fájlra konvertálását. A szöveggént tárolt szekvenciákat minőségük szerint szűrni kell, a rossz minőségű szakaszokat és az adapterszekvenciákat ki kell vágni. Illumina paired-end szekvenálásnál szükséges az összetartozó szekvenciák egymáshoz illesztése. A megfelelően feldolgozott szekvenciákat hozzá kell rendelni a megfelelő taxonhoz, referencia adatbázisok adatai alapján. Végül pedig vizuálisan értelmezhetővé kell tenni a jellemzően szöveges vagy táblázatos formátumú eredményeket, hogy értelmezni tudjuk azokat. Mindezekhez a munkafolyamatokhoz szükség van a megfelelő bioinformatikai szoftverekre. Számos szoftver elérhető metagenomi szekvenciák elemzésére, melyek a tudomány és technológia előrehaladtával felmerülő különböző problémákra kínálnak megoldást.

1.3.3.1. DADA2

A DADA2 (Callahan, McMurdie, Holmes 2017) egy innovatív megközelítésű, R alapú szoftver az Illumina amplikon szekvenálások kiértékelésére. A program újszerű megközelítése abban rejlik, hogy újraértelmezi a molekuláris rendszertan alapfogalmait. A korábban alkalmazott módszerek műveleti alapegysége az OTU, azaz az operációs taxonómiai egység volt. Az OTU-k egymással, vagy egy referencia adatbázis egy elemével megfelelő mértékű - általában 97% fölötti - egyezést mutató szekvenciák klaszterei. Az OTU-alapú programok tehát először csoportosítják a mintában talált hasonló szekvenciákat, majd a csoportokon végzik a további műveleteket. Ennek a megközelítésnek az a hátránya, hogy az egymással nagy hasonlóságot mutató, de nem teljesen

azonos szekvenciák, például két közeli rokon faj marker génszekvenciái könnyen kerülhetnek azonos OTU-ba, így a közeli rokon fajok nehezen megkülönböztethetők.

A DADA2 ugyanakkor ASV-kkel, azaz amplikon szekvencia variánsokkal dolgozik, azaz minden egyes egyedi leolvasást összevet a használt referencia adatbázissal: a közeli rokon fajokat nem veszi egy kalap alá, a be nem azonosított taxonokat pedig meghagyja az outputban, lehetőséget adva azok későbbi azonosítására, akár új fajok klasszifikációjára. Mivel az ASV-ket egyedileg kezeli a program, az egyes leolvasások számából jól kirajzolódik azok abundanciaértéke. A DADA2 emellett hibamodellt futtat a leolvasott szekvenciákon, amelynek célja a nagyon hasonló szekvenciákban található különbségek okának a megállapítása. A hibamodell felismeri, hogy egy kismértékű eltérés szekvenálási hibának tudható be, vagy valódi biológiai jelnek tekinthető, így a fals pozitív eredmények aránya is kisebb az outputban.

Kutatócsoportunk a DADA2-t a SILVA 16S rRNS génbankkal használta referenciaként. A DADA2 által generált output fájlok táblázatos formában tartalmazzák a szekvenálási adatokat, a könnyebb értelmezhetőség kedvéért fontos ezek vizualizációja. Erre a Phyloseq programot használtuk.

1.3.3.2. Emu

Az Emu (Curry és mtsai. 2022) a magas hibaarányú hosszú leolvasások fajszintű azonosítására lett kifejlesztve. A program két lépésben dolgozik; a leolvasásokat először referencia adatbázishoz hasonlítja, majd egy várható érték-maximalizáló (EM) algoritmuson alapuló hibakorrekció során finomítja a kalkulált fajszintű relatív abundanciaértékeket. Az Emu a hasonló programokhoz képest jóval kisebb arányban ad fals pozitív eredményeket, és kiváló a közeli rokon taxonok megkülönböztetésében. Ugyanakkor az algoritmus működési elveinek következtében az elfogadott abundanciaérték alsó határa 1000-ból 1, 1000 fölött 10 read egy fajtól, így az Emu a nagyon kis abundanciájú taxonok detektálására alkalmatlan. A program további hátránya a magas számítási-és időigény, illetve a zárt referencia: az Emu, szemben például a DADA2-vel, csak a használt referencia adatbázisban már szereplő taxonok azonosítására képes. Ez bizonyos tekintetben limitálja az alkalmazhatósági körét, bár az új fejlesztésű program jövőbeli fejlesztései átléphetik a jelenlegi korlátokat.

Csoportunk az Emu-t a shotgun leolvasások kivételével minden szekvenálás elemzésére használta.

1.3.3.3. EPI2ME

Az EPI2ME (Oxford Nanopore Technologies) egy felhőalapú, integrált és könnyen kezelhető platform long-read szekvenciák elemzésére. A szoftver az NCBI BLAST adatbázist használja a 16S rRNS szekvenciák referenciájaként, amely egy több, mint 22 ezer szekvenciát tartalmazó, folyamatosan fejlődő adatbázis. Az EPI2ME outputja vizuális, további bioinformatikai feldolgozást nem igényel. A programot az ONT MinION 16S hosszú leolvasások azonosítására használtuk.

1.3.3.4. Kaiju

A Kaiju (Menzel 2016) shotgun metagenomi szekvenciákon alapuló taxonómiai azonosításra lett kidolgozva. A fejlesztők célja egy olyan program írása volt, amely a nagy áteresztőképességű metagenomi minták feldolgozását a korábban elérhető programoknál gyorsabban és hatékonyabban végzi, és emellett a kevésbé ismert vagy ismeretlen szekvenciák azonosításában is hatékony. A korábbi metagenom-elemző szoftverek esetében ugyanis jellemzően fennállt egy tradeoff a gyorsaság és pontosság tekintetében. A helyi illesztésen alapuló módszerek pontosan, de lassan, a k-mer alapú megközelítések gyorsan, de kevésbé pontosan dolgoznak. A Kaiju k-merek, azaz fix k hosszúságú rövid szekvenciák pontos illesztésén alapul, így gyors, illetve nem nukleotid-, hanem aminosav-szekvenciákkal dolgozik, ami a pontosságot fokozza.

Mivel a bakteriális, illetve virális genomokban sűrűn helyezkednek el fehérjekódoló gének, az aminosavsorrend szintjén pedig kisebb evolúciós sebességgel halmozódnak fel mutációk, mint az örökítőanyagban, a Kaiju aminosavszekvenciára alapozza a taxonómiai azonosítást. A leolvasott szekvenciákat a hat lehetséges leolvasási keret szerint aminosavsorrendre fordítja, és ezeket a szekvenciákat hasonlítja referencia adatbázisokhoz. A nukleinsavaknál kisebb mutációs gyakoriság miatt a beazonosítatlan szekvenciák is könnyebben azonosíthatóak egy fajnál magasabb taxonómiai szinten.

A leolvasások hossza - ami a DNS-izolálás, a könyvtárkészítés és a szekvenálás minőségétől mind függ - erősen befolyásolja a Kaiju-elemzések sikerét: a rövid nukleinsav-leolvasások nagyon rövid proteinszekvenciákat adnak, melyek növelik a hamis pozitív eredmények arányát azzal, hogy az adatbázis jóval több elemére felilleszthetők, mint egy hosszabb szekvencia. A hosszabb paired-end readok esetén a Kaiju szignifikánsan jobban teljesít, mint a rövidebb single-end leolvasásokon. Emellett a Kaiju nem tud a fehérjét nem kódoló szekvenciák, így a 16S rRNS gén szekvenciája alapján dolgozni, tehát a V régió alapú bioinformatikai programoktól eltérő taxonómiai markereket

használ azonosításra. A különböző munkafolyamatokból származó eredmények összevetése esetén ez eltérések forrása lehet.

1.3.3.5. Minitax

A csoportunkon belül fejlesztett Minitax a minimap2 (Li 2018) output adatainak feldolgozására szolgáló bővítmény. A minimap2 egy térképező szoftver, amely kérdéses szekvenciákat referencia genomokra illeszt, azonban nem kimondottan metagenomikai elemzésekre lett kifejlesztve. A Minitax a minimap2-vel nyert térképezési adatokat dolgozza fel úgy, hogy a több referencia genomra is feltérképezett kérdés szekvenciák között a térképezés minősége alapján rangsort állít, illetve ha ezek minősége azonos, akkor a térképezések legközelebbi közös őséhez rendeli a kérdés szekvenciát. A Minitax munkafolyamat során a proGenomes adatbázist használtuk és az Illumina shotgun szekvenálások térképezését végeztük el.

2. Célkitűzés

A székletmintából izolált nukleinsavak szekvenálásán és bioinformatikai elemzéseken alapuló mikrobiom-analízis számos élettudományi területen került a figyelem középpontjába, ugyanakkor az alkalmazott munkafolyamatok számos szűk keresztmetszetet mutatnak, melyek egy vizsgálat eredményeit jelentősen befolyásolhatják. A mintafeldolgozáskor alkalmazott DNS-izoláló kit, a szekvenálási megközelítés és a megszekvenált genomi régió, illetve a bioinformatikai elemzés során alkalmazott programok, eltérő megközelítésükből adódóan, különböző biológiai és technikai eredetű adattorzításokat eredményezhetnek.

Kutatócsoportunk azzal a céllal, hogy az egyes módszerek által generált torzításokat jobban megismerje és hozzájáruljon a mikrobiom-kutatási metódusok optimalizációjához, egy kutya székletminta - elsősorban - prokarióta metagenomját több módszerrel zajló párhuzamos elemzésnek vetette alá. Ugyanazon mintán leteszteltünk négy különböző DNS-izoláló kitet, hat különböző rövid-és hosszú leolvasású, 16S rRNS amplikon és shotgun metagenom szekvenálási könyvtárat, két hosszú-és egy rövid leolvasású szekvenáló platformot, Illumina szekvenálás esetén kétféle reagenssel, valamint öt bioinformatikai adatfeldolgozó programot. A kapott eredményeket összevetettük egymással, illetve egy ismert összetételű kontroll mikroba mix feldolgozásával nyert adatokkal.

Az összehasonlítások alapján célunk volt azonosítani a munkafolyamat egyes lépéseinél használt alternatívák előnyeit, hátrányait, esetleges adattorzító hatásait. Kutatásunk célja a jelenleg a mikrobiom-kutatásban alkalmazott eszközök és módszerek értékelése, reményeink szerint hozzájárulva a pontos, reprezentatív módszereken alapuló mikrobiomikai kutatás és - elsősorban humán diagnosztikai és terápiás - alkalmazások fejlesztéséhez.

