



UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

Corso di Laurea Magistrale di Biologia Molecolare e Applicata

**VALUTAZIONE DELLE VARIANTI ETEROPLASMICHE NEL GENOMA
MITOCONDRIALE UMANO NEI FUSTI DI CAPELLI MEDIANTE
MASSIVE PERALLEL SEQUENCING: APPLICAZIONI E
PROBLEMATICHE IN GENETICA FORENSE**

**EVALUATION OF HETEROPLASMIC VARIANTS IN THE HUMAN
MITOCHONDRIAL GENOME IN HAIR SHAFTS USING MASSIVE
PARALLEL SEQUENCING: APPLICATION AND PROBLEMS IN
FORENSIC GENETICS**

Tesi di Laurea Magistrale di:

Miriana Leggieri

Relatore:

Chiar.ma Prof.ssa

Chiara Turchi

Sessione estiva

Anno Accademico 2021/2022

INDICE

Capitolo 1	5
INTRODUZIONE	5
1.1 <i>Genetica forense</i>	5
1.2 <i>Tipizzazione del DNA e marcatori genetici</i>	5
1.3 <i>Morfologia e funzione mitocondriale</i>	9
1.4 <i>Il genoma mitocondriale</i>	11
1.5 <i>Variabilità mitocondriale</i>	12
1.5 <i>Eteroplasmie mitocondriali</i>	14
1.6 <i>Eteroplasmie mitocondriali in vari tessuti</i>	16
1.6.2 <i>Eteroplasmie mitocondriali nei campioni di capelli</i>	17
1.7 <i>Sequenze mitocondriali nucleari umane (NUMT)</i>	19
1.8 <i>DNA mitocondriale nell'indagine forense</i>	20
1.9 <i>Database del DNA mitocondriale</i>	21
SCOPO DELLO STUDIO	23
Capitolo 2	25
MATERIALI E METODI	25
2.1 <i>Progettazione dello studio</i>	25
2.2 <i>Raccolta dei campioni</i>	26
2.3 <i>Estrazione</i>	29

2.3.1 Estrazione dai tamponi buccali	30
2.3.2 Estrazione da sangue	32
2.3.3 Estrazione da capelli.....	33
2.4 Quantificazione del mtDNA nei frammenti di capelli.....	37
2.5 Quantificazione dei campioni di riferimento	39
2.6 Sequenziamento con HID Ion GeneStudio S5 System.....	41
2.6.1 Reazione di amplificazione del mtDNA target	42
2.6.2 Digestione degli ampliconi.....	45
2.6.3 Legame degli adattatori e purificazione della libreria.....	45
2.6.4 Quantificazione della libreria mediante qPCR	48
2.6.5 Diluizione, raggruppamento e archiviazione delle librerie	49
2.6.6 Preparazione del template su HID Ion Chef	50
2.6.7 Sequenziamento con HID Ion GeneStudio S5 TM System.....	53
2.7 Analisi dei dati	54
Capitolo 3	55
RISULTATI	55
3.1 Risultati della quantificazione del mtDNA.....	55
3.2 Risultati della quantificazione delle librerie.....	64
3.3 Risultati analisi delle prestazioni del sequenziamento	67

<i>3.4 Analisi del coverage e performance del pannello.....</i>	<i>71</i>
<i>3.5 Analisi delle varianti eteroplasmatiche rilevate tramite MPS.....</i>	<i>75</i>
Capitolo 4	84
DISCUSSIONE	84
Capitolo 5	93
CONCLUSIONE.....	93
RINGRAZIAMENTI	96
BIBLIOGRAFIA.....	98
SITOGRAFIA IMMAGINI	101

Capitolo 1

INTRODUZIONE

1.1 Genetica forense

La *genetica forense* è una moderna branca della genetica che ha il compito di assegnare a un campione biologico una corrispondente identità. Questo comporta la sua necessità per risolvere casi giudiziari o contenziosi in ambito civile e penale. Il materiale biologico di partenza può essere di vario tipo: capelli, sangue, tamponi buccali, liquido seminale, materiale autoptico etc. Da questo materiale biologico si procede estraendo il DNA per ottenere così ciò che ci consentirà di delineare il profilo di un individuo. Infatti, la genetica forense è utilizzata, soprattutto, per: individuare l'autore di un reato, mettendo in correlazione i dati genetici di un sospettato con il materiale biologico rinvenuto sulla scena del crimine; compiere le indagini di parentela, volte ad accertare un rapporto di paternità contestato o dubbio; assegnare un'identità a dei resti umani rinvenuti a seguito di una tragedia o una catastrofe. Per fare questo, la genetica forense comprende saperi e tecniche di varie discipline quali la genetica, la biologia molecolare, la medicina legale etc.

