

Szegedi Tudományegyetem

Természettudományi és Informatikai kar

Genetikai Tanszék

36. OTDK Biológia Szekció

Archaikus DNS módszertanok összehasonlító vizsgálata

Kiss Petra

Molekuláris biológia MSc. – I. évfolyam

Témavezető: Dr. Neparáczki Endre Dávid

Szeged, 2022/23

1. Tartalomjegyzék

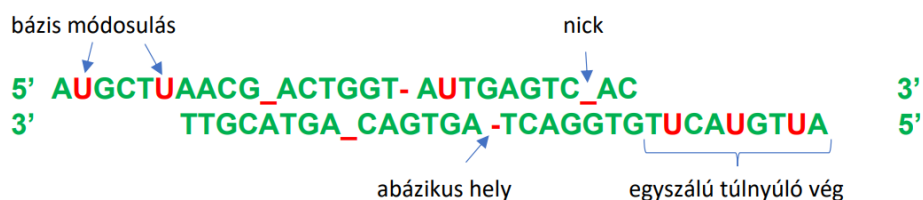
1.	Tartalomjegyzék.....	2
2.	Bevezetés.....	4
2.1.	Az archaikus DNS alapvető tulajdonságai és felhasználása.....	4
2.2.	Az aDNS könyvtárkészítés és szekvenálás főbb fejlődési lépései	5
2.2.1.	Az Illumina szekvenálás.....	6
3.	Célkitűzés	9
4.	Anyagok és módszerek.....	10
4.1.	Mintavétel	10
4.2.	A minták előkészítése a mintavételre	10
4.2.1.	A fogak előkészítése.....	10
4.3.	Mintavétel: a koponya és a sziklacsont feldolgozása - porítása	11
4.4.	DNS kivonás	12
4.4.1.	A minták feltárása	12
4.4.1.1.	A minták feltárása - DNS kivonás csontporból.....	13
4.4.1.2.	A minták feltárása - DNS kivonás egész fogból	13
4.4.2.	DNS tisztítás.....	13
4.4.2.1.	DNS tisztítás szilika mátrixon.....	13
4.4.2.2.	DNS tisztítás MinElute oszlopokkal	15
4.5.	DNS koncentráció mérése	16
4.6.	aDNS könyvtárkészítése új generációs (NGS) szekvenáláshoz	17
4.6.1.	Előamplifikálás.....	17
4.6.2.	Kettős indexálás	18
4.7.	Szekvenálás	19
4.7.1.	Átlagos fragmentméret meghatározás	20
4.7.2.	Shotgun szekvenálás	20
4.7.3.	NGS könyvtár komplexitás becslése Picard's Tools programmal	21
5.	Eredmények és értékelésük	23
5.1.	DNS kivonási módszerek és források összehasonlítása	23
5.2.	Az NGS könyvtár készítési módszerek összehasonlítása.....	24

5.3. Teljes genomszekvenálás	25
5.4. Fragment eloszlás	27
6. Összefoglalás	28
7. Köszönetnyilvánítás	30
8. Irodalomjegyzék.....	31
9. Mellékletek.....	33
9.1. Melléklet: az NGS könyvtárak és azok részleges genomszekvenálása során kinyert információik	33
9.2. Melléklet: a teljes genomszekvenálás eredményei.....	35
9.3. Melléklet: Tételes szerzői hozzájárulás.....	36

2. Bevezetés

2.1. Az archaikus DNS alapvető tulajdonságai és felhasználása

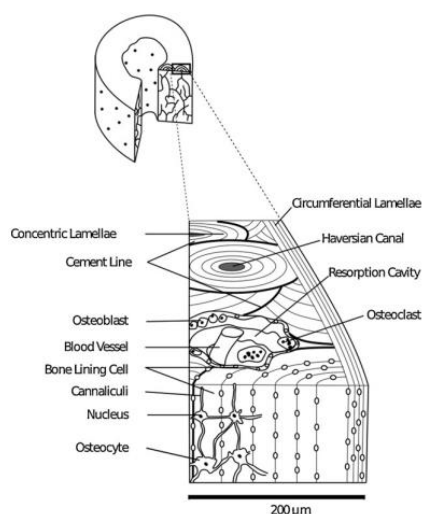
A DNS a halált követően akár több százezer éven keresztül is megőrződhet a csontokban (Rohland et al., 2018). Az évek során az aDNS számos biológiai, fizikai, kémia változáson megy keresztül. Jellemzően 150-60 bázispár méretűnél kisebb fragmentekre töredezik (Prüfer et al., 2010), abázikus helyeket, bázis módosulásokat, egyesszálú túlnyúló végeket és egyesszálú töréseket tartalmaz (1. ábra). Ezek a tulajdonságok és mivel nagyon kis mennyiségben őrződik meg csontmaradványokban, az aDNS munkáknak speciális követelményei vannak. A pre-amplifikációs munkálatok szigorúan steril laboratóriumban zajlanak (aDNS labor), ami a lehető legtávolabb helyezkedik el a poszt-PCR (molekuláris) labortól (Knapp et al., 2012).



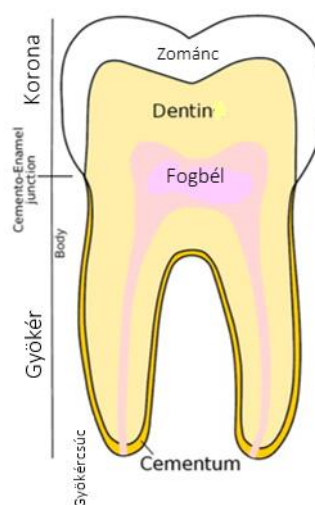
1. ábra – post-mortem (halál után bekövetkező) DNS degradáció

Humán szekvenciákból igazságügyi, antropológiai, evolúció biológiai, környezeti és régészeti kérdéseket tudunk megválaszolni (Orlando et al., 2021).

Csöves csont esetén a DNS kivonás a tömör állományból történik, mivel a kollagén rostok közé szorult osteocyták őrzik meg legjobban a DNS-t (2. ábra). Az endogén DNS megőrződése a sziklacsonthoz (*pars petrosa*) a legjobb, de a foggyökér cementum rétege (3. ábra) is jó forrásként szolgáltathat (Campos et al., Adler et al., 2011).



2. ábra – a csontszövet felépítése, az azt alkotó sejtek (Campos et al., 2012)



3. ábra- humán nagyörlő – a fog szövetei és régiói (Adler et al., 2011)

2.2. Az aDNS könyvtárkészítés és szekvenálás főbb fejlődési lépései

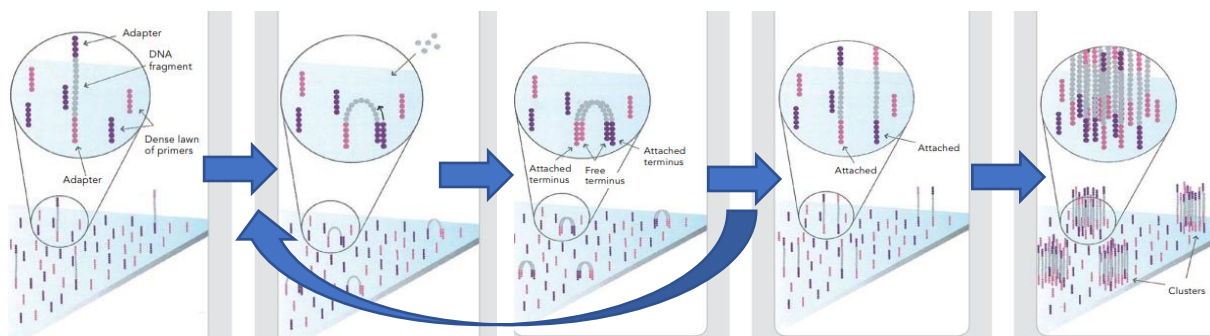
A nagy áteresztő képességű szekvenálás (high-throughput sequencing = HTS) megjelenése előtt, maximum 1000 bázispárnyi méretű fragmentek szekvenálására volt lehetőség a klasszikus, Sanger szekvenálással. Az első aDNS könyvtárakat úgy készítették, hogy az aDNS fragmenteket bakteriális plazmidokba klónozták, és több független klónt is megszekvenáltak (Orlando et al., 2021). Ez a folyamat sok munkával járt és időigényes volt, valamint ma már azt is tudjuk, hogy az így kinyert szekvenciák nagy része szennyező forrásból származott. (Rizzi et al., 2012). Ezt követően polimeráz láncreakció (PCR) segítségével történt az aDNS feldúsítása (Rizzi et al., 2012), ez jelentősen meggyorsította a munkafolyamatokat, azonban a kontamináció elkerülésére így is nagyon nehéz, mert az amplifikáció egyszerre indulhat a szennyező és az endogén forrásról, így az ezzel a módszerrel kinyert információk jelentős részének hitelessége megkérdőjelezhető. A HST-t megelőzően alkalmazott Sanger szekvenálás, egyidőben csupán egyetlen klónozott vagy amplifikált fragment szekvenálására képes. A Sanger szekvenálás során a vizsgált egyszálú DNS fragmenthez hibridizált primer meghosszabbításával a polimeráz enzim a komplementer szálnak megfelelő fluoreszcens színekkel jelölt nukleotidokat épít be. A lánc szintetizálását mindegyik nukleotidnál random beépülő dideoxinukleotid molekulák szakítják meg, így különböző hosszúságú szintézis termékek keletkeznek, melyek meghatározott nukleotidban végződnek. A dideoxi molekulák abban különböznek a normál nukleotidoktól, hogy 3' végükön nem rendelkeznek OH csoporttal, így nem képesek a következő nukleotiddal foszfodiészter kötés kialakítására. Ezután a fragmenteket kapilláris elektroforézisen futtatják, ahol egy lézer és egy szenzor segítségével leolvasták a fluoreszcens jeleket (Jeremy S., 2015).

Archaikus DNS-ből a 2000-es évek elején az új generációs szekvenálás (next-generation sequencing – továbbiakban NGS) megjelenésével vált lehetségessé megbízható, akár teljes genom szekvenciák kinyerése (Orlando et al., 2021). Az NGS fő jellemzője, hogy a szekvenálás magas teljesítményű, azaz kis mennyiségű ultrarövid fragmentekből képes több millió lókuszt egyidejű szekvenálására (Orlando et al., 2021). Erre ma már számos protokoll és kit áll rendelkezésünkre. Két fő NGS könyvtárkészítési módszer ismeretes, az egyik a kettős szálú (dsDNS), a másik pedig egyszálú (ssDNS) DNS-ből indul ki. Laborunkban a kettős szálú, kétszeresen indexált DNS könyvtár készítést használjuk, melynek részletes leírása a 4.6. pont alatt olvasható. Az aDNS szekvenálására az Illumina termékei bizonyultak legalkalmasabbnak. Általánosságban elmondható az NGS könyvtárakról, hogy készítésük során úgynevezett

Illumina specifikus adapter szekvenciákat ligálunk a DNS fragmentekhez. A szekvenálást mindig megelőzi egy alacsony ciklusszámú, lineáris PCR amplifikáció, mert a csekély mennyiségű aDNS könyvtárból egyébként nem lehet szekvenálható koncentrációt elérni. A nagy kapacitású szekvenátorokban több mint a DNS-ét egyszerre szekvenáljuk, de mivel az adapterek mindegyik mintára specifikus, 4-8 nukleotid hosszúságú, ismert szekvenciájú egyedi index kombinációkat tartalmaznak, ezek leolvasásával az egyes mintákhoz tartozó szekvenciák elkülöníthetők (Orlando et al., 2021).