3. Anyagok és módszerek

3.1. Feldolgozott minták

3.1.1. Kutya székletminta

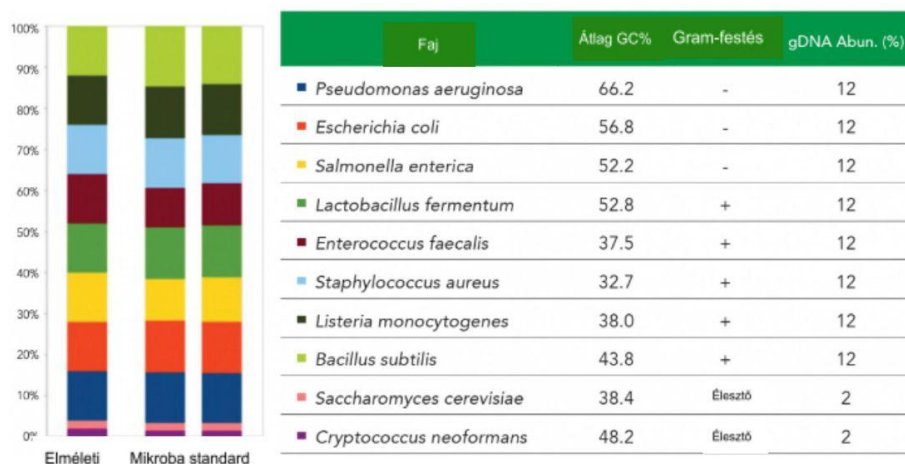
A metodikai elemzés során feldolgozott minta egy 13 éves egészséges kutyától (hím pumi) származott. A kutya bélflóra a humán mikrobiom kiváló modellje, a nagy hasonlóság feltehetően az evolúciós szempontból is jelentős hosszúságú szoros együttélés következménye (Coelho és mtsai. 2018). Kutatócsoportunk, mint azt az 1.2.1. alfejezetben említettem, humán és kutya eredetű minták feldolgozásán keresztül keres összefüggéseket az étrend összetétele, a mikrobiom mintázatai és a metabolikus szindróma egyes tünetei között, ennek megfelelően kutya és humán székletmintákhoz is van hozzáférésünk. A minták természetét figyelembe véve a módszertani tesztelési folyamatokat kutya mintákon végeztük: a donor állat étrendje lényegesen szabályozottabb, az egyazon egyedről való újbóli mintavétel pedig szükség esetén könnyebben kivitelezhető, mint emberi donorok esetén. Emellett, mivel a széklet metagenomi szekvenálás gyakran detektál gazda eredetű DNS-t, a humán genomi szekvenciák esetén szükséges bonyolult adatvédelmi és etikai intézkedések megelőzhetőek kutya minták használatával.

A használt minta begyűjtése az ürítéstől számított fél percen belül megtörtént. A gyűjtés során a mintát a nukleinsavak integritását megőrző DNA/RNA Shield-be (ZymoResearch) helyeztük, majd a laboratóriumban -80°C-on tároltuk.

A begyűjtött ürülekéből 4 izoláló kittel, minden esetben 4 párhuzamossal tisztítottunk DNS-t. A kísérletek ismert összetételű kontrolljaként a ZymoBIOMICS Microbial Community Standard mikroba mixet (ZymoResearch) használtuk.

3.1.2. Kontroll mikroba mix

A vizsgálatok pozitív kontrolljaként ismert összetételű standard mikroba mixet alkalmaztunk (ZymoBIOMICS Microbial Community Standard, ZymoResearch). Az 5. ábrán a mix pontos összetétele látható, az egyes baktérium fajok Gram-festődését is jelölve. A mixben található két gombafajt nem detektáltuk, mivel a mixekből izolált DNS-t csak a bakteriális 16S rRNS génre dúsítottuk. A mix feldolgozásával elsődleges célunk az izolálók kitek Gram-szelektivitásának feltérképezése volt, azaz annak, hogy az egyes kitek mutatnak-e eltérést a Gram-pozitív sejtfal feltárásának hatékonyságában. Amennyiben ez felmerülne, az komoly torzítást eredményezne az adott kittel kinyerhető adatokon.



5. ábra: a pozitív kontroll mikroba mix összetétele

3.2. A mintafeldolgozás összefoglalása

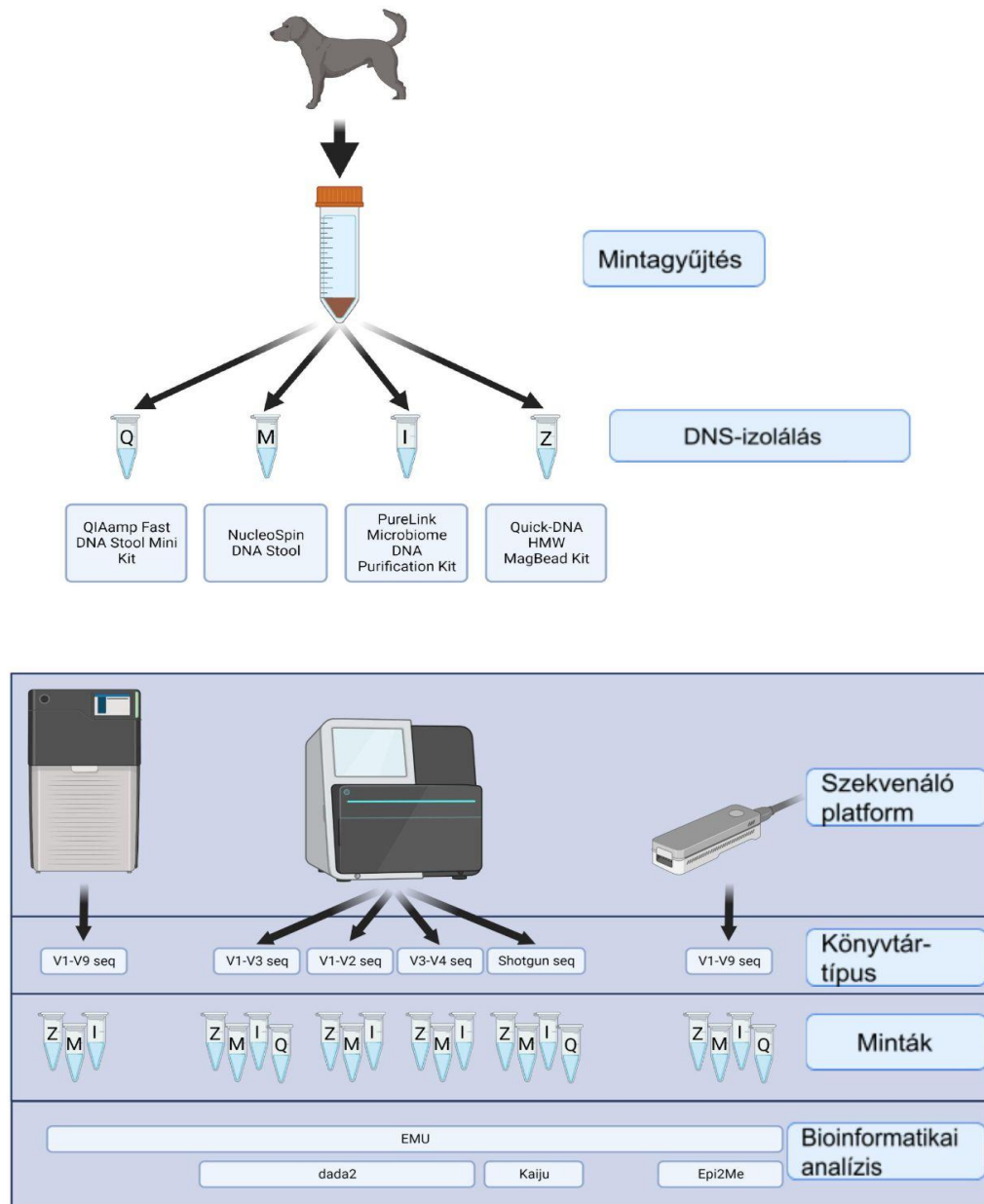
A 6. ábrán a kutatás munkafolyamatainak összefoglalása látható. A Qiagen, Macherey-Nagel, Invitrogen és ZymoResearch kitekkel 4-4 párhuzamos izolálást végeztünk egyenként 0,2 vagy 0,1 gramm mintából (az adott kit protokolljának megfelelően). A DNS mennyiségét és minőségét ellenőriztük. A Macherey-Nagel, Invitrogen és ZymoResearch kitekkel izolált DNS-ből Illumina V1-V2, V3-V4 és V1-V3 ampikon és shotgun, illetve ONT és PacBio hosszú leolvasású V1-V9 ampikon könyvtárakat készítettünk. A Qiagen kittel izolált DNS-ből (az alacsonyabb minőségből és mennyiségből kifolyólag) csak Illumina V1-V3, shotgun és ONT V1-V9 könyvtárakat készítettünk. A DNS-könyvtárakat minőségi és mennyiségi ellenőrzésnek vetettük alá.

Az Illumina V1-V2, illetve V3-V4 könyvtárakat Illumina MiSeq-kel, 500 cikluson szekvenáltuk. A V1-V3, illetve shotgun könyvtárak szekvenálását 300 és 500 ciklussal is elvégeztük. Az ONT és PacBio V1-V9 könyvtárakat a megfelelő szekvenátoron futtattuk. A leolvasások számát és minőségét ellenőriztük.

A DADA2 programmal az Illumina szekvenálásokat, Kaijuval a shotgun metagenomi eredményeket, Emuval minden ampikon szekvenálást, Minitax-szal az 500 ciklusos Illumina V1-V3 és shotgun eredményeket, EPI2ME-vel az ONT V1-V9 adatokat dolgoztuk fel.

Összehasonlító elemzéseket végeztünk a következő szempontokra nézve: a DADA2 és az Emu programok teljesítménye az Illumina ampikon readeken, az Emu és az EPI2ME teljesítménye az ONT readeken, a Kaiju és a Minitax teljesítménye az Illumina shotgun és V1-V3 readeken, a 300 és 500 ciklusos Illumina szekvenálások teljesítménykülönbsége shotgun és V1-V3 readeken, a

V1-V2, V1-V3, V3-V4 és V1-V9 amplicon szekvenálások, illetve az ONT és PacBio szekvenátorok egymáshoz viszonyított teljesítménye.



6. ábra: a mintafeldolgozás lépései és eszközei

3.3. DNS-izolálás és minőségellenőrzés

A -80°C-on tárolt, majd szobahőmérsékleten elszívófülke alatt 30-40 percig olvasztott mintából a Qiagen QIAamp Fast Stool DNA Mini Kit (50) - Pathogen Protocol, a Macherey – Nagel Nucleospin Stool Mini Kit Genomic DNA, az Invitrogen PureLink Microbiome DNA Purification Kit és a ZymoResearch Quick-DNA HMW MagBead Kit DNS-izoláló kitek protokolljai szerint 4 technikai párhuzamosban, egyenként 0,2, a ZymoResearch kit esetén 0,1 gramm minta felhasználásával izoláltunk DNS-t. A protokollok lényeges technikai eltéréseit az 1.3.1. alfejezet tárgyalja.

Az izolált DNS mennyiségét és minőségét minden izolálás esetében elvégeztük.

Mennyiségi vizsgálatot Qubit 4.0 fluoriméterrel (Thermo Fisher Scientific), Qubit HS DNA Assay Kit használatával végeztünk. Az izolált gDNS minőségét, a fragmensek méretét Agilent TapeStation 4150 készüléken a Genomic DNA Screen Tape assay-vel vizsgáltuk. A TapeStation gyors és megbízható elektroforetikus elválasztást tesz lehetővé, a méréssel a DNS integritási érték (DIN) meghatározható.