1.2 Tipizzazione del DNA e marcatori genetici

La maggior parte del DNA (99,7%) è identico tra un individuo e un altro, solo una piccola frazione del nostro genoma (0,3%) è variabile. Grazie a questa "piccola" variabilità caratteristica di ogni individuo è possibile andare a tipizzare, ossia utilizzare l'informazione contenuta in queste regioni variabili del DNA per

identificare e distinguere un individuo da un altro. Per poter fare questo nella genetica forense vengono presi in considerazione dei marcatori del DNA, ossia un locus genetico che mostra un'elevata variabilità. Inoltre, considerando che spesso i campioni forensi contengono DNA degradato, ossia frammentato in piccoli pezzi, a causa di agenti chimici e/o fisici, i marcatori oltre ad avere un elevato grado di variabilità all'interno della popolazione, dovranno anche avere una lunghezza in nucleotidi ridotta, in genere al di sotto delle 400 bp. Tra i marcatori genetici più conosciuti e studiati in ambito forense ci sono i microsatelliti o *short tandem repeats* (STRs). Essi sono infatti, i migliori candidati per l'analisi forense per la loro ridotta lunghezza della sequenza di DNA che costituisce l'unità ripetuta in tandem, chiamata anche unità ripetitiva, che può variare da 2 a 6 paia di basi (Fig. 1).

L'insieme delle unità ripetitive costituisce la cosiddetta "*regione ripetuta*" del microsatellite, essa è costituita da un numero variabile di ripetizioni nei vari individui, il che li rende utili nell'identificazione umana. Gli STRs vengono inoltre definiti dinucleotidici, trinucleotidici, tetranucleotidici, pentanucleotidici ed esanucleotidici, cioè costituiti rispettivamente da 2, 3, 4, 5 e 6 nucleotidi per unità ripetitiva.

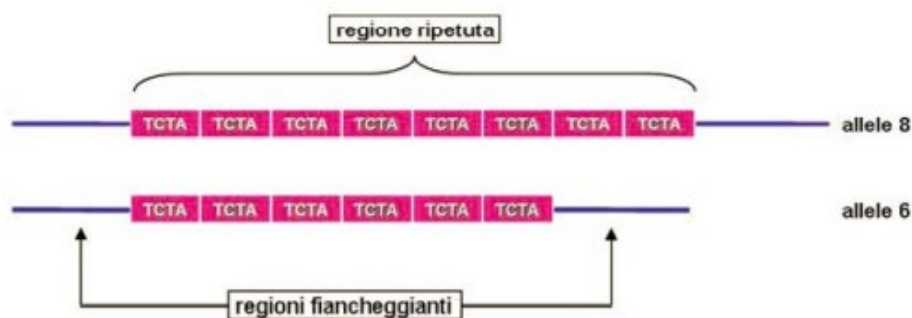


Fig. 1 Struttura di un STR.

Gli STRs costituiti da ripetizioni tetranucleotidiche sono più utilizzati in ambito forense poiché nel corso dell'amplificazione PCR vi è una minore probabilità di formazione delle *stutter*, ossia degli ampliconi (regioni di DNA prodotte in PCR) più corti di una o più unità ripetitive rispetto all'allele. Questi ultimi vengono generati durante il processo di duplicazione del DNA in seguito a uno scivolamento della polimerasi sul filamento di DNA stampo. La loro presenza rende più difficile l'interpretazione di tracce miste di DNA (tracce in cui sono presenti più profili genetici appartenenti ad individui diversi). Inoltre, gli STRs tetranucleotidici sono più facilmente distinguibili durante la separazione elettroforetica, basata sulla diversa lunghezza dei frammenti di DNA. È possibile estrarre e amplificare più STRs simultaneamente, utilizzando un mix di primer diretti verso diversi STRs, tramite la cosiddetta *multiplex PCR*. Nel 1996, l'FBI Laboratory ha fornito 13 loci da poter utilizzare nell'allestimento del database nazionale del DNA, meglio conosciuto come Combined DNA Index System (CODIS). Dal 2017 è stato approvato l'ampliamento di questo sistema aggiungendo 7 loci, fino a un totale di 20 STR, per aumentare il potere discriminatorio.