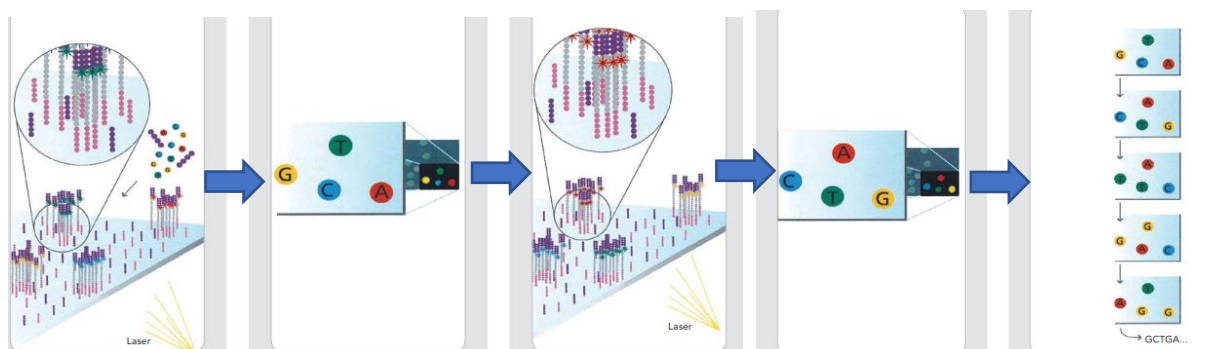
2.2.1. Az Illumina szekvenálás

Ez a technológia egy ciklus alatt több millió párhuzamos, egyenként 75-300 bázis pár (bp) méretű fragment szekvenciájának leolvasására képes. A rendszerbe épített nagy felbontású kamera minden ciklusban lefotózza a fluoreszcensen jelölt nukleotidok beépülését (sequencing by synthesis, SBS). A szekvenáló felszínt flow-cellnek nevezik, ez olyan kihorgonyozott oligonukleotidokat tartalmaz, amelyek komplementerek a könyvtárkészítés során használt adapter molekulákkal, így az egyszálú könyvtár DNS fragmentek véletlenszerűen kikötődnek azokhoz az egyik végüknél. Ezt követően híd amplifikáció történik, ami során a kikötött egyszálú DNS fragment mellé a polimeráz komplementer szálát szintetizál. Ezt további lokális amplifikációs lépések követik, végül a flowcell minden pontján egy DNS cluster képződik, ahol mindegyik cluster egyetlen molekula egy helyre tömörült klonja (4.ábra) (Illumina Inc. 2009).



4. ábra: híd amplifikáció folyamata: kikötés → híd amplifikáció → kettőszálú fragmentek → denaturálás → clusterek kialakulása (Illumina Inc. 2009)

A szekvenálás a ciklikusok reverzibilis termináción alapszik. Minden egyes ciklus során a clusterek mindegyik molekulája egyetlen fluoreszcensen jelölt nukleotiddal hosszabbodik. Amikor egy jelölt nukleotid beépül, a szintézis megáll, és megtörténik a nukleotid leolvasása. Ezután a fluoreszcens molekulát és a láncterminációs csoportot lehasítják, lemossák, majd az egész ciklust megismétlik 70-150-szer. Mivel minden molekula clusterben van jelen, így a fluoreszcens jel felerősödik, és számítógépes segítséggel, lézerekkel gerjesztve detektálhatóvá válik (5.ábra) (Illumina Inc. 2009).



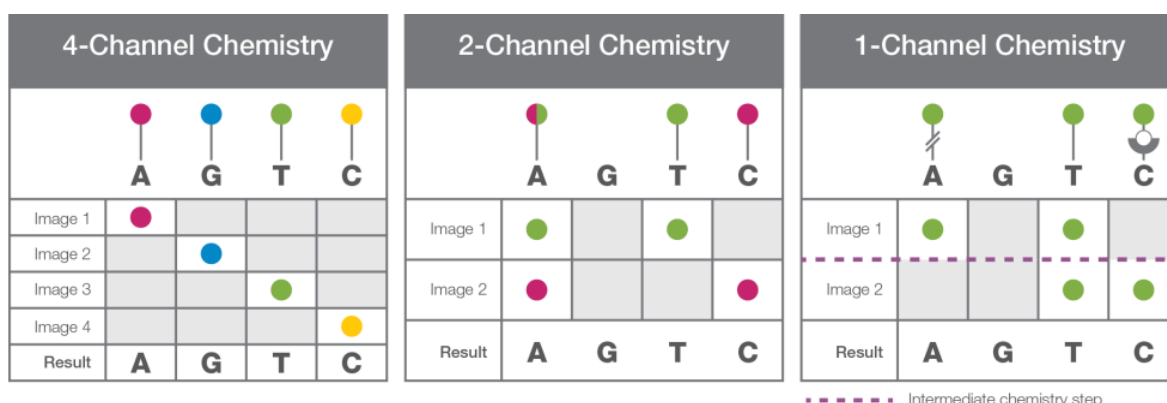
5. ábra: 1. ciklus - fluoreszcens nukleotid beépülése → leolvasás → 2. ciklus – fluoreszcens nukleotid beépülése → leolvasás → ismétlődés a teljes fragment szekvenciájának meghatározásáig. (Illumina Inc. 2009)

Az szekvenálási kémia alapján három eljárást különböztetünk meg attól függően, hogy hány féle fluoreszcens jelölést és képcsatornát használnak ciklusonként a beépült nukleotid leolvasására (6. ábra).

A négycsatornás SBS kémia egy ciklus során négy különböző színű jelölést és négycsatornás képet használ. Minden bázis (adenin=A, guanin=G, timin=T és citozin=C) eltérő színnel jelölt, és amikor egy nukleotid beépült, négy különböző hullámhosszágon gerjesztve négyféle színű fotót készít a rendszer, majd azokat kielemezve megállapítja, hogy melyik clusterbe pontosan melyik színnek megfelelő nukleotid épült be (Illumina Inc. 2018).

A kétszatornás SBS kémia csak kétféle festéket, pirosat és zöldet használ, és kétszatornás képet készít ciklusonként. Az egyes nukleotidok eltérő jelölése itt úgy valósul meg, hogy a timin zöld fluoroforral jelölt, a citozin pirossal, az adenin mindkettővel egyszerre, a guanin pedig egyikkel sem (Illumina Inc. 2018).

Az egycsatornás SBS kémia, ami ötvözi a komplementer fém-oxid félvezető (complementary metal-oxidesemiconductor, CMOS) technológiát a szintézis alapú szekvenálással, minden ciklusban egyetlen fluorescens festéket és két kémiai lépést használ, valamint két egymást követő képet készít mindkét reakció után. A timinnek állandó fluoreszcens jelölése van, a guanin nincs festve, az adeninnek eltávolítható jelölése van, ami csak az első fotón látszik, a citozin pedig egy linker csoporttal rendelkezik, ami a második lépésben képes megkötni a jelölést, így csak a második fotón látható (Illumina Inc. 2018). Egycsatornás kémiát használnak jelenleg az Illumina gyártó Novaseq6000 és iSeq 100-as szekvenátorai.



6. ábra – négycsatornás kémia (balról), kétszatornás kémia (középen), egycsatornás kémia (jobbról) (Illumina Inc. 2018).

3. Célkitűzés

A kutatás célja a NGS könyvtárak minőségi becslésének kiegészítése egy új objektív mérőszámmal. Ennek érdekében különböző archaikus DNS kivonási és könyvtárkészítési protokollok szekvenciaadatokon történő összehasonlítása. A különböző módszerekkel készült aDNS kivonatok és újgenerációs szekvenáló könyvtárak objektív összehasonlítását azok fragment eloszlása, endogén DNS tartalma és komplexitás értéke alapján végezzük. A laborunkban eddig a könyvtárak jellemzésére nem használtuk a Picard's tools szoftver komplexitás értékét ezért ennek megbízhatóságát kísérletes alapon vizsgáltuk.

Ezek segítségével ajánlást tudunk tenni a jelenleg elérhető legjobb aDNS extrakciós és könyvtár készítő módszerre, továbbá arra, hogy használható-e alacsony lefedettségű szekvenálás után a Picard's tools szoftver komplexitás értéke.

4. Anyagok és módszerek

4.1. Mintavétel

A mintavétel II-III századi szarmata korú csontokból történt. A minták lelőhelye a Hódmezővásárhelyi Fehértó környéke, ahol összesen 22 szarmata sírt tártak fel. A minták megőrződésüket tekintve antropológiailag jobb és rosszabb minősítést kaptak, valamint csontozatuk és sírmellékleteik alapján meghatározták a valószínű antropológiai nemüket is (Párducz. 1946-48). A vizsgálatba három jobb és három rosszabb megtartású leletet vontunk be, amit az 1. táblázat foglal össze.

Minta Neve	Lelőhely	Sírszám	Leltári szám	Megőrződés	Antropológiai nem
HVFTper3	Hódmezővásárhely - Fehértó	3	1661	jobb	Nő
HVFTper4	Hódmezővásárhely - Fehértó	4	1662	jobb	Férfi
HVFTper15	Hódmezővásárhely - Fehértó	15	1673	jobb	Nő
HVFTper2	Hódmezővásárhely - Fehértó	2	1659	rosszabb	Nő
HVFTper10	Hódmezővásárhely - Fehértó	10	1668	rosszabb	?
HVFTper19	Hódmezővásárhely - Fehértó	19	1676	rosszabb	Férfi

1. Táblázat: A minták nevei, lelőhelye, antropológiai jellemzők alapján a megtartása és a nemük.

4.2. A minták előkészítése a mintavételre

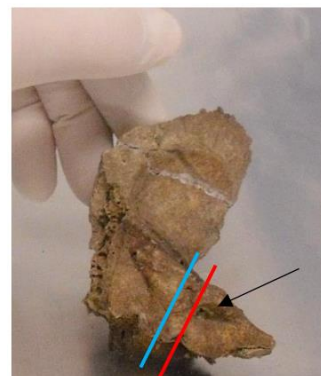
A mintákat az erre a célra kialakított steril laborban dolgoztuk fel. A laborba való beléptetés során hipóval alaposan áttöröltem a csontokat, majd 2x30 percig (az első 30 perc elteltével a mintákat megfordítottam) UV sugárzásnak vettem alá (Fulton, 2012). A hipó a felszíni exogén eredetű DNS csökkentésében játszik szerepet, azonban az UV sugárzás hozzájárul a felszín közeli szennyező DNS mennyiségének redukálásához is (Fulton, 2012).

4.2.1. A fogak előkészítése

A fogak esetében nem volt szükség porításra, mivel a foggyökér cement rétegéből vontuk ki a DNS-t (Harney et al., 2020). Kihúzást követően átmostam őket hipóval majd Ultra Pure vízzel, ezután az elszívó fülkében hagytam őket megszáradni. Száradást követően a fogkoronákat parafilmmel körbetekertem, így biztosítva, hogy a DNS kivonás csak a foggyökérből történjen.

4.3. Mintavétel: a koponya és a sziklacsont feldolgozása - porítása

A beléptetést követően a koponyák vagy koponya darabok feldolgozása a csontporító helyiségben történt. Laborunkban jellemzően két módszert alkalmazunk a koponyából történő mintavételre, attól függően, hogy milyen állapotban van az adott minta. Mindkét esetben speciális elszívó fülke alatt dolgozunk, illetve mindkét módszerhez Dremel 3000-es eszközt használunk, amire gumikesztyűt húzunk a keresztfertőzés elkerülésének az érdekében. A fülkéket, az egyéb munkafelületeket, illetve az összes használt eszközt minden feldolgozott minta után DNS mentesítjük

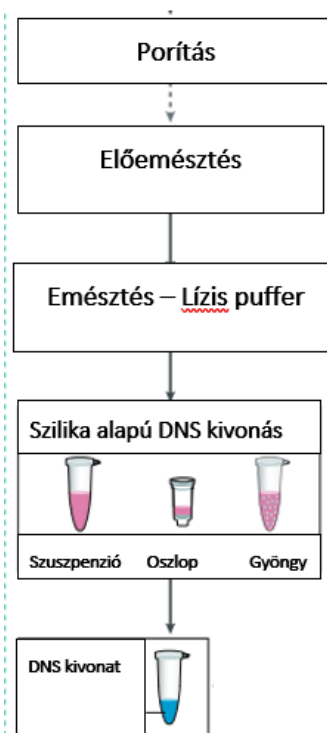


7. ábra: példa a sziklacsonton tett vágások helyének bemutatására

hipóval és 3%-os hidrogén-peroxiddal (H_2O_2), Ultra Pure vízzel átmoszuk azokat, majd a fülkéket és az eszközöket 20 perces UV sugárzásnak tesszük ki. Intakt koponya esetén a Benoit és társai (Benoit et al., 2012) által kidolgozott módszert alkalmazzuk, ami a sziklacsont fúrásával jár. Ez a módszer nem okoz jelentős roncsolódást a koponyán. Amennyiben a koponya szétesett és ezáltal a sziklacsont jól hozzáférhető, vagy ha sziklacsont letört, akkor vágással tisztítjuk körbe a belső fülnek megfelelő csontrészt. Ezzel a módszerrel a maradvány jelentősen roncsolódik, de így több és tisztább csontport tudunk kinyerni. A vágást oldal irányú elszívás mellett végezzük, és az elszívott csontpor kidobásra kerül. Vágás előtt drótkefés fejjel megtisztítottam a sziklacsont teljes felületét, ezután a Dremel eszközre gyémántvágó korongot helyeztem és a belső hallónyílás (*porus acusticus internus*) (7. ábra - nyíl) felett néhány mm-re (7. ábra – piros vonal) végig vágtam a csontot, majd kb. 0,5 cm-el feljebb (7. ábra – kék vonal) is megismételtem a vágást. Ezzel egy olyan csontszeletet kapunk, melynek közepén a sziklacsont található. Ezt a szeletet fúrófejjel tovább tisztítottam, a szivacsos állománytól maximálisan megfosztottam, valamint igyekeztem a különféle elszíneződéseket és egyéb szennyeződésre utaló jeleket is eltávolítani. A tisztított sziklacsont szeletet őrlőmalomban maximális rázófrekvencián 30 másodpercig porítottam. Ezt a módszert Gamba és társai dolgozták ki (Gamba et al., 2014).