3.4. Könyvtárkészítés

A különböző kitekkel izolált DNS-ből hosszú és rövid leolvasású amplikon, illetve shotgun szekvenálási könyvtárakat készítettünk kereskedelmi forgalomban kapható kitek segítségével, azok protokollját követve. A szekvenálási könyvtárak az adott szekvenálási platformmal kompatibilis módon módosított DNS-t eredményeznek. Az amplikon könyvtárkészítő kitek a 16S rRNS gén hiperkonzervatív régióira specifikus primerek és PCR-reakció segítségével sokszorozzák fel a szekvenálás során megcélzott génszakaszt (rövid leolvasású Illumina amplikon szekvenálás esetén a V1-V2, V1-V3 vagy V3-V4 régiók, hosszú leolvasású szekvenálás esetén a teljes 16S rRNS gén a V1-V9 régiókat beleértve), emellett az amplikonokra felkerülnek a szekvenáláshoz és a minták későbbi azonosításához szükséges Illumina, SMRTbell (PacBio), és ONT adapterek. Az általunk alkalmazott shotgun könyvtárkészítő kit (Illumina DNA Prep Kit) 'tagmentációt' alkalmaz, ahol a DNS fragmentálása és a fragmensek adapterligálása (tagging) egy lépésben, egy speciális ún. *bead-linked transposome complex* felületén zajlik.

A V1-V2, illetve V3-V4 amplikon könyvtárak elkészítésére a Quick-16S NGS Library Prep Kit-et (ZymoResearch), a V1-V3 könyvtárakhoz a NEXTFLEX® 16S V1-V3 Amplicon-Seq Kit-et

(PerkinElmer) használtuk. A ZymoResearch kit a V1-V3 régió targetálására nem alkalmas, ezért alkalmaztuk két eltérő gyártó termékét.

A teljes gént lefedő V1-V9 amplikon szekvenálási könyvtárakat a 16S Barcoding Kit (ONT) és a SMRTbell Express Template Prep Kit 2.0 (PacBio) kitékkel készítettünk. A shotgun könyvtárak esetén a DNA Prep Kitet (Illumina) alkalmaztuk.

A Macherey-Nagel, Invitrogen és ZymoResearch kitek megfelelő mennyiségű és minőségű DNS-t izoláltak, melyből az összes kívánt szekvenálási könyvtárat el tudtuk készíteni. A Qiagen kittel kis mennyiségű és alacsony minőségű DNS-t nyertünk ki, melyet csak 16S V1-V3 Illumina, valamint shotgun szekvenálási könyvtárak elkészítésére tudtunk alkalmazni. Mivel ez utóbbi kit igen kis mennyiségű DNS-t nyert ki, a 0,2 gramm ürülékből kinyert összes DNS nem volt elegendő mindegyik könyvtár elkészítésére, továbbá, az alacsony DIN értékek alapján a teljes V1-V9 régió megszekvenálása is kérdésesnek bizonyult a Qiagen DNS mintákból. Ezekből végül V1-V9 ONT könyvtárat készítettünk, melyből csupán néhány leolvasást nyertünk.

A könyvtárak minőségét TapeStation 4150-en, D1000 ScreenTape Assay-vel ellenőriztük.

3.5. Szekvenálás

3.5.1. Rövid leolvasású Illumina szekvenálás

A 16S rRNS V régió és shotgun könyvtárakat Illumina MiSeq-kel szekvenáltuk 500 cikluson. A V1-V3 és a shotgun könyvtárakat emellett 300 ciklus beállítással is megszekvenáltuk. A nagyobb ciklusszám akár 100 bázispárral is növelheti a leolvasások hosszát - 500 cikluson 2×250 , 300 cikluson 2×150 bázispáros leolvasási kerettel számolhatunk - ugyanakkor a magasabb ciklusszámú szekvenálás időtartama, költsége és a feldolgozandó adat mennyisége is nagyobb. Illumina szekvenálás esetén tehát indokolt esetben érdemes optimalizálni a ciklusszámot; ha egy alacsonyabb ciklusszám is megfelelő minőségű adatot generál, akkor ennek a használata anyagi szempontból indokolt. A szekvenálásokhoz a MiSeq Reagent Nano Kit v2-t használtuk 300, illetve 500 cikluson.

3.5.2. Hosszú leolvasású ONT és PacBio szekvenálás

A long-read szekvenálásokat az ONT MinION platformon MinION Flow Cell-el és PacBio Sequel IIe platformon PacBio SMRT Cell 8M-el végeztük.

3.6. Bioinformatikai analízis

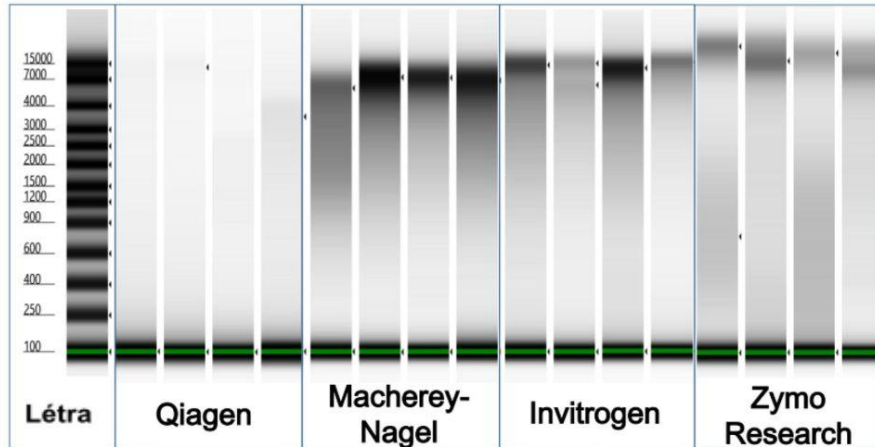
Az Illumina V régió szekvenálásokat DADA2 és Emu, az Illumina shotgun szekvenálásokat Kaiju és Minitax, az ONT V1-V9 szekvenálásokat EPI2ME és Emu, a PacBio V1-V9 szekvenálásokat Emu programokkal dolgoztuk fel. Az ONT szekvenciák bázishívására Guppy-t használtunk. Az adatok vizualizációjához a Phyloseqet alkalmaztuk.

A használt programok verziószáma: Guppy v6.3.8, DADA2 v.12.20.0, Kaiju v.1.9.0, Emu v3.4.4, EPI2ME Desktop Agent v3.5.7, Phyloseq v.1.36.0.

A következő adatbázisokat használtuk referenciaként: SILVA v.138 (DADA2), proGenomes v.3 (Kaiju, Minitax), az NCBI BLAST adatbázisa (EPI2ME). Az Emu saját adatbázist használ, mely az rrnDB v5.6 és az NCBI 16S RefSeq génbankokon alapul.

4. Eredmények és értékelésük

4.1. DNS-izolálás minősége



7. ábra: az egyes kitekkel izolált DNS TapeStation elemzése

A 7. ábra a TapeStation 4150-en végzett Genomic DNA ScreenTape Assay elemzést mutatja, melyen látható, hogy a négyből három kit nagy mennyiségű genomi DNS-t tudott izolálni a 7000-15 000 bázispáros mérettartományban mind a négy technikai párhuzamos esetében. Az izolált fragmensek hossza a marker sáv alapján, a mennyisége a futott sávok intenzitása alapján következtethető ki: ezek szerint a Macherey-Nagel kit tudott a legnagyobb mennyiségben, a ZymoResearch kit pedig a legnagyobb mérettartományban izolálni (a ZymoResearch kit ugyanakkor a többihez képest fele mennyiségű DNS-ből indul ki). Az Invitrogen kit esetében a

négy technikai párhuzamosban izolált DNS mennyisége változó. A Qiagen kit egyik párhuzamos esetében sem izolált megfelelő mennyiségű és méretű DNS-t. A Qiagen izolátumok alacsony minősége nem minden esetben tette ezeket alkalmassá további elemzésre, könyvtárkészítésre, így a Qiagen kittel izolált DNS-ből csak Illumina V1-V3, Illumina shotgun és ONT V1-V9 könyvtárakat készítettünk.

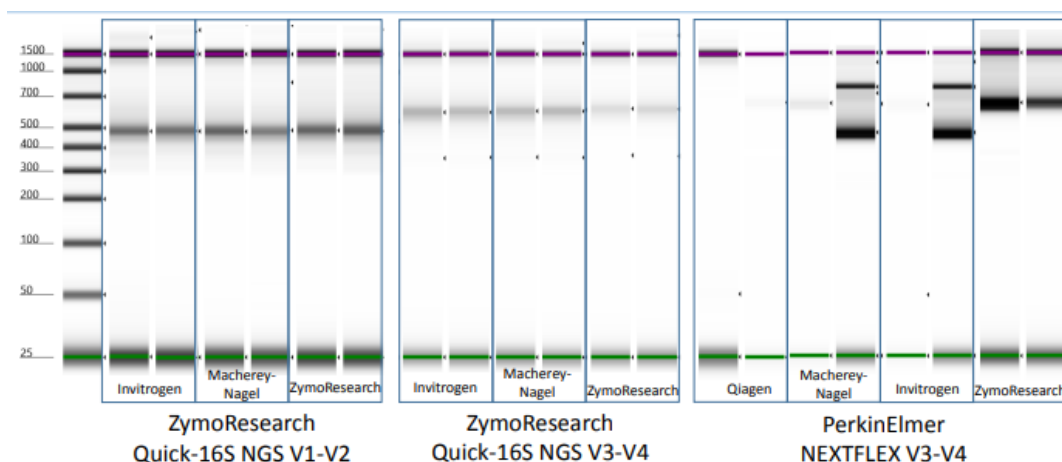
Az 1. táblázat bemutatja, hogy mennyi kiindulási anyagból pontosan milyen koncentrációjú DNS-t tudtunk az egyes kitekkel átlagosan (4 párhuzamos alapján) izolálni.

	Qiagen	Macherey-Nagel	Invitrogen	ZymoResearch
Kiindulási minta tömege (g)	0,2 g	0,2 g	0,2 g	0,1 g
Kinyert DNS koncentrációja (ng/μl)	1,815 ± 0,655 ng/μl	24 ± 7 ng/μl	14,6 ± 6,2 ng/μl	12,4 ± 1,5 ng/μl

1. táblázat: az egyes kitekkel adott mennyiségű székletmintából izolált DNS koncentrációja 4 párhuzamos átlaga alapján

4.2. Szekvenálási könyvtárak minősége

A 8. ábra az Invitrogen, Macherey-Nagel és ZymoResearch kitekkel izolált DNS-ből készített V régió amplikon könyvtárak minőségét mutatja TapeStation-ön, D1000 ScreenTape Assay-vel ellenőrizve, minden izolálásból két párhuzamosban. A ZymoResearch könyvtárkészítő kitjei mindhárom izolálás termékén jól teljesítettek, a könyvtárak megfelelő koncentrációban készültek el, bár a V1-V2 könyvtárak koncentrációja lényegesen magasabb. Ugyanakkor a PerkinElmer kit csak a ZymoResearch izoláló kittel izolált DNS-ből tudott jó minőségű könyvtárakat készíteni, a Macherey-Nagel és Invitrogen izolálásokból több esetben aspecifikus termékek képződtek, a Qiagen izolálásból készített könyvtár pedig alacsony koncentrációban tartalmaz DNS-t.



8. ábra: a V régió ampikon könyvtárak TapeStation elemzése

A shotgun, illetve long-read könyvtárakat egy-egy kittel készítettük, melyeket az Illumina, ONT illetve PacBio gyártók termékei voltak. A long-read könyvtárak mérete megközelítőleg 1700, a shotgun könyvtárak hossza megközelítőleg 650 bázispár volt. A könyvtárak további összehasonlítására ez esetekben a pályamunkában nem térek ki.