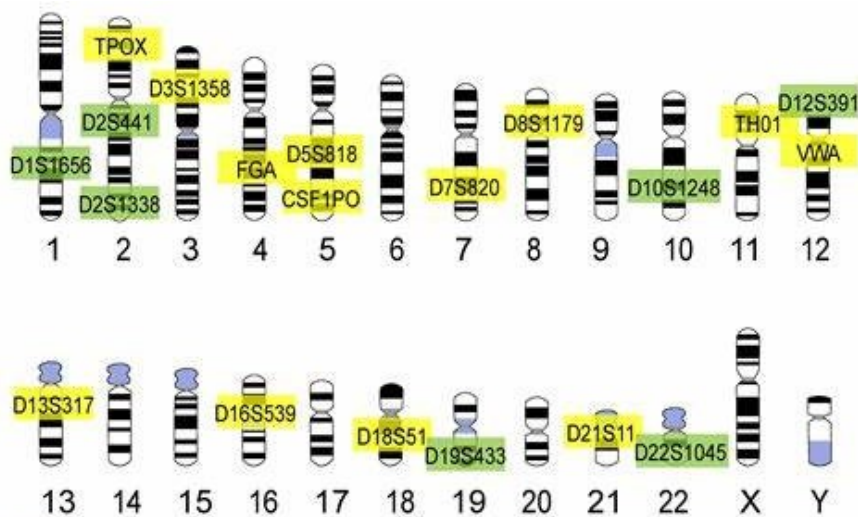


FIG.2 20 loci STR del CODIS con posizioni cromosomiche (in verde i loci aggiunti dal 2017)

Andando a tipizzare i 13 microsatelliti del CODIS è possibile identificare in maniera inequivocabile un individuo poiché la *random match probability* (la probabilità che due individui non imparentati abbiano lo stesso profilo genetico) è pari a uno su mille miliardi di individui non imparentati. È possibile andare a tipizzare contemporaneamente questi microsatelliti in poco tempo e con poco DNA di partenza utilizzando vari kit presenti in commercio. La genetica forense però, non si occupa soltanto di analizzare microsatelliti a livello di cromosomi autosomici, ma anche a livello di cromosomi sessuali come il cromosoma Y, e a livello del DNA mitocondriale. Essi vengono utilizzati per la ricostruzione delle linee parentali, in particolare l'analisi di polimorfismi sul cromosoma Y serve per accertamenti di paternità, al contrario i polimorfismi del DNA mitocondriale possono essere utilizzati per confermare l'ereditarietà materna, in quanto ereditato solo per via materna¹.

1.3 Morfologia e funzione mitocondriale

I *mitocondri* sono organuli cellulari, essi hanno lunghezza compresa tra 1 e 6 μm e spessore tra 0,2 e 1 μm , anche il numero è variabile per cellula a seconda del suo fabbisogno energetico. Essi si presentano come cilindretti allungati con bordi arrotondati, sono costituiti da due membrane, una *membrana esterna* e una *interna*, le quali delimitano lo *spazio intermembrana*. La membrana esterna ha un più alto contenuto lipidico e sono costituiti essenzialmente da fosfolipidi, inoltre, è caratterizzata da un'elevata permeabilità che consente il passaggio di diverse molecole. La membrana interna si ripiega verso l'interno formando delle *creste mitocondriali*, essa delimita la camera interna definita *matrice mitocondriale*, costituita da un gel discretamente viscoso nel quale si trovano enzimi per l'ossidazione degli acidi grassi e per il funzionamento del ciclo di Krebs, e i *granuli* contenenti calcio e magnesio. Inoltre, la membrana interna non è permeabile quanto quella esterna e contiene diverse proteine di trasporto che controllano il passaggio selettivo di alcune molecole coinvolte nei processi metabolici che avvengono nel mitocondrio (Fig. 3).