4.4. DNS kivonás

A szakirodalomban több archaikus DNS-re optimalizált kivonási módszert is közöltek, és melyek mindegyike a 8. ábrán bemutatott főbb lépésekből áll. Először fel kell tártani a csont mátrixot, hogy a DNS az oldatba kerüljön. Ez két lépésből áll, egy előemésztésből és az emésztésből. Feltárást követően a DNS az oldatba kerül, ezt követően kötő pufferrel kikötjük a DNS-t szilika alapú mátrixhoz, a mátrixot etanolos oldattal mossuk, majd a DNS-t leoldjuk a mátrixról. A módszerek közötti legfőbb különbség abban rejlik, hogy milyen szilika-alapú anyaghoz történik a DNS kikötése (Orlando et al., 2021). Az egyik esetben egy szilika szuszpenziót használunk, másik esetben gyárilag elkészített szilika oszlopokat alkalmazunk. Labormunkám során mindkét eljárást alkalmaztam. Laborunkban kezdetben az olcsóbb szilika szuszpenziós módszert alkalmazták, de ma már az egyszerűbben kezelhető MinElute oszlopos módszert preferáljuk: nagyobb hatékonysága, és gyorsabb eljárás módja miatt. Azonban a szuszpenziós eljárásnak is vannak előnyei, például sokkal kisebb fragmentek megkötésére is képes lehet. Egy harmadik elterjedt módszernél a DNS adszorpciója szilikával bevont mágnesgyöngyökhöz történik, azonban ezzel a technikával nem dolgoztam.



8. ábra – aDNA kivonás alapvető lépései (Orlando et al., 2021)

Összesen három sorozat kivonást végeztünk sziklacsontból és két sorozatot fogakból. A fogak esetében kizárólag a MinElute oszlopos eljárást alkalmaztuk. Minden esetben minden kivonási lépést egy csontpor mentes negatív kontrollal is elvégeztünk.

4.4.1. A minták feltárása

Minden DNS tisztítást megelőzően egy 30 perces előemésztés, amikor a csontport alkotó szemcsék felszíne, valamint a foggyökerek felülete emésztődött rövid ideig, ezzel eltávolítva a szemcsék, illetve a foggyökerek felszínén maradt esetleges szennyeződéseket. Ezt a folyamatot minden esetben egy 72 órás emésztés követte, ami során a DNS kioldódik, és vizes oldatba kerül. A felhasznált protokollok alapjául (Rohland & Hofreiter, 2007b) és (Gamba et al., 2014) által kidolgozott leírások szolgáltak.

4.4.1.1. A minták feltárása - DNS kivonás csontporból

MinElute-os tisztítás esetén a DNS extrakció 200 mg csontporból történt, míg szilika mátrixos tisztítás esetén 260 mg csontporból indultunk ki.

A csontport előemésztettük 1ml EDTA (etilén-diamin-tertaecetsav) (0,5 M) és 5µl Proteináz-K (20mg/ml) keverékében 48°C-os rázóinkubátorban 1500 rpm-es fokozaton. 30 perc inkubálást követően a mintákat 1 percig 13000 rpm-en (maximális fordulatszámon) centrifugáltuk. A felülúszó leöntését követően a leülepedett csontporra mintánként 1,5 ml extrakciós puffert mértünk rá. A puffert 6 minta esetében összesen 8,25 ml EDTA (0,5M), 111 µl Proteináz-K (20mg/ml), 90 µl Triton X-100 (1%) és 549 µl ddH₂O alkotta. A centrifugacsövek felső részét parafilmmel körbetekertük, ezzel megakadályozva az oldat esetleges párolgását a viszonylag hosszú inkubációs idő alatt. Vortexelést követően a mintákat 48°C-os rázóinkubátorban 1500 rpm-es fokozaton 72 órán át emésztettük.

4.4.1.2. A minták feltárása - DNS kivonás egész fogból

A már DNS kivonásra előkészített fogakra annyi oldatot mértünk, hogy éppen ellepje a foggyökeret. Az extrakciós negatív kontrollt is beleszámolva kivonásonként összesen hét mintával dolgoztunk. A mastermix összesen 7 ml EDTA-t (0,5 M) és 35 µl Proteináz-K-t (20 mg/ml) tartalmazott, ezt mértük szét a fogakra. A minták 30 perces 48°C-os inkubálása hibridizációs kamrában történt. Az inkubációs idő leteltével leöntöttük a fogokról a folyadékot. és extrakciós puffert mértünk szét rájuk, ami összesen 6416,7 µl EDTA-t (0,5M), 86,3 µl Proteináz-K-t (20 mg/ml), 70 µl Triton X-100-t (1%) és 430 µl ddH₂O-t tartalmazott. Mindegyik mintára annyit mértünk, hogy a puffer csak a foggyökereket lepje el, majd parafilmmel lezártuk a centrifuga csövek tetejét. Az emésztés ideje a fogak esetében is 72 óra volt 48°C-on, azonban az előemésztéshez hasonlóan az inkubáció ennél a lépésnél is a hibridizációs kamrában történt.

4.4.2. DNS tisztítás

4.4.2.1. DNS tisztítás szilika mátrixon

A tisztítás lényege, hogy a DNS-t szilika alapú mátrixhoz kötjük, ami képes megkötni akár még az ultrarövid (>25 bp) DNS fragmenteket is (Rohland et al., 2018). Ezzel a módszerrel csak a sziklacsontból vontunk ki DNS-t. A protokoll alapjául (Rohland & Hofreiter, 2007a) leírása

szolgált. Mivel a szilika szuszpenzió érzékeny a tárolásra, a kivonáshoz szükséges szilika szuszpenziót a kivonást megelőző napon készítettük.

Szilika szuszpenzió készítése: 2.4 g szilika port (Sigma, cat. no. S5631) felszuszpendáltunk 20 ml DNS mentes vízben, majd egy órán át sötétben (mivel a szilika fényérzékeny) szobahőmérsékleten ülepedni hagytuk. Ezt követően a felülúszóból 19,5 ml-t átpipettáztunk egy tiszta falcon csőbe, majd további 4 órán keresztül sötétben szobahőn ülepítettük a mintát. A felülúszóból 17 ml-t lepipettáztunk, így megszabadulva a túl nagy és a túl kicsi szemcséktől. A maradék 2,5 ml üledékkel dolgoztunk tovább. 24 µl 30%-os sósavval beállítottuk a szuszpenzió pH-át, majd 850 mikroliterenként szétmértük a szuszpenziót, amit a felhasználásig sötétben szobahőn tároltunk.

DNS szilika mátrixhoz kötés: A kötőpuffer 45 ml végérfogatra készült, és tartalmazott 25,065g 5,83 M GuHCl-ot (guanídium-hidroklorid), 1575 µl 105mM-os NaOAc-ot (nátrium-acetát), 2115 µl 46,8% izopropanolt, 26,1 µl Tween-20-at, amit ddH₂O-al felöntöttünk 45 ml-re. Az emésztett mintákról 2 percig 13000 rpm-en történő centrifugálás követően leöntésre került a felülúszó, ezzel dolgoztunk tovább. A felülúszókhoz 150 µl szilika szuszpenziót és 6-6 ml kötő puffert adtunk. Ecetsavval beállítottuk a pH-t 4 és 6 közé. Kb. 110 µl ecetsav hozzáadása után az összes minta pH-ja 5 körüli volt. A falcon csöveket parafilmmel lezártuk és 3 órán át sötétben forgattuk. Ezt követően 5 percig 3000-es fordulatszámon centrifugáltuk a csöveket, hogy leülepedjen a szilika a hozzá kötődött DNS-el. Leöntöttük a felülúszót, majd további 1 ml kötő puffert pipettáztunk a szilika üledékre, azzal alaposan felszuszpendáltuk az üledéket és 2 ml-es centrifugacsövekbe vittük át.

Mosás: 1 percig maximális fordulatszámon (13000 rpm) centrifugáltuk a szilika szuszpenziót, majd leöntöttük a felülúszót. A szilika üledékhez 800 µl mosópuffert mértünk, ami ebben az esetben 80%-os etanol (EtOH) volt, ebben alaposan lassan felszuszpendáltuk a felülúszót, és újra 1 percig maximális fordulatszámon centrifugáltuk a mintákat. A teljes mosási lépést újra megismételtük, majd a mosó puffer eltávolítása után a csöveket elfektetve a szilika üledéket 15 percig fülke alatt nyitott tetővel szobahőmérsékleten szárítottuk.

DNS leoldás: a szárított szilikához 50 µl 50°C-ra előmelegített TE puffert mértünk, és alacsony fordulatszámon vortexeltük, amíg a szilika felszuszpendálódott. Ezt követően 48°C-os termosztátban inkubáltuk a mintákat 10 percig úgy, hogy néhány percenként újravortexeltük őket, ezzel megakadályozva a szilika leülepedését. Az inkubációs idő leteltével 1 percig

maximális fordulatszámon centrifugáltuk, a DNS tartalmú felülúszót tiszta 1,5 ml-s centrifuga csövekbe vittük át, az üledéken pedig megismételtük a DNS leoldást. A teljes DNS tartalmú felülúszó térfogata összesen 100 µl volt, amit a teljes szilikamentesítés céljából újra centrifugáltunk, végül 1,5 ml-s LoBind centrifugacsövekbe vittük át és -20°C-on tároltuk a következő lépésig.

4.4.2.2. DNS tisztítás MinElute oszlopokkal

A DNS tisztítás QIAGEN MinElute oszlopokhoz gyárilag hozzákötött szilika mátrix segítségével történt, melyhez a gyárilag előállított oldatokat használtunk. Ezt a fajta kivonást az összes mintából kétszer végeztük el, egyrészt összehasonlítás céljából, másrészt hogy több DNS-t kinyerjünk, a többféle módszerrel tervezett NGS könyvtár készítéshez. A tisztítás QIAvac vákuumos tisztító rendszer (a vacuum manifold) használatával történt, ami minimális előkészületeket igényel. A MinElute oszlopokat a vákuumos tisztítórendszerhez steril csatlakozók segítségével hozzákapsoltuk, azonban az oszlopok térfogata túl kicsi a tisztításhoz használt oldatok térfogatához képest, ezért az oszlopokhoz steril infúziós harangokat erősítettünk parafilm segítségével.

DNS kötése az oszlopokhoz: A 72 órán át 48°C-on inkubált mintákat 1 percig 13000-es fordulatszámon centrifugáltuk, majd a felülúszókat egyenként 15 ml-es feliratozott falcon csövekbe öntöttük. A felülúszókhöz 5x-ös mennyiségű kötő puffert (PB puffer) öntöttünk, csontporhoz tartozó felülúszó (1,5 ml) esetén ez 7,5 ml-t jelentett mintánként, fogak esetében mintánként eltérő érték volt. A keveréket az előkészített és vákuum rendszerre felhelyezett infúziós harangokba öntöttük. Akkor helyeztük az oszlopokat vákuum alá, amikor a nyomás - 900mB alá csökkent. A minták átfolyását követően, az oszlopról levettük a harangokat, rátettük a kupakokat, és a MinElute gyűjtőcsövekbe helyeztük őket.