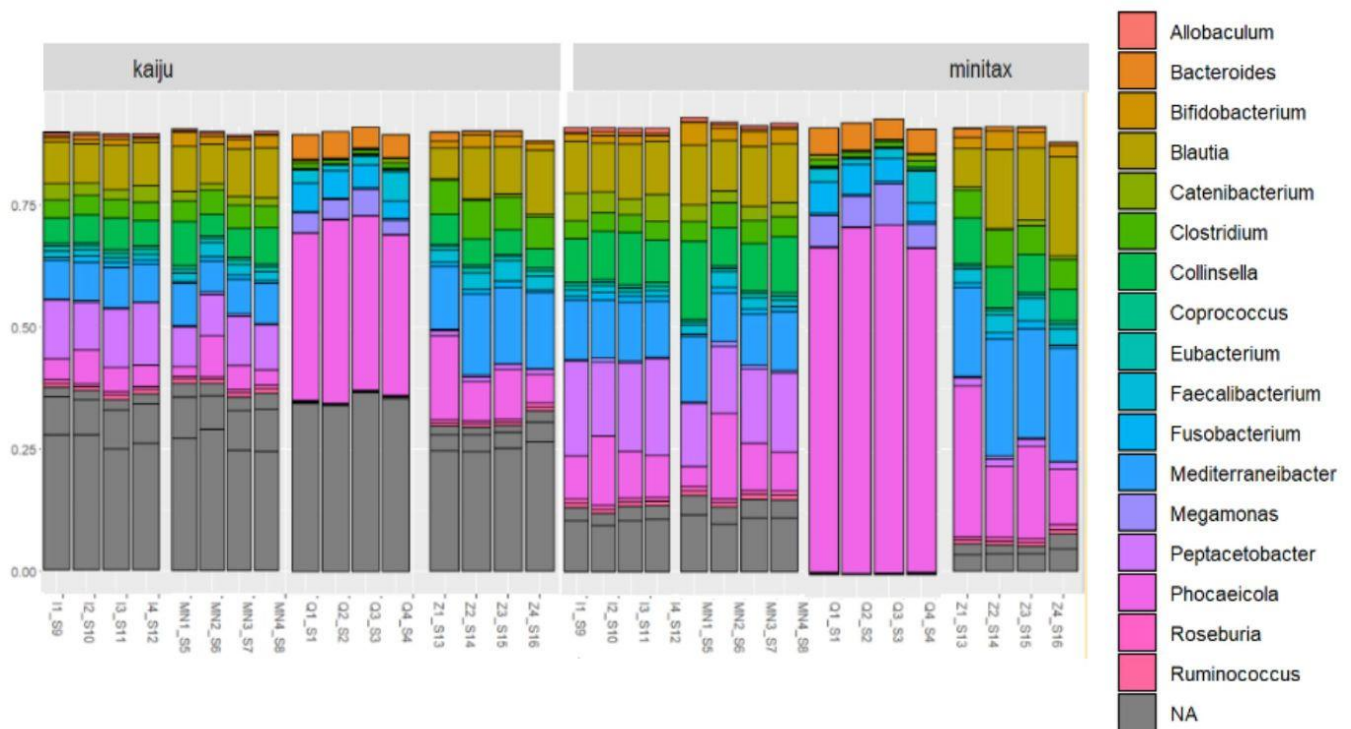
4.3. Szekvenálási eredmények, a minta összetétele

4.3.1. Shotgun metagenomi szekvenálás

A 9. ábrán az Illumina shotgun szekvenálások eredményei láthatók mind a négy izoláló kit alapján. A szekvenciák azonosítása Kaijuval és Minitax-szal a proGenomes adatbázis alapján történt. Az oszlopdiagramok a 20 leggyakoribb fellelt nemzetséget jelölik, a sötétszürke színek be nem azonosított taxonra utal. A minimap2 alapú Minitax a Kaiju-nál jóval kisebb arányban mutat beazonosítatlan leolvasásokat. Ennek lehet az az oka, hogy a minimap2 által térképezett szekvenciák effektíve hosszabbak, mivel a nukleotidszekvencia nem fordítódik át aminosavszekvenciára; emellett minimap2-vel a fehérjét nem kódoló régiók is elemezhetők. A Kaiju előnye ugyanakkor, hogy az aminosav-szinten kevesebb mutáció miatt az illesztések pontosabbak lehetnek. Mindazonáltal a két program nemzetség szinten jelentős hatékonyságbeli különbségének oka további kutatásra szorul.

Az Invitrogen és Macherey-Nagel kitek mindkét bioinformatikai pipeline szerint hasonló eredményeket hoztak. A ZymoResearch kit eredményei elsősorban a *Fusobacterium*, *Bacteroides* és *Peptacetobacter*-abundanciák terén mutatnak eltéréseket az Invitrogen és Macherey-Nagel kitektől; emellett kevesebb ismeretlen faj és nagyobb egyenletesség tapasztalható a Zymo mintákon. A

Qiagen kit a többihez képest nagymértékben felülreprezentálja a *Phocaeicola* taxont. A technikai párhuzamosok összetétele az Invitrogen kit esetén hasonlóan egymásra.



9. ábra: A shotgun metagenomi szekvenálással azonosított 20 leggyakoribb nemzetség a 4 különböző kittel izolált DNS, illetve a Kaiju és Minitax programok alapján

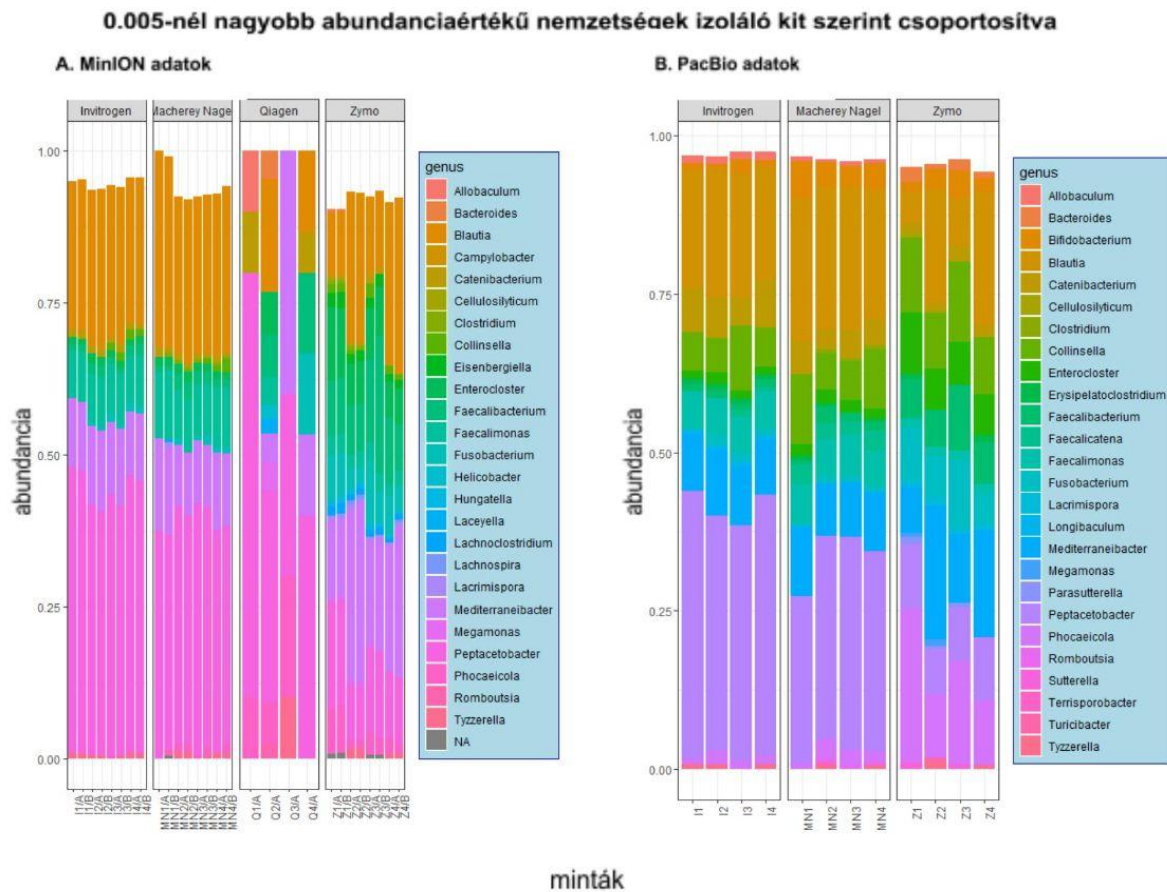
4.3.3. 16S rDNS hosszú leolvasású szekvenálás

A 10.A és 10.B ábrákon a 20 leggyakoribb nemzetség relatív abundanciája látható a MinION (A) és PacBio (B) szekvenálásokat összehasonlítva. A bioinformatikai elemzéshez az Emu-t használtuk. Az Invitrogen és Macherey-Nagel-izolátum alapú könyvtárak mindkét szekvenálás szerint hasonló összetételűek, bár egyes alacsony abundanciájú taxonok esetén tapasztalható jelentős eltérés közöttük. A Qiagen izolálás minősége alacsony volt, ez alapján csak nanopore szekvenálással tudtunk néhány, biológiailag értékelhetetlen leolvasást kinyerni.

Az Invitrogen és Macherey-Nagel minták az ONT szekvenálás szerint elsősorban *Blautia*, *Peptacetobacter*, *Mediterraneibacter* és *Faecalimonas* nemzetségekből álltak. Ezekben a mintákban a PacBio-szekvenálás is jelentős *Blautia*, *Peptacetobacter* és *Faecalimonas*-abundanciát detektált, emellett kimutatott bizonyos, az ONT-vel nem fellelt genusokat, pl. *Allobaculum*, *Bifidobacterium*.

A Zymo minták közül kettő a többinél jelentősen nagyobb *Blautia*-és *Mediterraneibacter*-abundanciát mutatott, amit mind a két hosszú leolvasási megközelítéssel

kimutatható volt. Az ONT-hez képest a PacBio kevesebb *Peptacetobacter*, de több *Phocaeicola*-t detektált.