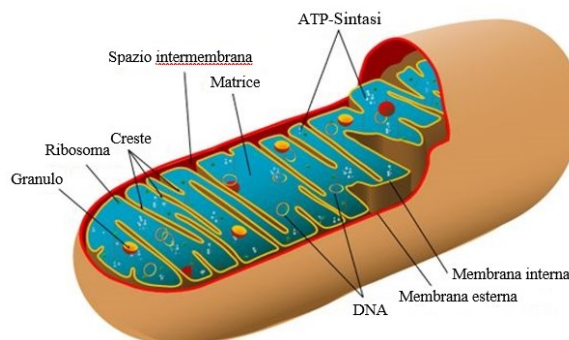


Fig. 3 Struttura di un mitocondrio

I mitocondri, infatti, sono coinvolti nel metabolismo energetico, ossia quell'insieme di reazioni che a partire dalla degradazione di una molecola organica porta alla produzione di molecole di ATP. Il metabolismo energetico consiste di una *fase anaerobia*, che si attua in assenza di ossigeno e si svolge nel citoplasma cellulare, e di una *fase aerobia* che richiede la presenza di ossigeno ed è questa che viene svolta all'interno dei mitocondri. Durante la prima fase viene svolta la *glicolisi*, nella quale, com'è intuibile dal nome, le molecole di *glucosio* ricche di energia sono scisse ognuna in due molecole di *piruvato*, vengono prodotte due molecole di ATP al netto e due di NADH (per riduzione del NAD^+). Il piruvato prodotto in presenza di ossigeno si trasforma in *radicale acetato* che, a propria volta legandosi al *coenzima A* forma l'*acetilcoenzima A*, il quale raggiungerà la matrice mitocondriale e innescherà l'inizio del *ciclo di Krebs*. In questo ciclo attraverso otto reazioni consecutive verranno prodotte due molecole di CO_2 . Dal punto di vista energetico importante è la produzione di una molecola di GTP (utilizzata per rigenerare una molecola di ATP), di tre molecole di NADH e una di FADH_2 . Infine, il NADH e FADH_2 si comportano come intermedi ossidoriduttivi in grado di trasportare elettroni fino alla *catena respiratoria mitocondriale*, dove vengono riossidati (a NAD^+ e FAD) e cedono gli elettroni alla catena stessa, che sarà così in grado di rigenerare molecole di ATP a partire da ADP^2 .

1.4 Il genoma mitocondriale

I mitocondri sono dotati di un proprio genoma definito appunto *genoma mitocondriale*, ereditato esclusivamente per via materna. Quest'ultimo non è legato a proteine e consiste di piccole molecole di DNA circolare adese alle creste mitocondriali, è presente in numerose copie, e contiene soltanto 37 geni, di cui 24 codificano per tRNA e rRNA, mentre gli altri 13 codificano per proteine. Il mtDNA è costituito da una catena a doppio filamento di circa 16,5 kb. Si possono distinguere due filamenti noti come: *filamento pesante* (H) e *filamento leggero* (L), il primo è ricco in G e codifica per 28 geni, il secondo è ricco in C e codifica per 9 geni. La maggior parte del DNA mitocondriale è codificante (ca. 93%), a differenza del DNA nucleare, in cui solo una piccola parte (ca. 2%) risulta codificante. L'unica regione priva di DNA codificante è quella denominata "*regione di controllo*", localizzata tra i geni per il tRNA della prolina e per il tRNA della fenilalanina. Questa regione, lunga 1.122 bp è compresa tra la posizione 16.024 e la posizione 576 e rappresenta il 5-7% del DNA mitocondriale. Per la presenza dell'OH la regione di controllo è chiamata anche "*regione contenente il D-loop*", in quanto la replicazione del DNA mitocondriale avviene secondo il modello dello spostamento dell'ansa (displacement loop o D-loop)¹.

1.5 Variabilità mitocondriale

L'apparente mancanza di meccanismi di riparazione del mtDNA e la bassa fedeltà della mtDNA polimerasi portano ad un tasso di mutazione significativamente più elevato nel genoma mitocondriale, rispetto al genoma nucleare. Alla *regione di controllo* è attribuibile il più alto tasso di variabilità rispetto al resto del genoma mitocondriale. In particolare, all'interno della regione di controllo abbiamo: la *regione ipervariabile 1* (HVR1) e la *regione ipervariabile 2* (HVR2), caratterizzate da un elevato polimorfismo. La HVR1 è compresa tra la posizione 16.024 e la 16.365, mentre la HVR2 tra la posizione 73 e la 340; oltre a queste sono presenti anche altre regioni polimorfiche come la HVR3, compresa tra la posizione 438 e 574 (Fig.4).