**Megjegyzés:* ezt a tisztítási technikát a könyvtárkészítések során is alkalmaztuk, azonban jelentősebb kisebb mennyiséggel dolgoztunk, így a protokolltól eltérően nem a vákuum rendszert használtunk, hanem centrifugáltuk a mintákat maximum fordulatszámon 1 percig.

Mosás: a mintákra 750 µl mosó puffert (PE puffer) pipettáztunk (minden esetben ezt a mennyiséget alkalmazzuk), majd 13000 rpm-en (maximális fordulatszámon) centrifugáltuk őket. Az összegyűlt folyadékot kiöntöttük a gyűjtőcsövekből és újra átcentrifugáltuk a mintákat, ezzel teljesen eltávolítva a mosó puffert, és megszáritva az oszlopokat. Ez azért fontos lépés,

mert a mosó puffer magas koncentrációban etanolt tartalmaz és annak esetleges jelenléte akadályozhatja a DNS leoldódását, valamint a későbbi enzimatikus reakciókat.

Leoldás: az oszlopokat tiszta centrifugacsövekbe helyeztük, majd közvetlenül a szilika mátrixra 25 µl 50°C-ra előmelegített elúciós puffert (EB puffer) mértünk. 2 percig pihentettük a mintákat, hogy a puffer leoldja a DNS-t majd 1 percig maximális fordulatszámon centrifugáltuk azokat. Ezt követően megismételtük a leoldást. A leoldott DNS oldat össztérfogata 50 µl volt, melyeket a következő lépésig 1,5 ml-s centrifugacsövekben -20°C-on tároltuk.

4.5. DNS koncentráció mérése

A DNS kivonást követően nagyon fontos a minták DNS tartalmának ellenőrzése, mert ha a minták nem tartalmazznak detektálható mennyiségű DNS-t, akkor azokkal a munkát nem érdemes folytatni. Ha a kivonás sikeres volt, akkor a negatív kontroll nem tartalmazhat DNS-t, ami azt igazolja, hogy a munka során nem szennyeztük be a felhasznált oldatokat, valamint vélhetően magukat a mintákat sem. Az DNS koncentráció mérésére Qubit® Fluorometert használtunk. A mérés lényege, hogy egy fényérzékeny fluoreszcens reagenst adunk az oldathoz, ami a kettősszalú DNS-spirálba interkalálódik, így a műszer által érzékelt fluorometriás jel erőssége arányos a DNS koncentrációjával. A méréshez rendelkezésünkre áll egy kit (Qubit dsDNA HS Assay Kit), ami tartalmazza a fluoreszcens reagenst, a reakcióhoz használandó puffert, valamint a műszer kalibrálásához két standard DNS koncentrációjú mintát, melyek egyikében a DNS koncentráció nagyon alacsony, míg a másikban magas. MinElute oszlopokon tisztított minták esetében 1 µl, míg szilika mátrixon tisztított minták esetében 6 µl mintán végeztük a DNS koncentrációmérést. A gyártó által javasolt protokollt alapján állítottuk össze a mérést. Először a mintákból kimértük a fentebb említett 1 vagy 6 µl-es mennyiségeket kifejezetten erre a célra használt Qubit csövekbe, valamint két külön csőbe 10-10 µl-t a DNS standardokból. Ezután $n \times 199$ µl Qubit puffert mértünk ki egy 15 ml-es falcon csőbe, és ebbe $n \times 1$ µl fluoreszcens reagenst pipettáztam. Mindkét esetben $n = \text{minták száma} + 2 \text{ standard minta}$. Szétmértük a reagenst tartalmazó puffert a mintákra úgy, hogy minden csőben a végtérfogat 200 µl legyen ($\rightarrow$ 190 µl a standard mintákra, 194-199 µl a DNS kivonatokot tartalmazó Qubit csövekbe). A műszert beállítottuk (dsDNS, high sensitivity), kalibráltuk a standardokra, megadtuk a DNS tartalmú minták kiindulási térfogatát, és megmértük egyesével a minták DNS koncentrációját, amit a műszer ng/µl-ben jelenít meg.

4.6. aDNS könyvtárkészítése új generációs (NGS) szekvenáláshoz

A kitisztított DNS-ből szekvenáló könyvtárat készítettünk, melynek során enzimátikus folyamatok révén szekvenálásra alkalmas állapotba kerülnek a DNS darabok. A könyvtár készítés során a DNS fragmentekhez specifikus adaptereket ligálunk. Az adapterek mintánként eltérő rövid index (jelölő) szekvenciákat tartalmaznak, melynek segítségével az egyszerre szekvenált nagyszámú minta különválasztható. Háromféle könyvtár készítési módszert alkalmaztunk, ezek mind megegyeztek a kettős indexálás lépéséig. Egyetlen módszernél egy opcionális lépést is alkalmaztam, az úgynevezett előamplifikálást. A laborunkban alkalmazott protokoll alapját több közölt módszer képezte. A felhasznált protokollok (Meyer & Kircher, 2010) és (Kircher, 2012) leírásain alapszanak, kiegészítve az első részleges UDG (Uracil-DNS glükóziláz) kezelés lépést, amit (Rohland et al., 2015) alapján végeztünk.

A könyvtárkészítés minden alkalommal 18 mintával történt egyszerre:

- sziklacsontból kivont MinElute oszlopokon tisztított DNS minták (6 db)
- fogból kivont MinElute oszlopokon tisztított DNS minták (6 db)
- sziklacsontból kivont szilika mátrixhoz kötötten tisztított DNS minták (6db)

4.6.1. Előamplifikálás

Ez egy opcionális lépés, melynek során a könyvtár molekulákat még az indexálást megelőzően egy alacsony ciklusszámú PCR reakcióval felsokszorozzuk, hogy elegendő mennyiség maradjon az esetleges későbbi vizsgálatokhoz. A reakcióban univerzális primereket használunk, amelyek az adapterekhez kötődnek, valamint olyan polimeráz enzimet, ami képes átírni a molekulákban bennmaradt uracilokat is. Ezt a lépést csupán a Kappa Hifi enzimmal készített könyvtáraknál alkalmaztuk.

A reakcióelegy mintánként tartalmazott 15 µl ddH₂O-t, 15 µl GoTaq puffert (5x), 7,5 µl MgCl₂-ot (25 mM-os), 7,5 µl dNTP mixet (deoxynucleotide mix (dATP, dCTP, dGTP és dTT)) (egyenként 2.5 mM-os), 3,75 µl BSA-t (Bovine serum albumin) (10mg/ml), 0,375 µl G2GoTaq enzimet (5U/ µl), 4,5-4,5 µl IS7 és IS8 primereket (10 µM) és 15 µl DNS templátot, így a végtérfogat mintánként 75 µl volt, amit egyenlő arányban 25 mikroliterenként szétmértünk három PCR csőbe. A beállított PCR program a következő volt: 96°C 6 percig, 94°C 30 másodpercig, 58°C 30 másodpercig, 72°C 30 másodpercig, 64°C 10percig, majd 4°C -on tartás. A három 30 másodperces lépes 12 cikluson keresztül ismétlődött. A program lejártával, a PCR

csövek kinyitása előtt a mintákat átvittük a molekuláris laborba és a munkát ott folytattuk tovább. Az azonos mintákat tartalmazó csövek tartalmát egyesítettük, majd MinElute oszlopokon a fentebb leírtakkal azonos módon tisztítottuk a mintákat. A DNS leoldás 25 µl EB pufferrel történt. Ezt követően megmértük a minták koncentrációját a 4.5.-ös pontban leírtak szerint.

4.6.2. Kettős indexálás

Ezen lépés során a DNS fragmentek mindkét végéhez olyan adaptereket kapcsolunk, amelyek mintánként egyedi index kombinációkat tartalmaznak. Az indexek specifikus 4-8 nukleotid hosszúságú egyedi, ismert szekvenciák, általuk képesek vagyunk azonosítani és szétválasztani szekvenálás során az egy egyedhez tartozó DNS szekvencia olvasatokat. Az rendelkezésünkre álló indexek kombinálása tetszőleges, a lényeg, hogy a szekvenálandó minták mindegyike egyedi indexkombinációval rendelkezzen. Az indexet tartalmazó adaptereket PCR reakcióval visszük fel a molekulák végeire (Orlando et al., 2021).

A) Kettős indexálás Accuprime Pfx Supermix-szel:

Ezen enzim használatakor nem alkalmazunk előamplifikálást, az adapter feltöltés után közvetlenül indexáltuk a könyvtárat. Ennek előnye, hogy egy PCR reakcióban amplifikáljuk a DNS molekulákat és különböző index-el jelöljük.

A reakcióelegy mintánként tartalmazott 58,5 µl Accuprime Pfx Supermixet, 3 µl BSA-t (10 mg/µl), 1,5 – 1,5 µl P5 és P7 indexet (primert) és 10,5 µl DNS templátot. Mintánként a végtérfogat 75 µl volt, ezt egyenlő arányban három PCR csőbe szétmértük és úgy helyeztük be a PCR gépbe. A PCR beállítása a következő volt: 95°C 5 percig, 95°C 15 másodpercig, 60°C 30 másodpercig, 68°C 30 másodpercig, 68°C 5 percig, majd 4°C-on tartás. A 15-30 másodperces lépések 12 cikluson át ismétlődtek. A program lejártával egyesítettük az azonos index kombinációval rendelkező 3 reakciót, majd MinElute oszlopokon tisztítottam azokat a fentebb leírt protokoll alapján. Az elúció 20 µl EB pufferrel történt. Megmértük a könyvtárak koncentrációját a fent leírt Qubit módszerrel, majd további felhasználásig -20°C-on tároltam azokat.

B) Kettős indexálás G2GO Taq Polimerázzal

Előamplifikálás ebben az esetben sem történt, az adapter feltöltést közvetlenül az indexálás követte.

A reakcióelegy a következő összetevőkből állt mintánként: 28,5 µl ddH₂O, 20 µl GoTaq puffer (5x), 10 µl MgCl₂ (25 mM-os), 10 µl dNTP mix (egyenként 2.5 mM-os), 5 µl BSA (10mg/ml), 0,5 µl G2GoTaq enzim (5U/ µl), 6-6 µl P5 és P7 primerek (10 µM) és 14 µl DNS templát. A végtérfigat mintánként 100 µl volt, amit 4 PCR csőbe egyenlő arányban szétmértük. A PCR program a következő volt: 96°C 6 percig, 94°C 30 másodpercig, 58°C 30 másodpercig, 72°C 30 másodpercig, 64°C 10percig, majd 4°C -on tartás. A három 30 másodperces lépés 12 cikluson keresztül ismétlődött. A program lejártával az azonos mintákat tartalmazó csövek tartalmát egyesítettük, majd MinElute oszlopokon a fentebb leírtakkal azonos módon tisztítottam a mintákat. A DNS leoldás 20 µl EB pufferrel történt. Ezt követően megmértük a könyvtárak koncentrációját Qubit módszerrel a fentebb leírtak alapján, majd további felhasználásig -20°C-on tároltuk azokat.

C) Kettős indexálás KAPA® HiFi HotStart ReadyMix-szel

Ebben az esetben az adapter feltöltést követően előamplifikáltuk a mintákat és csak azután következett a minták indexálása. A Kappa polimeráz nem képes átírni az uracilokat, így csak előamplifikált könyvtárhoz használható. Emellett ez az enzim enyhe nukleáz aktivitással is rendelkezik, így magas templát és primer koncentrációt igényel. Minden mintából azonos mennyiségű, 50 ng DNS templátot amlifikáltunk. Az előamplifikálás során mért koncentrációkból kiszámolható, hogy melyik mintából hány mikroliterre van szükség az 50 ng-os mennyiség elérésének céljából, ezt a mennyiséget X-el jelölöm.