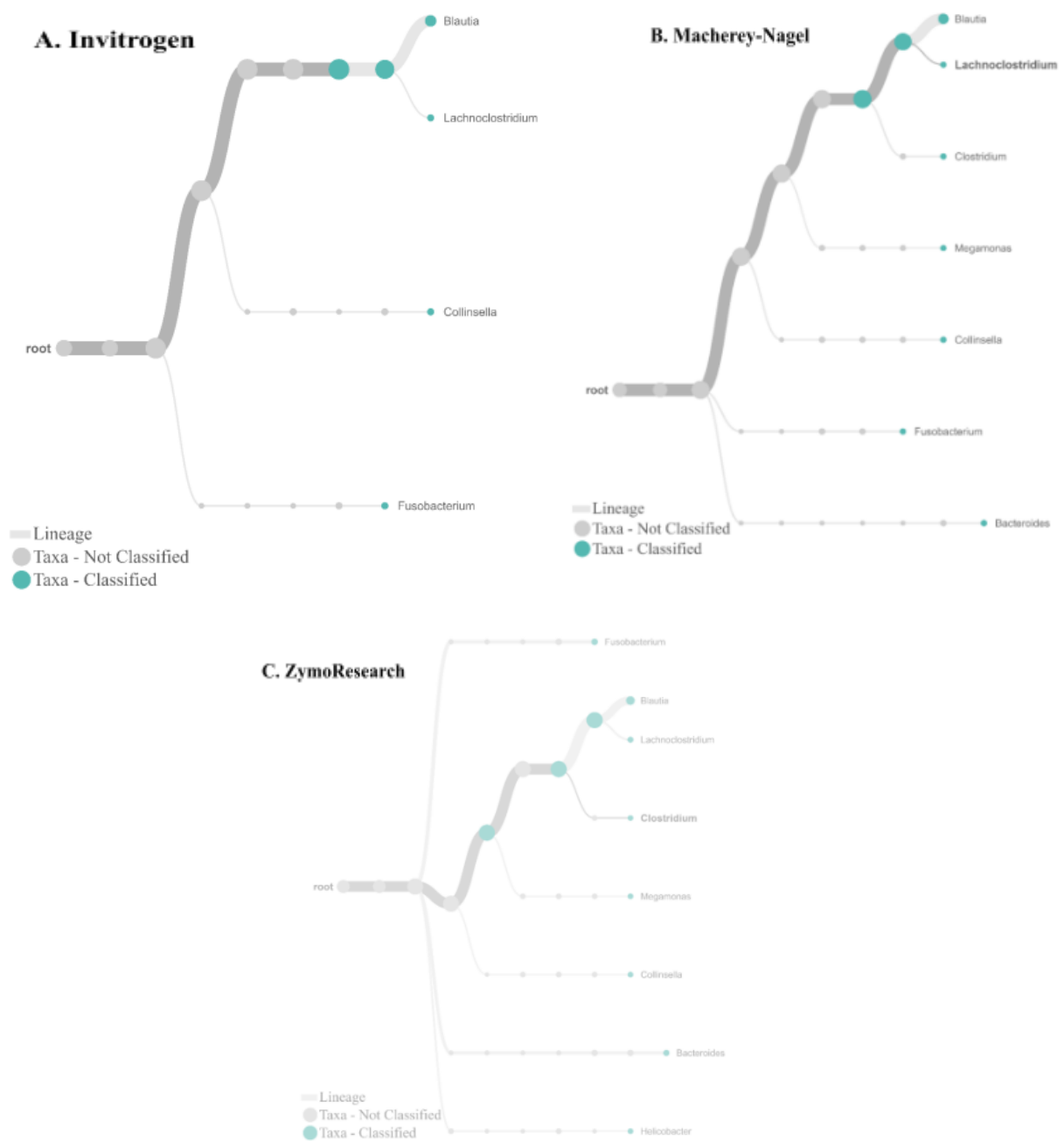


10. ábra: az ONT (A) és PacBio (B) 16S V1-V9 ampikon szekvenálással azonosított leggyakoribb nemzetségek

Az Invitrogen és Macherey-Nagel izolálások hasonlóságára a hosszú leolvasású szekvenálás további bizonyítékot szolgáltat. A két megközelítést összehasonlítva bizonyos különbségek láthatók a minták összetételében, a PacBio szekvenálás jellemzően több taxont tudott azonosítani. Ennek állhat a hátterében a két technológia különbsége, ugyanis PacBio szekvenálással egy cirkularizált könyvtárról több, egymásra illeszthető, így nagyobb megbízhatóságú leolvasás keletkezhet, míg a nanopore szekvenálás esetén az egyes bázisok csak egyszer leolvashatók. Mivel a két szekvenálás eredményei nem mutatnak drasztikus eltérést, a PacBio Sequel IIe-nél lényegesen költséghatékonyabb ONT MinION is ideálisnak mutatkozik a mikrobiom összetételének megállapítására a teljes 16S rRNS gén ampikonjainak szekvenciája alapján.

A 11. ábra az ONT long-read szekvenálások EPI2ME elemzését mutatja be a 3 izoláló kit szerint (a Qiagen kittel kinyert kis mennyiségű DNS-ről mindössze néhány leolvasást kaptunk, így az onnan származó adatok nem reprezentatívak). Az Invitrogen és Macherey-Nagel izolálások eredményei

egymásra jobban hasonlítanak, mint a ZymoResearch kit adataira, viszont minden kit szerint egységes, hogy a minták legabundánsabb nemzetsége a Blautia.



11. ábra: az egyes kitek izolálása alapján EPI2ME-vel azonosított nemzetségek 1-1 mintán

A Blautia az egyéb megközelítések döntő része szerint is jelentős mennyiségben található meg a mintákban, ugyanakkor nem mutat abszolút dominanciát. Az ONT szekvenálási adatokat Emu programmal elemezve a leggyakoribb nemzetség jellemzően a Peptacetobacter, amelyet a Blautia (Lachnospiraceae) követ. A Peptacetobacter nemzetséget az EPI2ME nem detektálta, bár annak jelenlétét a shotgun elemzés is kimutatta. A V régió amplikonok alapján a Peptoclostridium

mutatható ki nagyobb mennyiségben, amely a *Peptacetobacter*rel együtt a *Peptostreptococcaceae* családhoz tartozik, tehát rokon genusokról beszélhetünk. Ezek elkülönítésére a lényegesen több szekvenciaadatot integráló long-read és shotgun módszerek alkalmasabbak a V régió amplikon szekvenálásánál (Zhang és mtsai. 2018); így a V régió amplikon szekvenálással detektált *Peptoclostridium* valójában akár *Peptacetobacter* is lehet. Ez utóbbi taxon detektálásában mutatott sikertelenség és a *Blautia* jelentős felülreprezentációja az EPI2ME program lehetséges hibájára utal. A *Blautia* felülreprezentációja mögött elképzelhető, mint ok, hogy a szintén a *Lachnospiraceae* családba tartozó és az egyéb megközelítések szerint 10% körüli abundanciát mutató *Mediterraneibacter* genust *Blautia*ként azonosította az EPI2ME.

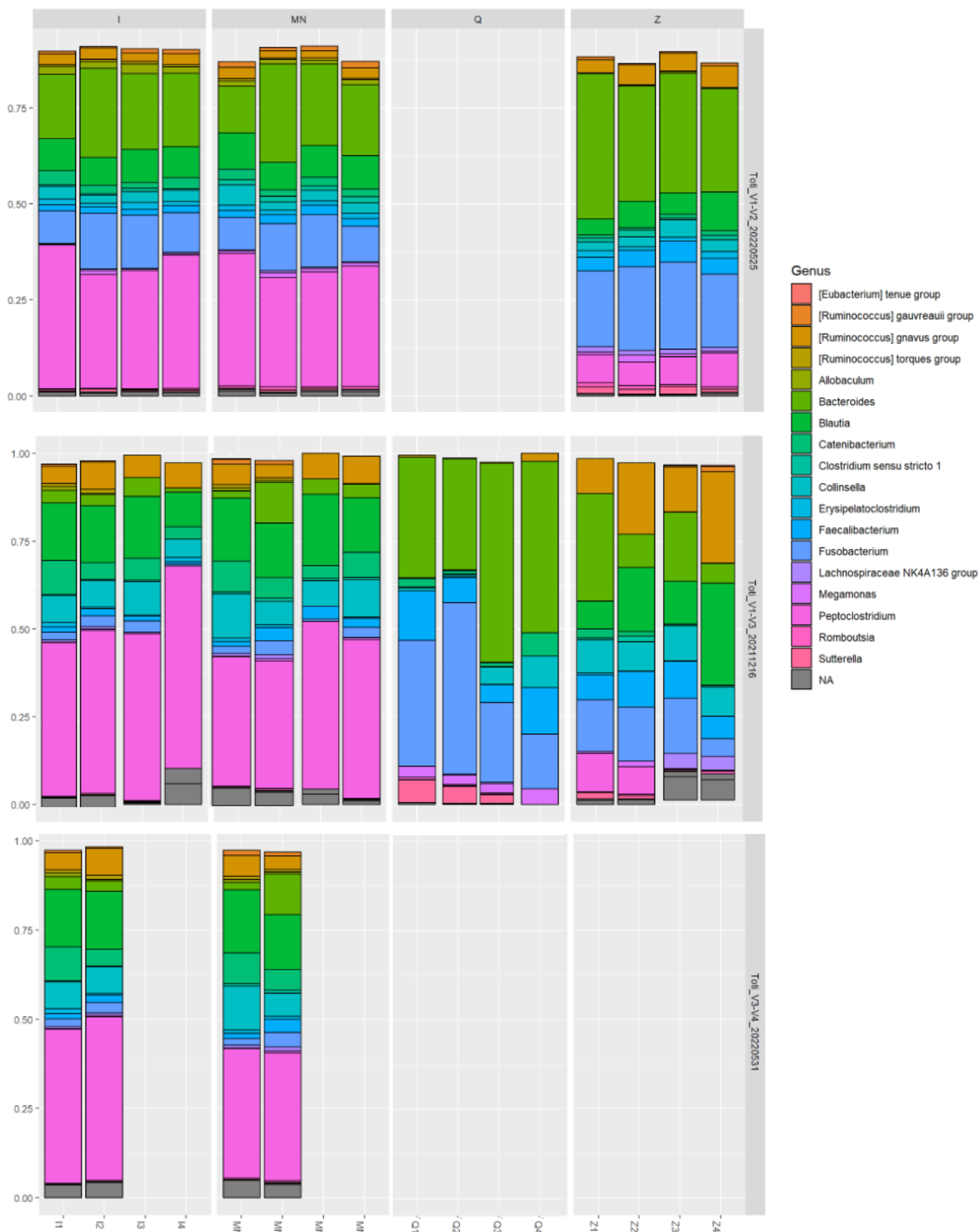
4.3.4. 16S rDNS V régiók amplikon szekvenálása

A V1-V2 amplikon szekvenálást az Invitrogen, Macherey-Nagel és ZymoResearch kitek alapján végeztük. Az Invitrogen és Macherey Nagel-eredmények nagyon hasonlóak az összetételek és a párhuzamosok közötti variabilitás tekintetében. A ZymoResearch kit alapján ezekkel szemben több *Bacteroides* és *Fusobacterium*, de kevesebb *Peptoclostridium* lett azonosítva. A V régió szekvenálásokkal azonosított *Peptoclostridium*-szekvenciák, ahogy azt az EPI2ME adatok kiértékelésénél részleteztem, származhatnak a shotgun szekvenálással és V1-V9 szekvenálással (Emu bioinformatikai elemzés szerint) nagyobb mennyiségben azonosított *Peptacetobacter*-fajoktól is.

Az Invitrogen- és Macherey-Nagel-eredmények a V1-V3 amplikonok esetén is hasonlítanak, a legjelentősebb nemzetség a *Peptoclostridium*, melynek detektálásában a Qiagen és a Zymo kitek is alulteljesítettek. A V1-V2-szekvenálással jelentős mennyiségű *Fusobacterium*ok azonosításában ugyanakkor ez utóbbi kitek teljesítettek jobban, az Invitrogen-és Macherey-Nagel izolálásokból a V1-V3 szekvenálás csupán kis mértékben detektált *Fusobacteriát*.

A V3-V4 szekvenciák alapján is megállapítható az Invitrogen- és Macherey-Nagel-izolálások hasonlósága, taxonösszetétel terén itt jelentős mennyiségű *Peptoclostridium*, ám kevés *Fusobacterium* található. A V3-V4 szekvenciák bioinformatikai elemzése a dolgozat leadásának időpontjában is folyik, melynek következtében jelen pályamunkában nem tudunk teljes értékű V3-V4 adatokat közölni.