A reakcióelegy mintánként 50 µl végtérfigatú és a következőket tartalmazta: X µl DNS templát (50 ng), 15-X µl ddH₂O, 25 µl KAPA® HiFi HotStart ReadyMix (2X), ami tartalmazza az enzimet is, 5-5 µl P5 és P7 indexek. A PCR program a következő: 98°C 3 percig, 98°C 20 másodpercig, 66°C 10 másodpercig, 72°C 15 másodpercig, majd 77°C további 30 másodpercig. A 20-10-15 másodperces lépések 5 cikluson át ismétlődtek. A program befejeztével a mintákat a fentebb leírtak szerint MinElute oszlopokon tisztítottuk, és 20 µl EB pufferrel eluáltuk a DNS-t, majd Qubit műszerrel megmértük könyvtárak koncentrációját. A további felhasználásig -20°C-on tároltuk őket.

4.7. Szekvenálás

Laborunkban az összes mintát a harmadik generációs Illumina iSeq 100 szekvenátorral szekvenáltuk, ami egycsatornás kémiával dolgozik. Alacsony readszámú (mintánként ~100,000

pair end read) shotgun szekvenálást végeztünk, ami által információt kaptunk többek között a minták endogén és exogén DNS tartalmáról, valamint az egykori személy genetikai nemét is megbecsültük. A shotgun szekvenálási stratégia lényege, hogy véletlenszerű, nagy számú genomdarabokat szekvenálunk, majd azokat a referencia genomhoz illesztjük (Motahari et al., 2013).

Az endogén DNS tartalom meghatározásához az összes kinyert szekvencia darabot a humán referencia genomra illesztjük, így az illeszthető szekvenciák aránya adja az endogén humán DNS arányt, míg a nem illeszthető szekvenciák külső, exogén szennyező forrásból származnak.

Minden minta esetében a legmagasabb endogén tartalommal bírót tovább küldtük alacsony lefedettségű teljes genomszekvenálásra is, amit a NovaSeq6000 eszközzel végeztünk.

4.7.1. Átlagos fragmentméret meghatározás

Az elkészült könyvtárak fragment méret eloszlásának meghatározása - azaz minőségük ellenőrzése - elengedhetetlen fontosságú, mielőtt a könyvtárakat megszekvenáljuk. Ennek mérésére az Aligent Tapestation 4200-as eszközt használjuk, ami magas érzékenységgű DNS Screen Tape (mikro gélelektroforézis) segítségével képes ezen információk meghatározására. A könyvtárak minőségellenőrzésén felül a fontossága abban rejlik, hogy ha meghatározzuk a DNS fragmentek méret eloszlását egy mintában, akkor az alapján megbecsülhető a minta DNS tartalma nM-ban, ami alapján kikalkulálható az az olvasat szám amit a mintából szeretnénk kapni a szekvenálás során.

A már Qubit-tal lemért, ismert koncentrációjú DNS könyvtárakból 1 µl -t 10, 100 vagy 1000-szeresre hígítottuk, hogy a bennük lévő DNS koncentrációja 10-1000 pg/µl legyen. Ezután a hígított mintákból 2-2 µl-t átvittünk a műszerrel kompatibilis plate-be, melyekbe már ki volt mérve 2-2 µl fluoreszcens puffer. 1 percig 2000 rpm-en vortexeltük a mintákat, majd a gépbe helyezve az önállóan elvégezte a futtatást, majd az egyes könyvtárak fragment eloszlásának a meghatározását.

4.7.2. Shotgun szekvenálás

Felhasznált kit: A szekvenáláshoz az Illumina cég által gyártott kit-et, az iSeq 1000 i1 Reagent v2 Single csomagot használtuk, ami egy futáshoz szükséges reagensekkel már előre megtöltött patron (cartridge-et) és egy flow cell-t tartalmazott. A flow cellt 4°C-on, a cartridge-et -20°C-on tároltuk, a felhasználást megelőző napon tettük szobahőre őket, ahol egy napig állni hagytuk.

Minták előkészítése a szekvenálásra: Először a mintákat a Qubit koncentráció és a TapeStation által meghatározott fragment eloszlás alapján 1 nM-osra hígítottuk. Ennek érdekében a 100ng/μl-t megközelítő vagy azt meghaladó koncentrációjú mintákból 1,5 μl-t, az attól alacsonyabb koncentrációjú minták esetében pedig 3 μl-t felhasználva végeztük el a hígításokat. Ezután mindegyik mintából 10-10 μl-t kivettünk és összemértük őket egyetlen mikrocentrifuga csőben, és alapos vortexelést követően tovább hígítottuk 84 μl ddH₂O és 16 μl könyvtár keverék arányban. Az így kihígított mixből 20 μl került szekvenálásra.

A szekvenátor beállítása: A gépben a következő paramétereket programoztuk: generate fastq modul → index reads:2, paired-end reads, read hossz:151, library preparation kit: custom. Egy ciklus alatt 150 nukleotid leolvasása történt, a futás alatt ez összesen 8 millió paired-end olvasatot jelent.

A kinyert nyers szekvenciák feldolgozását bioinformatikus végezte, aki a Cut Adapt szoftvert használva a szekvenciákról levágta megmaradt adapter részeket (Martin, 2011.), majd a szekvencia olvasatokat a humán referencia genomra illesztette a Burrows-Wheeler Aligner programot használva (Li & Durbin, 2009). Ebből kiszámította a mintánk endogén DNS tartalmát. A PCR duplikátumok eltávolításához és a további adatok kinyeréséhez, mint a könyvtár komplexitás becsléséhez a Picard's Tools programot használta (Broad Institute, 2016).

4.7.3. NGS könyvtár komplexitás becslése Picard's Tools programmal

A Picard's Tools programmal lehetőség nyílik egy olyan minőségi mérőszám megadására, ami a könyvtár egyedi molekuláinak a számát, azaz a könyvtár komplexitását becsüli meg (Broad Institute, 2016). A PCR-, és optikai duplikátumok miatt ez egy igen hasznos mérőszám lehet, hisz hiába a több molekula, ha azok csupán egymás másolatai.

A becslés az olvasat párok szekvenciája alapján történik. A rossz minőségű-, valamint a nem párosodott olvasatok kiszűrésre kerülnek, a program azokat nem veszi figyelembe. A molekulákat akkor tekinti duplikátumnak, ha legfeljebb 3% az eltérés közöttük. Ezután az eredmény pontosítása érdekében az adatok további szűrőkön mennek keresztül (Broad Institute, 2016).

A számítás alapja a Lander-Waterman-egyenlet: $C/X = 1 - \exp(-N/X)$

X = a könyvtárban lévő egyedi molekulák száma, N = összes read szám csökkentve az optikaiduplikátumok számával, C = az egyedi vizsgált read szám (a read párokban megfigyelt egyedi fragmentumok száma csökkentve a PCR és optikai duplikátumok számával)

A program az X értéket határozza meg, azaz a könyvtárban lévő egyedi molekulák számát becsüli meg.

Ezt egy úgy teszi az algoritmus, hogy egy függvény segítségével keresi azt az r faktort (m, M) egy intervallumban (ahol $m=1$ az elején és M -re igaz, hogy $M \cdot C > X$), amivel ha megszorozzuk C -t megkapjuk a valós X -et. Az r -t úgy becsüli, hogy $1 < r < M$, amivel veszi az m és M range középértékét, ahol $r = (m+M)/2$.

Azaz feltételezi, hogy $X = r \cdot C$ és kiszámolja a $C/X - 1 + \text{EXP}(-N/X)$ képlet értékét.

Amennyiben ez $= 0$, akkor pont megkaptuk azt az r értéket, amivel C -t megszorozva megkapjuk X -et. Ha ez nagyobb nulla akkor tudjuk hogy az r az kisebb mint a jó faktor, így a valós értéke r és M között van (r, M), azaz ilyenkor az alsó limitet (m) növeljük $m=r$, majd újra számoljuk $r = (m+M)/2$ -t, ha kisebb mint nulla akkor tudjuk, hogy az r az nagyobb mint a jó faktor, így a valós értéke (m, r) rangeban van, ilyenkor a felső limitet (M) csökkentjük, $M=r$, majd újra számoljuk az $r = (m+M)/2$ -t.

Ezt 40 iterációban közelítjük, és a végén az algoritmus visszaadja a $C \cdot (m+M)/2$ -öt mint a legjobb értéket amire igaz, hogy $C/X - 1 + \text{EXP}(-N/X) \sim 0$.

5. Eredmények és értékelésük

Összesen 6 egyénből származó mintával dolgoztunk. Mind a hat egyén sziklacsonitját sikeresen porítottuk, illetve minden egyénből a rendelkezésünkre állt fog minta is. A DNS kivonás minden esetben sikeresnek bizonyult a porított sziklacsonitokból MinElute oszlopokhoz és szilika mátrixhoz kötött módszerrel is, valamint a fogakból végzett extrakció is sikerült. A negatív kontrollok egyike sem tartalmazott DNS-t. Az NGS könyvtárkészítés az összesen 18 DNS extraktumból mintánként 3-3 enzimet alkalmazva szintén sikeres volt, így $3 \times 18 = 54$ aDNS könyvtár került részleges shotgun szekvenálásra, majd egyénenként a legmagasabb endogén DNS tartalommal bírók (= 6 DNS könyvtár) teljes genomszekvenálására is sor került.

5.1. DNS kivonási módszerek és források összehasonlítása

A 2. táblázatban látható a minták DNS koncentrációja. Sziklacsonitból MinElute oszlopon tisztítva kétszer vontunk ki DNS-t (sziklacsonit A és sziklacsonit B). A koncentráció különbségek ezek esetében adódhatnak az elúciós lépésből.

Minta:	DNS koncentráció ng/μl					Megőrződés:
	MinElute				Szilika	
	fog	Újraemésztett fog	sziklacsonit A	sziklacsonit B	sziklacsonit	
HVFTper3	0,112	0,118	0,714	1,05	0,198	jó
HVFTper4	0,156	1,15	0,64	0,638	0,417	jó
HVFTper15	0,41	0,368	1	0,61	0,0963	jó
HVFTper2	0,35	0,492	14	16,6	2,97	rossz
HVFTper10	2,08	0,42	0,922	0,668	0,64	rossz
HVFTper19	0,244	0,458	0,496	0,33	0,229	rossz

2. Táblázat – a minták DNS koncentrációja, és antropológiailag megállapított megőrződése

Látható, hogy az antropológiailag megállapított jobb és rosszabb megőrződésű minták DNS koncentrációja nem mutat szignifikáns különbséget. Ennek az a legvalószínűbb magyarázata, hogy a rossz megőrződésű minták hasonló, vagy akár magasabb DNS tartalma jórészt exogén eredetű. A táblázatban feltüntetett koncentráció értékek csupán informatív jellegűek a kivonás sikerességéről, analízisre nem felhasználhatók.

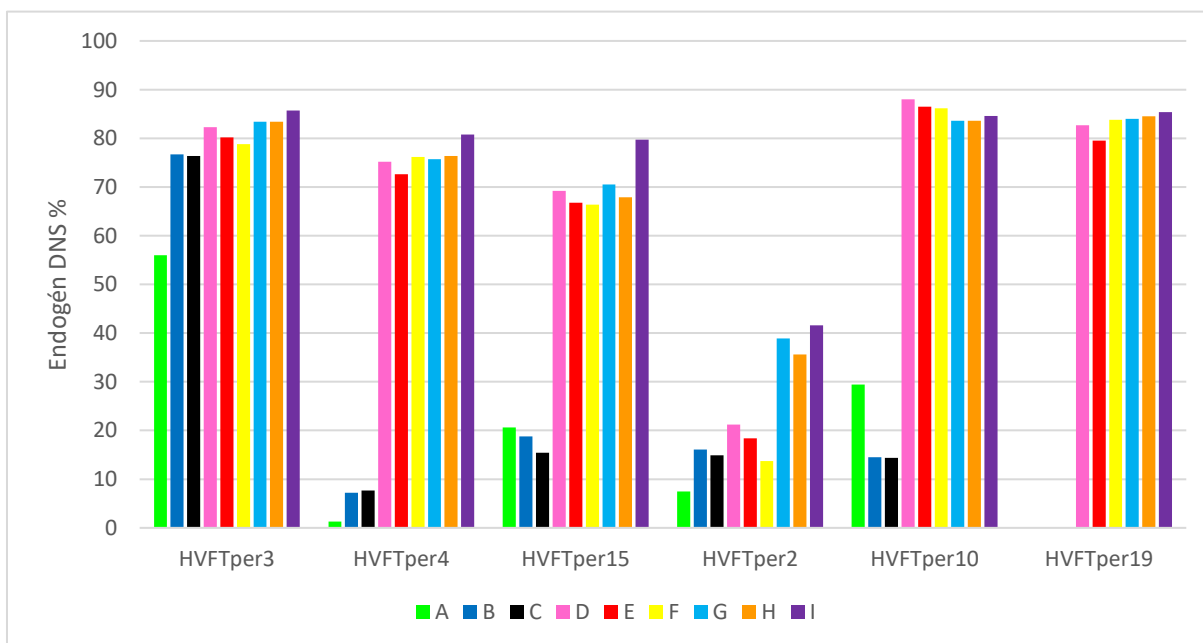
5.2. Az NGS könyvtár készítési módszerek összehasonlítása

A kísérlet során azt vizsgáltuk, hogy könyvtár készítés során mely enzimek a leghatékonyabbak a különböző forrásból származó, különböző módszerrel tisztított DNS minták esetében. A különböző módszerek kombinációit a 3. táblázat foglalja össze, a továbbiakban ezek a jelölések lesznek használatosak. A mért értékek a 9.1. mellékletben megtalálhatók.

Könyvtár jelölése	DNS forrása	DNS kivonás	Könyvtárkészítéshez használt enzim
A	Fog	MinElute oszlop	Kappa Hifi (+preamplifikáció)
B	Fog	MinElute oszlop	Accuprime Pfx. Supermix
C	Fog	MinElute oszlop	G2Go Taq polimeráz
D	Sziklacsont	MinElute oszlop	G2Go Taq polimeráz
E	Sziklacsont	MinElute oszlop	Accuprime Pfx. Supermix
F	Sziklacsont	MinElute oszlop	Kappa Hifi (+preamplifikáció)
G	Sziklacsont	Szilika mátrix	G2Go Taq polimeráz
H	Sziklacsont	Szilika mátrix	Accuprime Pfx. Supermix
I	Sziklacsont	Szilika mátrix	Kappa Hifi (+preamplifikáció)

3. Táblázat: a különböző forrásból, különböző DNS kivonási módszerrel kivont és különböző enzimmel készített aDNS könyvtárak kombinációi és azok jelölése

A 9. ábra összehasonlítja a különböző módszer kombinációkkal készített könyvtárak endogén DNS tartalmát. Öt könyvtár igazolja azt a szakirodalomból ismert tényt, miszerint a sziklacsont endogén DNS hozama az esetek nagy részében jelentősen jobb a fogakénál. A HVFTper3 mintára nem érvényesül ez az állítás, mivel ezesetben a sziklacsontot megközelítő endogén DNS hozamot kaptunk. Megfigyelhető az is, hogy a sziklacsont esetében bár nincs jelentős



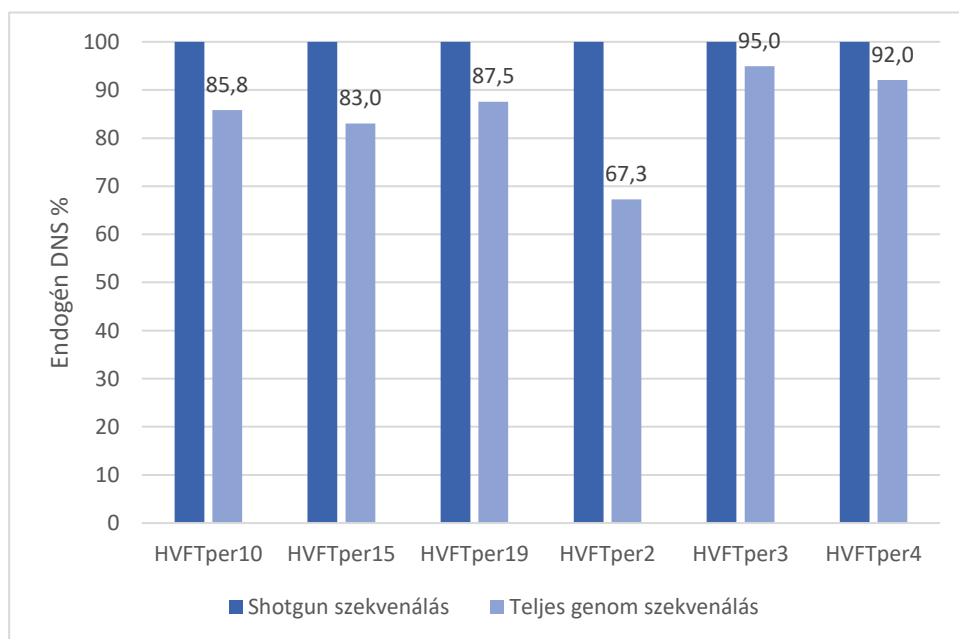
9. ábra: Az NGS aDNS könyvtárak százalékos endogén DNS tartama

különbség az eredmények tekintetében, mégis az I könyvtárak, azaz a sziklacsontról kivont, MinElute oszlopokon tisztított, G2GoTaq polimerázzal előamplifikált, majd Kappa Hifi enzimmel indexált könyvtárak, majdnem minden esetben minimálisan, de jobb eredményt mutatnak az endogén DNS hozam tekintetében a többi módszerrel szemben. Ez egy klasszikusan alkalmazott módszer archaikus minták esetén, azonban ez a módszerkombináció a legidő-, illetve munkaigényesebb.

Továbbá ezen minták eredményei alapján, a három antropológiailag rosszabb megtartásúnak minősített minta (HVFTper2, HVFTper10, HVFTper19) közül, csupán egyetlen esetben, a HVFTper2-nél figyelhető meg szignifikáns különbség az antropológiailag jobb megtartású minták endogén DNS tartalmához képest. Emiatt összességben ezen minták információi alapján nem lehet szignifikáns különbséget tenni, az antropológiailag megállapított jobb és rosszabb megőrződésű minták endogén DNS tartalma között, ebből adódóan megéri minden rendelkezésre álló koponyából DNS mintát venni, még akkor is, ha a csont rossz megőrződésűnek mutatkozik.

5.3. Teljes genomszekvenálás

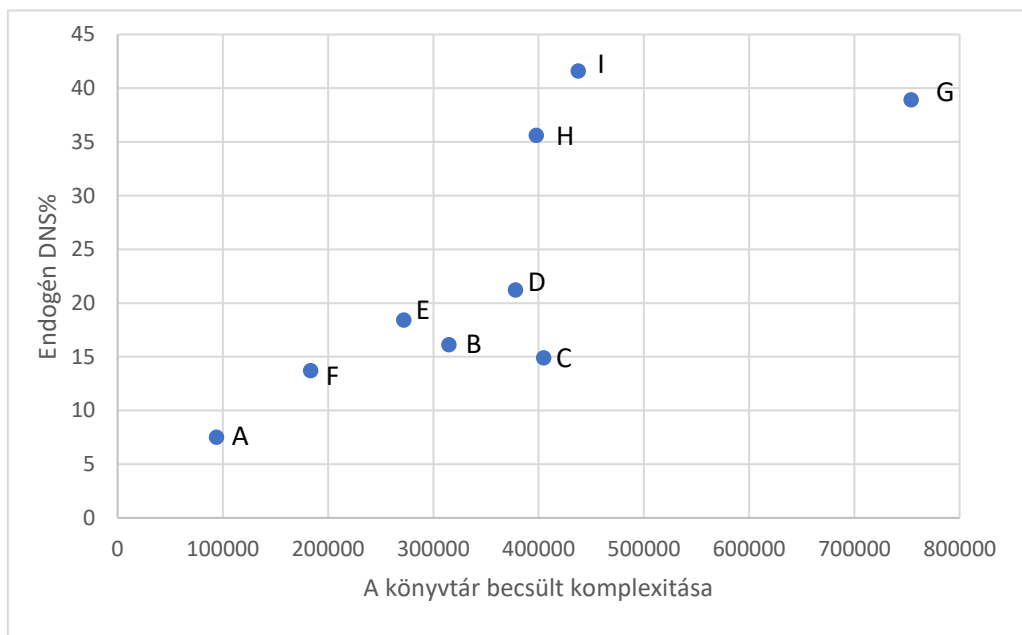
A teljes genomszekvenálásra kiküldött minták kiválasztása minden esetben egyetlen paraméter, a könyvtárak endogén DNS tartalma alapján történt. Öt minta esetében az I könyvtárak kerültek kiküldésre. Kizárólag a HVFTper10 minta esetében a 'D' könyvtár készítési módszerű került teljes genomszekvenálásra.



10. ábra: a részleges-, és teljes genomszekvenálás százalékos endogén DNS hozamának összehasonlítása

A 10. ábrán összehasonlítottuk a minták részleges genomszekvenálása során nyert endogén DNS hozam és a teljes genomszekvenálása során kapott eredményekkel (az értékek a 9.1. és 9.2. mellékletben megtalálhatók) arányát. Az összehasonlítás során a shotgun szekvenálás eredményeit tekintettük 100%-nak és ahhoz képest vizsgáltuk meg, hogy annak mekkora részét teszik ki a teljes genomszekvenálás eredményei. Kivétel nélkül minden esetben azt láthatjuk, hogy alacsony readszámú shotgun szekvenálás túlbecsüli a minta endogén DNS tartalmát, mivel teljes genomszekvenálás esetén alacsonyabb az endogén DNS értéket kaptunk. A legtöbb esetben ez az eltérés 5% és 17% közötti, azonban a HVFTper2 minta esetében az érték 32,7%. Ezen kiugróan magas érték miatt megvizsgáltuk a minta más paramétereit is, amik a 9.1. mellékletben is fel vannak tüntetve. A HVFTper2 minta shotgun szekvenálás során kapott endogén DNS tartalma jelentősen alacsonyabb volt a többihez képest, ezt már a 9. ábra is tükrözte. Ezen értékeket megvizsgáltuk a könyvtárak becsült komplexitásának a függvényében (11.ábra).

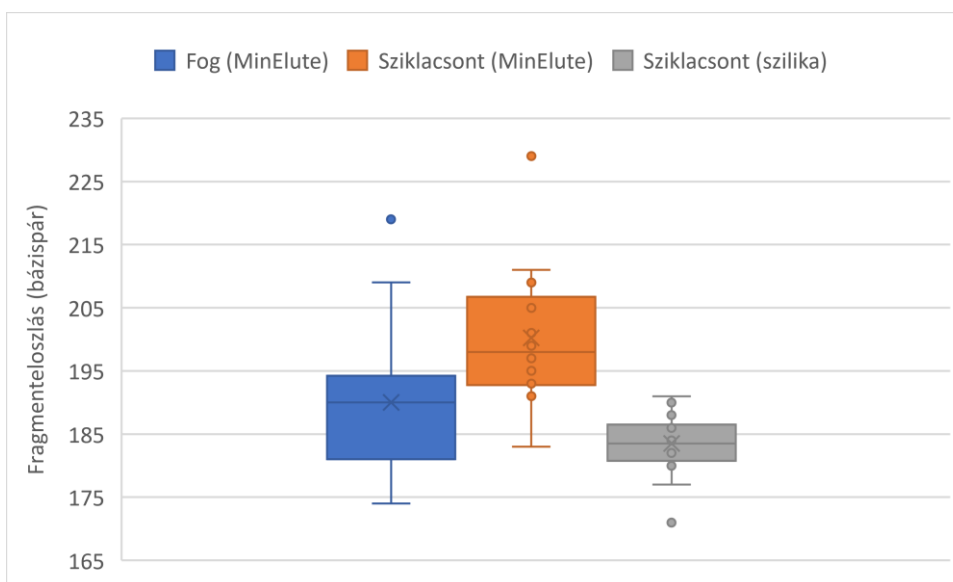
Mivel csupán az endogén DNS hozamot vettük figyelembe a minta kiválasztása során, így a legmagasabb endogén DNS tartalommal bíró könyvtár került teljes genomszekvenálásra, tehát az I könyvtár. A 11. ábrát megfigyelve láthatjuk, hogy az G könyvtár, endogén DNS tartalma csupán kevéssel marad le, azonban a könyvtár becsült komplexitása közel a duplája, így érdemes lehet ezt a mérőszámot is figyelembe venni archaikus minták teljes genom szekvenálása előtt.