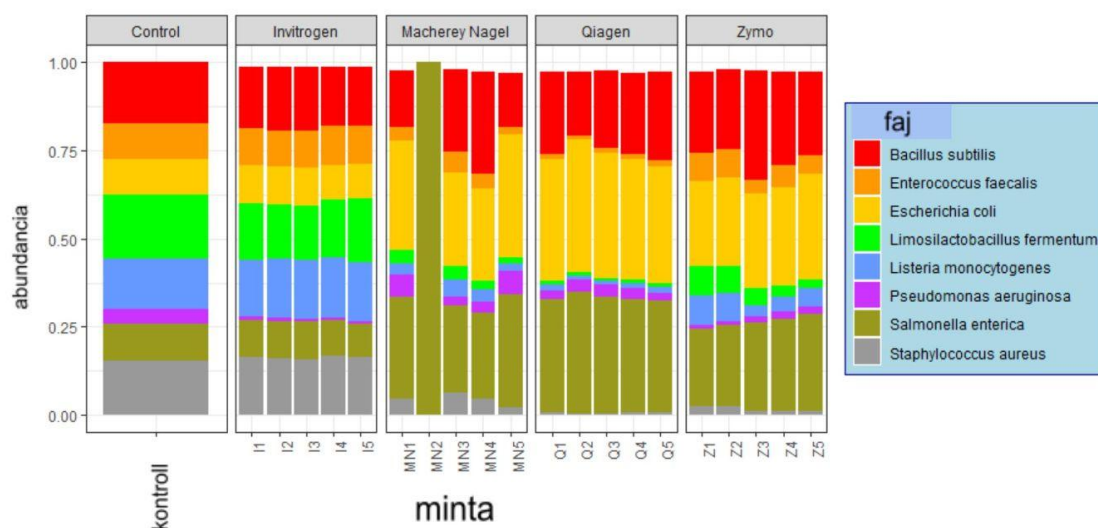
A V1-V2, V1-V3, V3-V4 amplikon szekvenálással azonosított leggyakoribb nemzetségek relatív abundanciáját a 12. ábra mutatja be.



12. ábra: a 16S V régió amplikon szekvenálás alapján azonosított 20 leggyakoribb nemzetség

A teljesen kiértékelt V1-V2 és V1-V3-adatok, illetve a jelenleg elérhető Invitrogen és Macherey-Nagel V3-V4 adatok arra mutatnak rá, hogy a V1-V3 és a V3-V4 adatok egymásra, illetve a shotgun-és long-read adatokra is jobban hasonlítanak, mint a V1-V2 szekvenciákra alapuló adatokra; habár a Qiagen-és ZymoResearch V1-V3 adatai nagyobb hasonlóságot mutatnak a V1-V2 adatokkal. A V1-V2 szekvenálással minden esetben, V1-V3 szekvenálással pedig a Qiagen-és Zymo-izolálásokról jelentősebb mennyiségben azonosított *Fusobacterium* az Emuval kiértékelt long-read, illetve a shotgun szekvenciák alapján ennél lényegesen kisebb abundanciával vannak jelen a vizsgált mintában; ez utalhat arra, hogy a V1-V2 régió - melyet a V1-V3 ampikon is tartalmaz - alapján a taxonösszetételi adatok torzulnak. Feltűnő, hogy a V1-V3 szekvenálásokról a *Fusobacterium* csak a Qiagen-és Zymo-izolálásokban felülreprezentált. A *Fusobacterium* egy Gram-negatív taxon; a két utóbbi kit a Gram-pozitív sejtfalak lízisében kevésbé bizonyult hatékonynak, mint az Invitrogen-és Macherey-Nagel-protokollok (lásd a 4.3.6. és 4.3.7. alfejezeteket). A shotgun és long-read (Emu) módszerek szerint is jelentős, 10% fölötti abundanciát mutató *Mediterraneneibacter* nemzetséget (*Lachnospiraceae*) egyik V régió alapján sem detektáltuk. A felsorolt torzítások, minták közötti eltérések alapján gyanítjuk, hogy a V régió ampikon szekvenálás nem kellően pontos egy széklet mikrobiom pontos összetételének meghatározására.

4.3.6. A kontroll mikroba mix szekvenálása

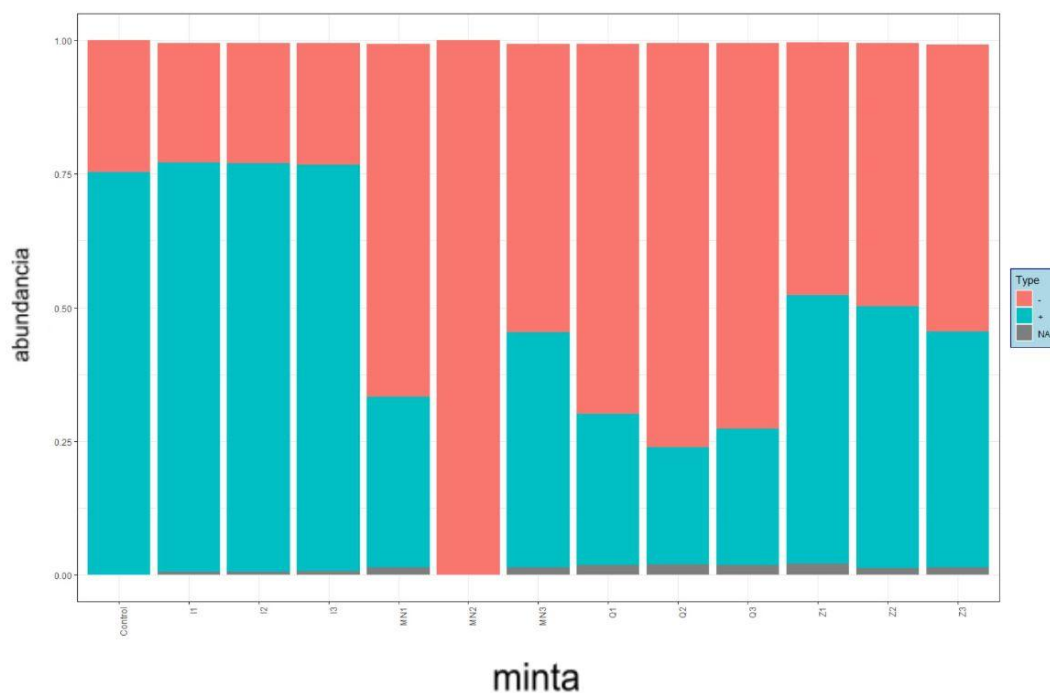


13. ábra: a kontroll mikroba mixből az egyes kitekkel izolált DNS alapján azonosított fajok

A ZymoResearch Microbial Community Standard mikrobamixből öt párhuzamosan végeztünk DNS-izolálást az egyes kitekkel. A 13. ábrán a várt értékek mellett az Emu-val kielemezett ONT long-read szekvenálások adatait látjuk. Eredményeink szerint a *Pseudomonas aeruginosa* (Gram-)

kismértékű alulreprezentációjától eltekintve az Invitrogen kit adta vissza legnagyobb pontossággal a várt adatokat, illetve a technikai párhuzamosok hasonlósága is itt volt a legnagyobb. A többi kitról általánosan megállapítható a *S. enterica* és az *E. coli* Gram-negatív fajok erős felülreprezentációja, illetve a Gram-pozitív *L. fermentum*, *L. monocytogenes* és *S. aureus* baktériumok jelentősen csökkent detekciója. A Macherey-Nagel kittel végzett párhuzamos izolálások eredménye egymástól a többi kittel összevetve nagyobb mértékben tér el, az egyik párhuzamos során csak a *S. enterica* (-) fajt sikerült kimutatni (ebben a párhuzamosban a szekvenálás összesen 3 leolvasást ért el, így ez az adat biológiailag valójában nem releváns). A Qiagen-izolátumokban a *B. subtilis* (+), *E. coli* (-) és *S. enterica* (-) jelentősen túlreprezentált, a *S. aureus* (+) detekciója minden esetben sikertelen, a párhuzamosok hasonlósága nagy. A ZymoResearch párhuzamosok egymástól közepes mértékben térnek el.

A 14. ábra a Gram-pozitív és Gram-negatív baktériumok arányát mutatja a kontroll mixben, illetve az egyes kitekkel nyert izolátumokban, ONT V1-V9 szekvenálás és Emu analízis alapján. Az Invitrogen kit megfelelő mértékben lizálja a Gram-pozitívok sejtfalát, hatékonyságban a ZymoResearch protokollja követi, ám a várt eredménytől való eltérése ennek a kitnek már jelentős. A Macherey-Nagel és Qiagen kitek kevésbé mutatkoztak hatékonyak a Gram-pozitív baktériumok DNS-ének kinyerésében.



14. ábra: a kontroll mixből az egyes kitekkel izolált Gram-pozitív és Gram-negatív fajok aránya

A vizsgált kitek közül a legnagyobb hatékonyságú Invitrogen termék gyöngyöket használ a fizikai sejtfeltárás során. A Gram-pozitívok kimutatásában és a várt abundanciák visszaadásában második leghatékonyabb ZymoResearch kit a gyöngyök helyett lizozimet használ. A gyöngyök nagyobb hatékonyságának kulcsa, hogy a sejtfalat fizikailag roncsolják. Ugyanakkor a behatás az izolálendő DNS-ben is kárt tehet, és amennyiben összehasonlítjuk a DNS-izolátumok TapeStation-értékeléseit (4.1. alfejezet), láthatjuk, hogy a ZymoResearch kittel izolált DNS nagyobb mérettartományba esik, illetve az Invitrogen-izolátumok mennyisége az egyes technikai párhuzamosok között eltérő.

4.3.7. Gram-pozitív kinyerés székletmintából

A standardizált mikrobamixek eredményeiről a széklet metagenomi izolálásban való alkalmazásra következő figyelembe kell venni a minták eltérő természetét, illetve a kinyert DNS downstream feldolgozásának szempontjait. A standard mixszel szemben a széklet nagy mennyiségben tartalmaz természetes PCR-inhibitorokat, a kitek értékelésénél az ezek eltávolításában elért hatékonyság, illetve a kinyert DNS tisztasága és integritása is fontos tényező a bakteriális sejtfalak egységes lízise mellett. Így az egyes kitekkel izolált Gram-pozitív DNS arányát érdemes a valós székletmintán is értékelni. A 15. ábra az egyes kitekkel nyert DNS-ből azonosított nemzetségek százalékos gyakoriságát mutatja hőtérképen, kiemelve az egyes taxonok Gram-festődését is. Az adatok V1-V3 amplikon szekvenáláson és DADA2 analízisen alapulnak.

A többi elemzés során tapasztalt jellegzetes mintázat, azaz az Invitrogen és Macherey-Nagel kitek jelentős hasonlósága, a Qiagen-alapú eredmények nagyfokú eltérése és a Zymo adatok átmeneti jellege itt is tapasztalható. Az Invitrogen és Macherey-Nagel kitek alapján az 5 leggyakoribb taxon Gram-pozitív. A Qiagen-izolálások szerint a Gram-negatív *Fusobacterium* és *Bacteroides* a két leggyakoribb nemzetség. A ZymoResearch protokolljával hasonló arányban azonosítottunk Gram-pozitív és Gram-negatív nemzetségeket a leggyakoribbak között. A párhuzamosok közti variabilitás az Invitrogen kit esetében volt a legkisebb.