11. ábra – A HVFTper2 minta endogén DNS hozama a könyvtárak becsült komplexitásának a függvényében

5.4. Fragment eloszlás

A Tape Station adatokból megtudtuk a könyvtárak fragment eloszlását. A felhasznált információk a 9.1. mellékletben megtalálhatók. A 12. ábrán látható a MinElute oszlopokon tisztított fog minták, a MinElute oszlopokon tisztított sziklacsont por minták és a szilika mátrixhoz kötötten tisztított sziklacsontpor minták fragment eloszlásának összehasonlítása. A MinElute oszlopokon tisztított sziklacsontpor minták átlag fragment mérete (>190) a magasabb, mint a szilika mátrixon tisztított mintáké (<185). Ez azt támasztja alá, hogy valóban igaz a szakirodalmakból ismert információ, miszerint a szilika mátrix kisebb méretű fragmenteket képes megkötni mint a MinElute oszlopok.



12. ábra - NGS könyvtárak fragment eloszlásának ábrázolása forrás és DNS kivonás szerint (balról – Minelute oszlopon tisztított fog minták, középen - MineElute oszlopon tisztított sziklacsontpor minták, jobbról szilika mátrixon tisztított sziklacsontpor minták)

6. Összefoglalás

Szakdolgozati munkám során három-három antropológiailag jobb és rosszabb megtartásúnak minősített leletből vettem fog és sziklacsontpor mintát, majd különböző DNS kivonási és NGS könyvtárkészítési módszereket kombinálva megszekvenáltuk ezeket.

Minden egyes sziklacsontpor mintából kétféle módszerrel vontam ki DNS-t - MinElute oszlopokkal és szilika mátrixhoz kötötten, míg fogak esetében csupán MinElute-os extrakciót alkalmaztam. A kivonás végére összesen 18 különböző DNS extraktummal rendelkeztem – leletenként három-hárommal. Mind a 18 DNS kivonatból három fajta enzimet (Accuprime Pfx Supermix, G2Go Taq polimeráz, Kappa Hifi) alkalmazva készítettem NGS könyvtárat, így összesen 54 könyvtárat készítettem, amik mind shotgun szekvenálásra kerültek. Minden minta esetében a legmagasabb endogén DNS tartalommal rendelkezőt könyvtárat teljes genomszekvenálásra küldtük.

A shotgun szekvenálás eredményei alapján, egy eset kivételével a fogakból kivonható endogén DNS mennyisége valóban szignifikánsan alacsonyabb a sziklacsontéhoz képest, ez alátámasztja a szakirodalmakból ismert állítást. Az antropológiailag jobb és rosszabb megtartásúnak minősített minták endogén DNS hozama között nem feltétlenül tehető különbség, így minden esetben érdemes az összes rendelkezésre álló mintából DNS kivonást végezni. Egy eset kivételével a sziklacsontból kivont, MinElute oszlopokon tisztított, G2GoTaq polimerázzal előamplifikált és Kappa Hifi enzimmel indexált könyvtárak rendelkeztek a legmagasabb endogén DNS tartalommal, így ezek a könyvtárak kerültek teljes genom szekvenálásra. A HVFTper10 esetében ugyanezen forrásból származó és módszerrel tisztított, de G2GoTaq polimerázzal indexált könyvtár bírt legmagasabb endogén DNS tartalommal, így az került szekvenálásra.

Az alacsony lefedettségű teljes genomszekvenálás eredményei alapján a minták valós endogén DNS tartama valamivel alacsonyabb a shotgun szekvenálásban kapott adatokhoz képest. Öt esetben 5-17%-al becsüli túl a részleges genomszekvenálás a mintát, azonban egy esetben több mint 30%-ot tévedett. Ez okot adott azon minta további paramétereinek alaposabb vizsgálatára. Ha figyelembe vesszük a Picard's tools által kiszámolt könyvtárkomplexitás értékét, ami a könyvtárban található egyedi molekulák számára ad becslést, láthatjuk, hogy van olyan könyvtár, ami endogén DNS hozama csupán kevéssel marad le legmagasabbhoz képest, azonban a becsült komplexitása közel a duplája. A többi könyvtárat megvizsgálva az adott szempontból szintén hasonló eredményt mutatnak. Ha csupán a könyvtárak becsült komplexitás

értékeiket vesszük figyelembe, akkor a G2GoTaq polimerázzal indexált könyvtárak minden esetben komplexebbnek mutatkoznak, mint a Kappa Hifi enzimmal készítek. Megjegyzendő, hogy a rendelkezésre álló adatok alapján a teljes genomszekvenálás során becsült könyvtárak komplexitása kivétel nélkül több nagyságrenddel magasabb, mint annak részleges szekvenáláskor becsült értéke. Mivel teljes genomszekvenálás során az elsődleges cél az, hogy a lehető legtöbb egyedi endogén molekulát kapjuk, így a minták teljes genomszekvenálása előtt érdemes lehet figyelembe venni azok endogén DNS tartalma mellett ezt a paramétert is. Ezen feltételezések indokolják a HVFTper2 G módszerrel készült könyvtár alacsony lefedettségű teljes genomszekvenálását is, amely után kísérletesen is igazolható a hipotézisünk.

7. Köszönetnyilvánítás

Köszönöm témavezetőmnek Dr. Neparáczki Endre Dávidnak a sok türelmet, segítséget és szakmai tanácsot, amelyekkel megkönnyítette számomra a dolgozat megírását és az eredmények kiértékelését.

Köszönöm Dr. Török Tibor egyetemi docensnek, a dolgozat elkészítéséhez nyújtott szakmai segítséget.

Köszönöm Kovács Bence, Maár Kitti és Gînguță Alexandra Ph.D. hallgatóknak a rengeteg segítséget és tanítást, az elméleti alapok és a gyakorlati ismeretek átadását.

Köszönöm Maróti Zoltánnak a bioinformatikai analízisek elkészítését.

Továbbá köszönetet szeretnék mondani a kutatócsoport minden tagjának, akikhez bármikor fordulhattam bármilyen kérdéssel.

8. Irodalomjegyzék

- Adler, C. J., Haak, W., Donlon, D., & Cooper, A. (2011). Survival and recovery of DNA from ancient teeth and bones. *Journal of Archaeological Science*, 38(5), 956–964. <https://doi.org/10.1016/j.jas.2010.11.010>
- Benoit, J. N., Quatrehomme, G., Carle, G. F., & Pognonec, P. (2012). An alternative procedure for extraction of DNA from ancient and weathered bone fragments. *Medicine, Science and the Law*. <https://doi.org/https://doi.org/10.1258/msl.2012.012026>
- Campos, P. F., Craig, O. E., Turner-Walker, G., Peacock, E., Willerslev, E., & Gilbert, M. T. P. (2012). DNA in ancient bone - Where is it located and how should we extract it? *Annals of Anatomy*, 194(1), 7–16. <https://doi.org/10.1016/j.aanat.2011.07.003>
- Gamba, C., Jones, E. R., Teasdale, M. D., McLaughlin, R. L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczy, L., Kovári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T. F. G., Hofreiter, M., Bradley, D. G., & Pinhasi, R. (2014). Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5. <https://doi.org/10.1038/ncomms6257>
- Harney, É., Cheronet, O., Fernandes, D. M., & Sirak, K. (2020). A Minimally Destructive Protocol for DNA Extraction from Ancient Teeth. *BioRxiv*.
- Kircher, M. (2012). Analysis of High-Throughput Ancient DNA Sequencing Data. In *Methods Mol Biol* (pp. 197–228). https://doi.org/10.1007/978-1-61779-516-9_23
- Knapp, M., Clarke, A. C., Horsburgh, K. A., & Matisoo-Smith, E. A. (2012). Setting the stage - Building and working in an ancient DNA laboratory. *Annals of Anatomy*, 194(1), 3–6. <https://doi.org/10.1016/j.aanat.2011.03.008>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. <https://doi.org/10.14806/ej.17.1.200>
- Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 5(6). <https://doi.org/10.1101/pdb.prot5448>
- Motahari, A. S., Bresler, G., & Tse, D. N. C. (2013). Information theory of DNA shotgun sequencing. *IEEE Transactions on Information Theory*, 59(10), 6273–6289. <https://doi.org/10.1109/TIT.2013.2270273>

- Orlando, L., Allaby, R., Skoglund, P., der Sarkissian, C., Stockhammer, P. W., Ávila-Arcos, M. C., Fu, Q., Krause, J., Willerslev, E., Stone, A. C., & Warinner, C. (2021). Ancient DNA analysis. *Nature Reviews Methods Primers*, 1(1). <https://doi.org/10.1038/s43586-020-00011-0>
- Prüfer, K., Stenzel, U., Hofreiter, M., Pääbo, S., Kelso, J., & Green, R. E. (2010). Computational challenges in the analysis of ancient DNA. In *Genome Biology* (Vol. 11). <http://genomebiology.com/2010/11/5/R47>
- Rizzi, E., Lari, M., Gigli, E., de Bellis, G., & Caramelli, D. (2012). Ancient DNA studies: New perspectives on old samples. In *Genetics Selection Evolution* (Vol. 44, Issue 1). <https://doi.org/10.1186/1297-9686-44-21>
- Rohland, N., Glocke, I., Aximu-Petri, A., & Meyer, M. (2018). Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nature Protocols*. <https://doi.org/https://doi.org/10.1038/nprot.2007.247>
- Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., & Reich, D. (2015). Partial uracil – DNA – glycosylase treatment for screening of ancient DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1660). <https://doi.org/10.1098/rstb.2013.0624>
- Rohland, N., & Hofreiter, M. (2007a). Comparison and optimization of ancient DNA extraction. *BioTechniques*, 42(3), 343–352. <https://doi.org/10.2144/000112383>
- Rohland, N., & Hofreiter, M. (2007b). Ancient dna extraction from bones and teeth. *Nature Protocols*, 2(7), 1756–1762. <https://doi.org/10.1038/nprot.2007.247>
- Broad Institute, 2016. Picard tools. <https://broadinstitute.github.io/picard/>. utolsó megtekintés: 2022. 11.10. (15:34)
- Jeremy Schoales (2015). How Does Sanger Sequencing Work? <https://www.thermofisher.com/blog/behindthebench/how-does-sanger-sequencing-work/> - utolsó megtekintés: 2022. 03. 11. (15:01)
- Illumina Inc. (2009). Illumina Sequencing Technology http://nextgen.mgh.harvard.edu/attachments/techspotlight_sequencing.pdf - utolsó megtekintés: 2022. 11. 1. (09:56)
- Illumina Inc. (2018). Illumina CMOS Chip and One-Channel SBS Chemistry <https://www.illumina.com/content/dam/illumina-marketing/documents/products/techspotlights/cmos-tech-note-770-2013-054.pdf> - utolsó megtekintés: 2022. 03. 21. (11:31)
- Párducz Mihály (1946-1948). Szarmata temető Hódmezővásárhely–Fehértón. [Nécropole sarmate à Hódmezővásárhely–Fehértó.] *Archaeologiai Értesítő VII–IX.* 283–290. - <http://real-j.mtak.hu/355/> - utolsó megtekintés: 2022. 04. 04. (21:16)