Ezen adatok alapján megállapíthatjuk, hogy bár ismert összetételű mikroba mix esetén a Macherey-Nagel protokollhoz képest a ZymoResearch kit adatai hasonlítottak jobban a várt, illetve Invitrogen eredetű adatokra, székletmintával dolgozva a Macherey-Nagel terméke nyer ki az Invitrogenhez jobban hasonlító adatokat. A kontroll, illetve természetes minták feldolgozása, illetve

ezen eredmények összevetése arra utalnak, hogy az Invitrogen kit adhatja vissza leghűségesebben a Gram-pozitív és Gram-negatív taxonok arányát.

Genus	I1	I2	MN1	MN2	Q1	Q2	Z1	Z2
Parasutterella (-)	0,04%	0,08%	0,02%	0,08%	0,99%	1,42%	0,81%	0,29%
Megamonas (-)	0,08%	0,38%	0,18%	0,63%	2,59%	2,65%	0,52%	1,39%
Sutterella (-)	0,11%	0,31%	0,06%	0,25%	4,79%	2,38%	1,28%	0,70%
Faecalitalea (+)	0,21%	0,10%	0,41%	0,37%	0,41%	1,40%	0,32%	0,15%
Clostridium sensu stricto 1 (+)	0,24%	0,27%	0,59%	0,79%	0,29%	0,20%	0,60%	1,51%
[Ruminococcus] gauvreauii group (+)	0,59%	0,38%	1,33%	1,00%	0,02%	0,01%	0,44%	0,38%
Lachnospiraceae NK4A136 group (+)	0,64%	0,60%	0,79%	1,03%	0,18%	0,28%	3,43%	2,35%
[Ruminococcus] torques group (+)	0,90%	1,14%	0,97%	0,74%	0,02%	0,03%	0,18%	0,24%
Candidatus Stoquefichus (+)	0,98%	0,42%	0,70%	0,52%	0,02%	0,01%	0,44%	0,48%
Allobaculum (+)	1,00%	0,38%	0,82%	0,54%	0,00%	0,01%	0,08%	0,06%
Erysipelatoclostridium (+)	1,21%	0,49%	0,88%	0,84%	0,03%	0,08%	1,19%	0,35%
Faecalibacterium (+)	1,63%	2,03%	1,51%	3,71%	7,14%	5,09%	6,28%	9,77%
Fusobacterium (-)	2,09%	3,08%	1,87%	4,07%	47,96%	22,34%	14,59%	15,87%
Bacteroides (-)	3,49%	3,00%	2,02%	11,41%	30,99%	55,48%	29,56%	9,66%
NA	3,72%	4,35%	5,02%	3,73%	0,12%	0,16%	1,35%	1,34%
[Ruminococcus] gnavus group (+)	4,84%	7,68%	5,88%	3,85%	0,30%	0,35%	7,71%	16,57%
Collinsella (+)	7,93%	7,56%	12,56%	6,66%	0,60%	4,82%	8,61%	7,69%
Catenibacterium (+)	9,60%	4,74%	8,66%	5,87%	0,99%	0,99%	2,26%	1,20%
Blautia (+)	16,02%	15,90%	17,47%	15,14%	0,17%	0,14%	7,72%	16,31%
Peptoclostridium (+)	42,88%	45,39%	35,95%	35,47%	0,39%	0,21%	8,75%	6,65%

15. ábra: a valós székletmintából az egyes kitekkel kimutatott leggyakoribb nemzetségek a Gram-festődést megelőzően.

I=Invitrogen, MN=Macherey-Nagel, Q=Qiagen, Z=ZymoResearch DNS-izoláló kitek.

5. Összefoglalás

Az intesztinális mikrobiom komplex, kiterjedt mechanizmusok útján hatást gyakorol a gazdaszervezet egészségi állapotára. Számos, elsősorban életmódalapú betegség esetén feltételezzük, hogy a bélflóra fontos szerepet játszik a kóros folyamatok kialakulásában. A mikrobiom jobb megismerése új diagnosztikai és terápiás eljárások kidolgozásához és finomításához járulhat hozzá, azonban a kutatómunka során alkalmazott módszerek kidolgozottsága jelenleg még nem kielégítő.

A bél mikrobiom vizsgálata elsősorban a székletmintából izolált DNS-és RNS-állomány alapján zajlik, hiszen ezek szekvenciája számos értékes információt tartalmaz. Probléma ugyanakkor, hogy a mintagyűjtéstől a szekvenciaelemzésig tartó munka során használható különböző módszerek megközelítése eltérő, így fennáll a kockázata annak, hogy ugyanarról a mikrobiomról eltérő információkat nyerjen két eltérő metodikával dolgozó kutatócsoport. Ez a tudományos kutatási folyamatok és a klinikai alkalmazhatóság során szükséges sztenderdizáció útjában áll, hiszen nem garantált, hogy a különböző módszerrel nyert adatokat egymással összevethetőként lehet kezelni.

Hogy az egyes módszerek esetleges torzításait jobban megértsük, kutatócsoportunk megvizsgálta egy kutya székletminta metagenomját összesen négy DNS-izoláló, hat könyvtár-készítő, négy szekvenáló és öt bioinformatikai módszer használatával, ismert összetételű kontroll mikroba mix feldolgozása mellett. Kísérleteink során megállapítottuk, hogy az alkalmazott DNS-izoláló kitek egységnyi mintából eltérő koncentrációjú, tisztaságú és integritású DNS-t vonnak ki, és nem minden kit alkalmas hosszú leolvasással megszekvenálható DNS kinyerésére. Az egyes kitek a Gram-pozitív mikrobák sejtfalának lízisében is eltérő hatékonyságot mutatnak, ez pedig döntően befolyásolja a kapott eredményeket. A 16S rRNS taxonómiai markergén azonosításra használt régiója, a szekvenálás során alkalmazott platform, megközelítés és leolvasási hossz, valamint az alkalmazott bioinformatikai elemző szoftver mind befolyásolta a vizsgált minta taxonómiai összetétel-adatait.

Eredményeink hangsúlyozzák annak fontosságát, hogy a mikrobiom-kutatás során nagy körültekintéssel kell megválasztani a használt módszereket, figyelembe véve az egyes változók eredményekre gyakorolt hatását. A módszertanilag tudatos kutatómunka a mikrobiom-kutatás további nagy fellendüléséhez vezethet, és nagyban segítheti a terület ismereteinek gyógyászati alkalmazását.

Köszönetnyilvánítás

Munkám végére érve szeretnék köszönetet mondani azoknak, akik nélkül ez a pályamunka nem jöhetett volna létre. Szeretném megköszönni az Orvosi Biológiai Intézet vezetőjének, prof. dr. Boldogkői Zsoltnak, és a GeMiNI projekt vezetőjének és témavezetőmnek, dr. habil. Tombácz Dórának, hogy részt vehettem a kutatásban. A témavezetésért, a folyamatos konzultációkért Tombácz Dóra mellett Gulyás Gábornak is hálás köszönettel tartozom. Szeretném megköszönni dr. Csontné dr. Kiricsi Mónikának, hogy vállalta a belső konzulens szerepét. Az adatok, melyek elemzésén ez a pályamunka alapul, Járay Tamás, Dörmő Ákos, dr. Csabai Zsolt, dr. Kakuk Balázs, illetve témavezetőim kitartó wet lab és bioinformatikai munkájának eredményeként jöttek létre, így külön köszönetet szeretnék mondani mindannyiuknak. A bioinformatikai elemzések során használt Minitax kifejlesztése dr. Kakuk Balázs érdeme. A projekthez felhasznált minta donorja: Harsányi-Tekergő Bátor "Toto".

A bemutatott kutatásnak keretet adó GeMiNI projekt a Magyar Tudományos Akadémia Lendület Program LP2020/8 (Tombácz Dóra) támogatásával jött létre. A pályamunka a Kulturális és Innovációs Minisztérium ÚNKP-22-1-SZTE-328. kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült. Hallgatói kutatómunkámat emellett a Kárpát-medencei Tehetségekutató Alapítvány is támogatja.

Irodalomjegyzék

Akbari, M., Hansen, M. D., Halgunset, J., Skorpen, F., Krokan, H. E. Low Copy Number DNA Template Can Render Polymerase Chain Reaction Error Prone in a Sequence-Dependent Manner. *The Journal of Molecular Diagnostics*, 7, 1, 36-39 (2005) doi = [https://doi.org/10.1016/S1525-1578\(10\)60006-2](https://doi.org/10.1016/S1525-1578(10)60006-2)

Amarasinghe, S.L., Su, S., Dong, X. et al. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol* 21, 30 (2020). <https://doi.org/10.1186/s13059-020-1935-5>

Andrade JC, Almeida D, Domingos M, Seabra CL, Machado D, Freitas AC, Gomes AM. Commensal Obligate Anaerobic Bacteria and Health: Production, Storage, and Delivery Strategies. *Front Bioeng Biotechnol*. 2020 Jun 5;8:550. doi: 10.3389/fbioe.2020.00550. PMID: 32582673; PMCID: PMC7291883.

Appleton J. The Gut-Brain Axis: Influence of Microbiota on Mood and Mental Health. *Integr Med (Encinitas)*. 2018 Aug;17(4):28-32. PMID: 31043907; PMCID: PMC6469458.

Banerjee, S., Schlaeppli, K. & van der Heijden, M.G.A. Keystone taxa as drivers of microbiome structure and functioning. *Nat Rev Microbiol* 16, 567–576 (2018). <https://doi.org/10.1038/s41579-018-0024-1>

Braune, A., Blaut, M. (2016) Bacterial species involved in the conversion of dietary flavonoids in the human gut, *Gut Microbes*, 7:3, 216-234, DOI: 10.1080/19490976.2016.1158395

Bukin, Y., Galachyants, Y., Morozov, I. et al. The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Sci Data* 6, 190007 (2019). <https://doi.org/10.1038/sdata.2019.7>

Callahan, B., McMurdie, P. & Holmes, S. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* 11, 2639–2643 (2017). <https://doi.org/10.1038/ismej.2017.119>

Chakravorty, S., Helb, D., Burday, M., Connell, N., & Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of microbiological methods*, 69(2), 330–339. <https://doi.org/10.1016/j.mimet.2007.02.005>

Chelakkot, C., Choi, Y., Kim, DK. et al. Akkermansia muciniphila-derived extracellular vesicles influence gut permeability through the regulation of tight junctions. *Exp Mol Med* 50, e450 (2018). <https://doi.org/10.1038/emm.2017.282>

Coelho, L.P., Kultima, J.R., Costea, P.I. et al. Similarity of the dog and human gut microbiomes in gene content and response to diet. *Microbiome* 6, 72 (2018). <https://doi.org/10.1186/s40168-018-0450-3>

Curry, K.D., Wang, Q., Nute, M.G. et al. Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nat Methods* 19, 845–853 (2022). <https://doi.org/10.1038/s41592-022-01520-4>

de Boer R, Peters R, Gierveld S, Schuurman T, Kooistra-Smid M, Savelkoul P. Improved detection of microbial DNA after bead-beating before DNA isolation. *J Microbiol Methods*. 2010 Feb;80(2):209-11. doi: 10.1016/j.mimet.2009.11.009. Epub 2009 Dec 6. PMID: 19995580.

Dietert RR, Dietert JM. The Microbiome and Sustainable Healthcare. *Healthcare (Basel)*. 2015 Mar 3;3(1):100-29. doi: 10.3390/healthcare3010100. PMID: 27417751; PMCID: PMC4934527

Fiedorová K, Radvanský M, Němcová E, Grombířiková H, Bosák J, Černochohá M, Lexa M, Šmajš D, Freiburger T. The Impact of DNA Extraction Methods on Stool Bacterial and Fungal Microbiota Community Recovery. *Front Microbiol*. 2019 Apr 17;10:821. doi: 10.3389/fmicb.2019.00821. PMID: 31057522; PMCID: PMC6479168.

Flekna G., Schneeweiss, W., Smulders, F. J. M., Wagner, M., Hein, I. Real-time PCR method with statistical analysis to compare the potential of DNA isolation methods to remove PCR inhibitors from samples for diagnostic PCR. *Molecular and Cellular Probes*, Volume 21, Issue 4 (2007) Pages 282-287, ISSN 0890-8508, <https://doi.org/10.1016/j.mcp.2007.02.001>.

Galloway-Peña J, Hanson B. Tools for Analysis of the Microbiome. *Dig Dis Sci*. 2020 Mar;65(3):674-685. doi: 10.1007/s10620-020-06091-y. PMID: 32002757; PMCID: PMC7598837.

Hasani A, Ebrahimzadeh S, Hemmati F, Khabbaz A, Hasani A, Gholizadeh P. The role of Akkermansia muciniphila in obesity, diabetes and atherosclerosis. *J Med Microbiol*. 2021 Oct;70(10). doi: 10.1099/jmm.0.001435. PMID: 34623232.

Henkrich, M., Járay, T. Mikrobiom elemzés során alkalmazott 16S rRNS gén V1-V3 és V4 ampikon szekvenálási módszerek összehasonlítása. SZTE TTIK Biológia TDK, 2022. Témavezetők: dr. habil. Tombácz Dóra egyetemi docens és Gulyás Gábor PhD-hallgató, SZTE SZAOK Orvosi Biológiai Intézet.

Hitch, T.C.A., Hall, L.J., Walsh, S.K. et al. Microbiome-based interventions to modulate gut ecology and the immune system. *Mucosal Immunol* 15, 1095–1113 (2022). <https://doi.org/10.1038/s41385-022-00564-1>

Huang, W., Wang, S., Jacobs-Lorena, M. Use of Microbiota to Fight Mosquito-Borne Disease. *Frontiers in Genetics*, 11 (2020). DOI=10.3389/fgene.2020.00196

Indira M, Venkateswarulu TC, Abraham Peele K, Nazneen Bobby M, Krupanidhi S. Bioactive molecules of probiotic bacteria and their mechanism of action: a review. *3 Biotech*. 2019 Aug;9(8):306. doi: 10.1007/s13205-019-1841-2. Epub 2019 Jul 24. PMID: 31355115; PMCID: PMC6656846.

Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol*. 2016 Nov 25;17(1):239. doi: 10.1186/s13059-016-1103-0. Erratum in: *Genome Biol*. 2016 Dec 13;17(1):256. PMID: 27887629; PMCID: PMC5124260.

Josefsen, M. H., Andersen, S. C., Christensen, J., Hoorfar, J. Microbial food safety: Potential of DNA extraction methods for use in diagnostic metagenomics. *Journal of Microbiological Methods* 114, 30-34 (2015) doi = <https://doi.org/10.1016/j.mimet.2015.04.016>

Kerrigan, Z, Kirkpatrick, J. B., D'Hondt, S. Influence of 16S rRNA Hypervariable Region on Estimates of Bacterial Diversity and Community Composition in Seawater and Marine Sediment. *Frontiers in Microbiology*, Volume 10, 2019. DOI=10.3389/fmicb.2019.01640

Kameoka, S., Motooka, D., Watanabe, S. et al. Benchmark of 16S rRNA gene amplicon sequencing using Japanese gut microbiome data from the V1–V2 and V3–V4 primer sets. *BMC Genomics* 22, 527 (2021). <https://doi.org/10.1186/s12864-021-07746-4>

Li, H. Minimap2: pairwise alignment for nucleotide sequences, *Bioinformatics*, Volume 34, Issue 18, 15 September 2018, Pages 3094–3100, <https://doi.org/10.1093/bioinformatics/bty191>

Nechvatal, J. M., Ram, J. L., Basson, M. D., Namprachan, P., Niec, S. R., Badsha, K. Z., Matherly, L. H., Majumdar, A. P. N., Kato, I. Fecal collection, ambient preservation, and DNA extraction for PCR amplification of bacterial and human markers from human feces. *Journal of Microbiological Methods*, Volume 72, Issue 2 (2008) Pages 124-132, ISSN 0167-7012, <https://doi.org/10.1016/j.mimet.2007.11.007>.

Menzel, P., Ng, K. & Krogh, A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7, 11257 (2016). <https://doi.org/10.1038/ncomms11257>

Monteiro, L., Bonnemaison, D., Vekris, A., Petry, K. G., Bonnet, J., Vidal, R., Cabrita, J., Mégraud, F. Complex polysaccharides as PCR inhibitors in feces: *Helicobacter pylori* model. *Journal of Clinical Microbiology*, 35, 4, 995-998, (1997), DOI = {10.1128/jcm.35.4.995-998.1997}

NIH Human Microbiome Portfolio Analysis Team. A review of 10 years of human microbiome research activities at the US National Institutes of Health, Fiscal Years 2007-2016. *Microbiome* 7, 31 (2019). <https://doi.org/10.1186/s40168-019-0620-y>

Ravi, R.K., Walton, K., Khosroheidari, M. (2018). MiSeq: A Next Generation Sequencing Platform for Genomic Analysis. In: DiStefano, J. (eds) Disease Gene Identification. Methods in Molecular Biology, vol 1706. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7471-9_12

Rhoads A, Au KF. PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics. 2015 Oct;13(5):278-89. doi: 10.1016/j.gpb.2015.08.002. Epub 2015 Nov 2. PMID: 26542840; PMCID: PMC4678779.

Salonen, A., Nikkilä, J., Jalanka-Tuovinen, J., Immonen, O., Rajilić-Stojanović, M., Kekkonen, R. A., Palva, A., de Vos, W. M. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: Effective recovery of bacterial and archaeal DNA using mechanical cell lysis. Journal of Microbiological Methods, 81, 2, 127-134 (2010) doi = {<https://doi.org/10.1016/j.mimet.2010.02.007>}

Santiago, A., Panda, S., Mengels, G. et al. Processing faecal samples: a step forward for standards in microbial community analysis. BMC Microbiol 14, 112 (2014). <https://doi.org/10.1186/1471-2180-14-112>

Schlunzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., Yonath, A. Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution, Cell, Volume 102, Issue 5, 2000, Pages 615-623, ISSN 0092-8674, [https://doi.org/10.1016/S0092-8674\(00\)00084-2](https://doi.org/10.1016/S0092-8674(00)00084-2).

Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., Doudna Cate, J. H. Structures of the Bacterial Ribosome at 3.5 Å Resolution. Science. Volume 310, number 5749, pages 827-834, year 2005. doi = {10.1126/science.1117230}

Sender, R., Fuchs, S., Milo, R. (2016) Revised estimates for the number of human and bacteria cells in the body. PLOS Biology. <https://doi.org/10.1371/journal.pbio.1002533>

Sorbara, M.T., Pamer, E.G. Microbiome-based therapeutics. Nat Rev Microbiol 20, 365–380 (2022). <https://doi.org/10.1038/s41579-021-00667-9>

Venegas, D. P., De la Fuente, M. K., Landskron, G., González, M. J., Quera, R., Dijkstra, G., Harmsen, H. J. M., Faber, K. N., Hermoso, M. A. (2019) Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. Frontiers in Immunology. Volume 10. DOI=10.3389/fimmu.2019.00277 ISSN=1664-3224

Weston AD, Hood L. Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. J Proteome Res. 2004 Mar-Apr;3(2):179-96. doi: 10.1021/pr0499693. PMID: 15113093.

Xu, G., Zhang, L., Liu, X. et al. Combined assembly of long and short sequencing reads improve the efficiency of exploring the soil metagenome. BMC Genomics 23, 37 (2022). <https://doi.org/10.1186/s12864-021-08260-3>

Zhang H, Jain C, Aluru S. A comprehensive evaluation of long read error correction methods. BMC Genomics. 2020 Dec 21;21(Suppl 6):889. doi: 10.1186/s12864-020-07227-0. PMID: 33349243; PMCID: PMC7751105.

Zhang J, Ding X, Guan R, Zhu C, Xu C, Zhu B, Zhang H, Xiong Z, Xue Y, Tu J, Lu Z. Evaluation of different 16S rRNA gene V regions for exploring bacterial diversity in a eutrophic freshwater lake. Sci Total Environ. 2018 Mar 15;618:1254-1267. doi: 10.1016/j.scitotenv.2017.09.228. Epub 2017 Oct 28. PMID: 29089134.

A 36. OTDK
BIOLÓGIA SZEKCIÓ FELHÍVÁSÁNAK
2. SZÁMÚ MELLÉKLETE

TÉTELES SZERZŐI HOZZÁJÁRULÁS

Ábrák forrása

1. ábra:

<https://pubmed.ncbi.nlm.nih.gov/?term=gut+microbiome&timeline=expanded>

2. ábra:

<https://biology.stackexchange.com/questions/54823/what-causes-the-variable-conserved-structure-in-the-16s-rna-gene>

3. ábra:

<https://blogs.iu.edu/ncgas/2021/09/17/introduction-to-illumina-sequencing/> alapján, magyar nyelvre átírva

4. ábra: <https://www.youtube.com/watch?v=E9-Rm5AoZGw> (A) és <https://www.yourgenome.org/stories/pacbio-and-third-generation-sequencing/> (B) alapján, magyar nyelvre átírva

5. ábra:

https://files.zymoresearch.com/datasheets/ds1706_zymobiomics_microbial_community_standards_data_sheet.pdf alapján, magyar nyelvre átírva

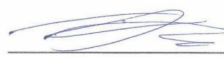
A 6-15. ábrák saját képek (TapeStation, 7-8. ábrák) és vizualizációk (Phyloseq, 9-10 és 12-15. ábrák, EPI2ME, 10. ábra)

Alulírott Henrich Miksa nyilatkozom, hogy a 36. OTDK-ra benyújtott Intesztinális mikrobiom-elemzési munkafolyamatok optimalizációja című pályamunkámban leírtak esetében az alábbi munkafolyamatok elvégzésében vettem részt:

Részfeladat*	A részvétel mértéke (ötfokozatú skálán* jellemezve)
Kutya székletminta begyűjtése	1
A minták előkészítése és feldolgozása, DNS izoláció	2
PCR amplifikációk, könyvtárkészítés, szekvenálás	2
Bioinformatikai elemzések és ábrakészítés: DADA2, Epi2ME, Kaiju, Minitax, Phyloseq	4
Manuális adatelemzés, adatok kiértékelése és interpretációja	4
Referencia irodalom áttekintése	5
Pályamunka és előadás kidolgozása	5

Kelt:

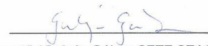
(Kérjük a dőlttel jelölt részeket értelemsszerűen adják meg. A táblázat szükség szerint értelemsszerűen módosítható).


Pályamunkát benyújtó hallgató (Henrich Miksa)
aláírása


Témavezető (dr. Tombác Dóra, SZTE SZAOK Orvosi Biológiai Intézet)
aláírása

A nyilatkozatban foglaltak helyességét megerősítem:

A nyilatkozatban foglaltak helyességét megerősítem:


Témavezető (Gulyás Gábor, SZTE SZAOK Orvosi Biológiai Intézet)
aláírása