9. Melléklet

9.1. Melléklet: az NGS könyvtárak és azok részleges genomszekvenálása során kinyert információik

Minta ID	Lab ID	DNS forrása	DNS kivonás típusa	DNS koncentráció (ng/ul)	Előamplifikálás	Könyvtár típusa	Előamplifikáció koncentráció (ng/ul)	Könyvtár koncentráció (ng/ul)	Tape Station fragment méret	Olvasatok száma	Humán genomra illesztett olvasat száma	Endogén DNS arány	Duplikátumok nélküli illesztett read szám	Duplikátumok aránya	Becsült könyvtár méret	Átlagos inszert méret	Az inszertek standard devianciája
HVFTper10	HVFTper10acc	petrous bone	MinElute column	0,922	-	Accuprime Pfx. Supermix	-	55	192	189032	163510	0.865	157110	0.039	2081100	60.91	22.38
HVFTper10	HVFTper10Facc	tooth	MinElute column	2,08	-	Accuprime Pfx. Supermix	-	29,6	193	124533	18084	0.145	17552	0.029	332211	68.97	29.14
HVFTper10	HVFTper10Fgo	tooth	MinElute column	2,08	-	G2GO Taq polymerase	-	65,4	191	146455	21038	0.144	20504	0.025	366794	58.41	24.61
HVFTper10	HVFTper10Fpre	tooth	MinElute column	0,42	G2GO Taq polymerase	Kappa Hifi	22,4	65,2	219	94276	27760	0.294	26900	0.031	497912	79.45	30.37
HVFTper10	HVFTper10go	petrous bone	MinElute column	0,922	-	G2GO Taq polymerase	-	59,4	191	231572	203686	0.880	196032	0.038	2557034	51.32	19.17
HVFTper10	HVFTper10pre	petrous bone	MinElute column	0,668	G2GO Taq polymerase	Kappa Hifi	25,8	81	206	111341	95946	0.862	92530	0.036	1336086	69.22	23.00
HVFTper10	HVFTper10SZacc	petrous bone	Silica	0,64	-	Accuprime Pfx. Supermix	-	64,4	182	119363	99784	0.836	96778	0.030	1558742	48.50	19.22
HVFTper10	HVFTper10SZgo	petrous bone	Silica	0,64	-	G2GO Taq polymerase	-	99,6	186	199927	167156	0.836	161844	0.032	2501170	46.98	17.32
HVFTper10	HVFTper10SZpre	petrous bone	Silica	0,64	G2GO Taq polymerase	Kappa Hifi	30,4	62,4	188	136114	115102	0.846	111258	0.033	1722691	52.04	18.43
HVFTper15	HVFTper15acc	petrous bone	MinElute column	1	-	Accuprime Pfx. Supermix	-	10,3	200	143260	95736	0.668	91648	0.043	1113047	67.15	26.63
HVFTper15	HVFTper15Facc	tooth	MinElute column	0,41	-	Accuprime Pfx. Supermix	-	9,5	188	48541	9110	0.188	8826	0.031	152753	84.16	48.86
HVFTper15	HVFTper15Fgo	tooth	MinElute column	0,41	-	G2GO Taq polymerase	-	68,4	179	166992	25686	0.154	24712	0.038	341604	67.62	33.73
HVFTper15	HVFTper15Fpre	tooth	MinElute column	0,368	G2GO Taq polymerase	Kappa Hifi	26	77,4	194	95859	19746	0.206	19104	0.033	304557	86.20	40.08
HVFTper15	HVFTper15go	petrous bone	MinElute column	1	-	G2GO Taq polymerase	-	59,2	199	203877	141168	0.692	134754	0.045	1515185	59.39	23.06
HVFTper15	HVFTper15pre	petrous bone	MinElute column	0,61	G2GO Taq polymerase	Kappa Hifi	24,6	38,4	211	102671	68164	0.664	65260	0.043	795606	84.38	30.59
HVFTper15	HVFTper15SZacc	petrous bone	Silica	0,0963	-	Accuprime Pfx. Supermix	-	3,64	171	154752	105078	0.679	101232	0.037	1333031	37.65	9.73
HVFTper15	HVFTper15SZgo	petrous bone	Silica	0,0963	-	G2GO Taq polymerase	-	15,3	177	202126	142544	0.705	137574	0.035	1895745	39.46	10.24
HVFTper15	HVFTper15SZpre	petrous bone	Silica	0,0963	G2GO Taq polymerase	Kappa Hifi	6,44	60	182	166569	132820	0.797	127182	0.042	1485737	42.67	10.59
HVFTper19	HVFTper19acc	petrous bone	MinElute column	0,496	-	Accuprime Pfx. Supermix	-	26,4	183	169574	134838	0.795	129866	0.037	1728596	59.85	21.34
HVFTper19	HVFTper19Facc	tooth	MinElute column	0,244	-	Accuprime Pfx. Supermix	-	16	181	123991	190	0.002	184	0.032	2177	61.21	33.23
HVFTper19	HVFTper19Fgo	tooth	MinElute column	0,244	-	G2GO Taq polymerase	-	61,8	178	155200	306	0.002	294	0.039	5500	51.73	28.04
HVFTper19	HVFTper19Fpre	tooth	MinElute column	0,458	G2GO Taq polymerase	Kappa Hifi	25,6	53,4	192	102793	170	0.002	168	0.012	NA	50.52	19.74
HVFTper19	HVFTper19go	petrous bone	MinElute column	0,496	-	G2GO Taq polymerase	-	80	197	199022	164540	0.827	158440	0.037	2135199	53.47	19.46
HVFTper19	HVFTper19pre	petrous bone	MinElute column	0,33	G2GO Taq polymerase	Kappa Hifi	28,2	53,2	210	158195	132514	0.838	127580	0.037	1784592	68.87	21.21
HVFTper19	HVFTper19SZacc	petrous bone	Silica	0,229	-	Accuprime Pfx. Supermix	-	14,7	180	199794	168748	0.845	162076	0.040	1957739	41.50	12.61

HVFTper19	HVFTper19SZgo	petrous bone	Silica	0,229	-	G2GO Taq polymerase	-	71	181	170537	143222	0.840	138668	0.032	2123793	41.80	12.50
HVFTper19	HVFTper19SZpre	petrous bone	Silica	0,229	G2GO Taq polymerase	Kappa Hifi	18,8	66,4	185	121788	104024	0.854	100222	0.037	1344114	45.78	12.83
HVFTper2	HVFTper2acc	petrous bone	MinElute column	14	-	Accuprime Pfx. Supermix	-	19,7	195	112821	20776	0.184	20016	0.037	271666	64.53	23.89
HVFTper2	HVFTper2Facc	tooth	MinElute column	0,35	-	Accuprime Pfx. Supermix	-	24	181	160395	25870	0.161	24800	0.041	314589	64.40	30.17
HVFTper2	HVFTper2Fgo	tooth	MinElute column	0,35	-	G2GO Taq polymerase	-	67,6	181	209099	31112	0.149	29948	0.037	404768	58.63	25.67
HVFTper2	HVFTper2Fpre	tooth	MinElute column	0,492	G2GO Taq polymerase	Kappa Hifi	30,6	106	195	67845	5076	0.075	4934	0.028	93861	74.74	29.65
HVFTper2	HVFTper2go	petrous bone	MinElute column	14	-	G2GO Taq polymerase	-	65,2	193	145917	30906	0.212	29668	0.040	377894	55.22	20.78
HVFTper2	HVFTper2pre	petrous bone	MinElute column	16,6	G2GO Taq polymerase	Kappa Hifi	53,6	39,6	229	91924	12600	0.137	12158	0.035	183453	73.20	25.31
HVFTper2	HVFTper2SZacc	petrous bone	Silica	2,97	-	Accuprime Pfx. Supermix	-	36	185	98882	35178	0.356	33728	0.041	397688	47.09	17.87
HVFTper2	HVFTper2SZgo	petrous bone	Silica	2,97	-	G2GO Taq polymerase	-	120	186	143234	55772	0.389	53766	0.036	753918	46.53	17.10
HVFTper2	HVFTper2SZpre	petrous bone	Silica	2,97	G2GO Taq polymerase	Kappa Hifi	26,2	68,8	190	94392	39252	0.416	37568	0.043	437675	51.32	18.24
HVFTper3	HVFTper3acc	petrous bone	MinElute column	0,714	-	Accuprime Pfx. Supermix	-	40,4	205	243458	195146	0.802	187572	0.039	2475181	65.52	27.02
HVFTper3	HVFTper3Facc	tooth	MinElute column	0,112	-	Accuprime Pfx. Supermix	-	25,2	174	163515	125472	0.767	120614	0.039	1639685	58.78	24.02
HVFTper3	HVFTper3Fgo	tooth	MinElute column	0,112	-	G2GO Taq polymerase	-	59	190	208043	159028	0.764	152734	0.040	1886177	55.26	21.59
HVFTper3	HVFTper3Fpre	tooth	MinElute column	0,118	G2GO Taq polymerase	Kappa Hifi	23	72,2	209	106853	59806	0.560	57542	0.038	758323	70.94	27.19
HVFTper3	HVFTper3go	petrous bone	MinElute column	0,714	-	G2GO Taq polymerase	-	75,6	201	177620	146132	0.823	140016	0.042	1713882	60.30	24.47
HVFTper3	HVFTper3pre	petrous bone	MinElute column	1,05	G2GO Taq polymerase	Kappa Hifi	57,8	37,6	191	100544	79238	0.788	76364	0.036	1068353	75.46	25.60
HVFTper3	HVFTper3SZacc	petrous bone	Silica	0,198	-	Accuprime Pfx. Supermix	-	11,9	184	163309	136244	0.834	132224	0.030	2170062	50.12	19.44
HVFTper3	HVFTper3SZgo	petrous bone	Silica	0,198	-	G2GO Taq polymerase	-	59,6	180	193658	161422	0.834	155920	0.034	2221208	46.88	16.99
HVFTper3	HVFTper3SZpre	petrous bone	Silica	0,198	G2GO Taq polymerase	Kappa Hifi	24,4	83,2	191	104140	89264	0.857	85776	0.039	1136744	54.46	18.99
HVFTper4	HVFTper4acc	petrous bone	MinElute column	0,64	-	Accuprime Pfx. Supermix	-	11,9	197	194532	141318	0.726	135676	0.040	1766388	64.07	24.40
HVFTper4	HVFTper4Facc	tooth	MinElute column	0,156	-	Accuprime Pfx. Supermix	-	8,78	190	94768	6830	0.072	6594	0.035	95913	69.87	34.00
HVFTper4	HVFTper4Fgo	tooth	MinElute column	0,156	-	G2GO Taq polymerase	-	31,4	188	184966	14166	0.077	13656	0.036	203802	68.99	33.45
HVFTper4	HVFTper4Fpre	tooth	MinElute column	1,15	G2GO Taq polymerase	Kappa Hifi	34	102	197	100300	1326	0.013	1282	0.033	14125	71.13	28.93
HVFTper4	HVFTper4go	petrous bone	MinElute column	0,64	-	G2GO Taq polymerase	-	62,6	195	196076	147486	0.752	142166	0.036	1977553	57.16	21.33
HVFTper4	HVFTper4pre	petrous bone	MinElute column	0,638	G2GO Taq polymerase	Kappa Hifi	30,4	67	209	99115	75534	0.762	72874	0.035	1103753	73.34	25.72
HVFTper4	HVFTper4SZacc	petrous bone	Silica	0,417	-	Accuprime Pfx. Supermix	-	22,4	181	128253	98034	0.764	94872	0.032	1407126	50.62	21.11
HVFTper4	HVFTper4SZgo	petrous bone	Silica	0,417	-	G2GO Taq polymerase	-	64	183	183370	138808	0.757	133998	0.035	1982991	46.95	17.60
HVFTper4	HVFTper4SZpre	petrous bone	Silica	0,417	G2GO Taq polymerase	Kappa Hifi	24,4	96,2	191	95915	77470	0.808	74942	0.033	1077585	55.31	20.35

9.2. Melléklet: a teljes genomszekvenálás eredményei

Lab ID	Teljes fastq olvasatok száma	Humán genomra illesztett olvasat száma	PCR és optikai duplikátumok száma	Optikai duplikátumok	A könyvtár becsült mérete	Átlagos genom lefedettség	Átlag fragment méret
HVFTper2SZpre	314 573 230	88 033 989	14 980 044	673 556	236 733 722	1,349676689	57,9576
HVFTper3SZpre	108 095 450	87 961 863	14 216 591	763 168	252 707 962	1,426188808	60,6725
HVFTper4SZpre	64 352 889	47 862 805	5 154 599	264 428	215 495 111	0,844056818	62,0068
HVFTper10go	97 492 495	73 611 130	11 010 466	616 547	231 356 295	1,189897992	59,6301
HVFTper19SZpre	114 107 892	85 298 383	12 496 723	595 805	272 477 369	1,167664113	50,3115
HVFTper15SZpre	105 831 365	70 033 973	11 817 540	368 881	188 035 609	0,866125947	46,6531