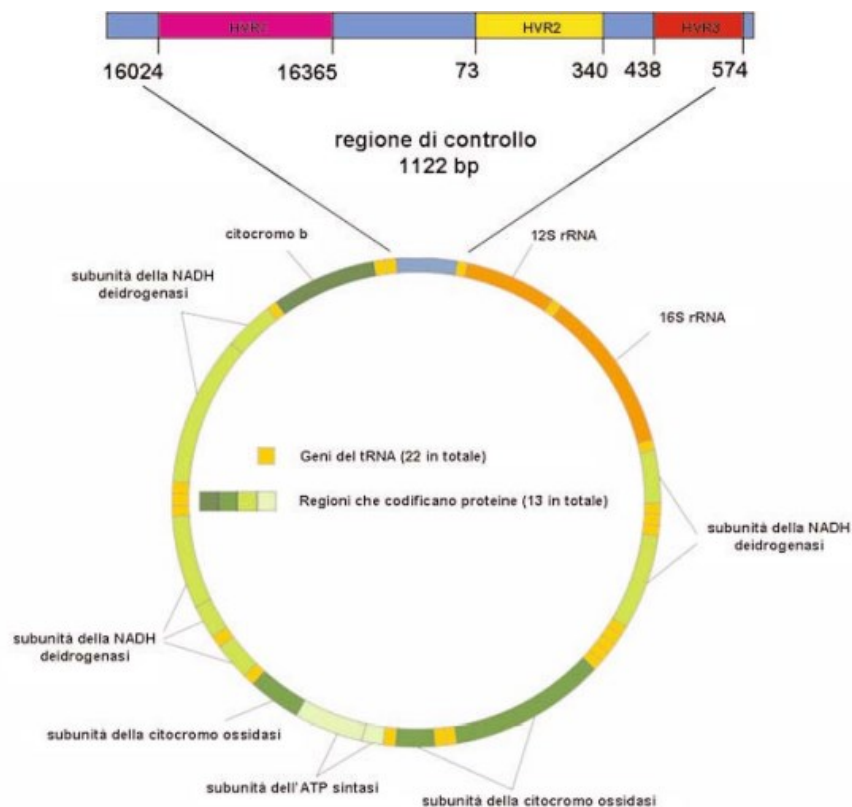


Fig. 4 Schema del genoma mitocondriale umano

L'analisi della sequenza del mtDNA definisce *l'aplotipo individuale* che viene riportato in base alle differenze rispetto all'rCRS (sequenza di riferimento di Cambridge). L'accumulo di mutazioni come i polimorfismi a singolo nucleotide (SNP), consentono di definire vari *aplogruppi*, ossia l'insieme di aplotipi simili derivati da un antenato comune. Grazie alla ricerca nel campo della genetica di popolazione questi diversi tipi mitocondriali sono stati definiti attraverso un albero filogenetico che permette di risalire ad un antenato comune.

I principali aplogruppi mitocondriali sono nominati in ordine di lettere, dalla A alla Z in base al momento della scoperta, in particolare: la grande maggioranza degli individui provenienti da popolazioni africane, e specialmente da popolazioni dell'Africa sub-sahariana, hanno avuto origine dal macro-aplogruppo L (che comprende: L0, L1, L2, L3, L4, L5 e L6); oltre il 90% degli individui delle popolazioni caucasiche europee e statunitensi sono classificati in 10 aplogruppi: H, I, J, K, M, T, U, V, W e X; per quanto riguarda le popolazioni afro-americane, gli aplogruppi più comunemente osservati sono L2a, L1c, L1b e L3b; mentre i principali aplogruppi trovati in individui provenienti da popolazioni asiatiche sono gli aplogruppi M e N³.

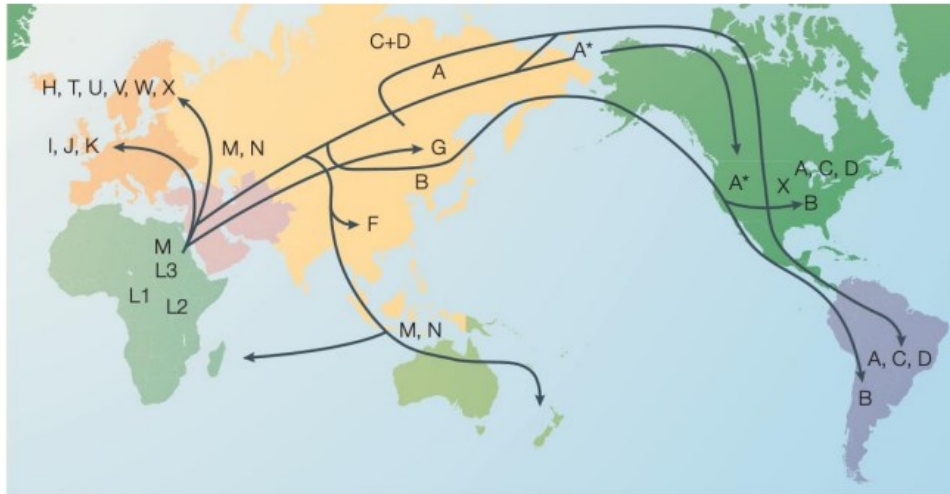


Fig.4 Distribuzione geografica dei principali aplogruppi mitocondriali

1.5 Eteroplasmie mitocondriali

Per *eteroplasmia* si intende la coesistenza di più di una variante del genoma mitocondriale (mtDNA) all'interno di un individuo. Esistono due classi di eteroplasmie: quelle legate a polimorfismi di lunghezza e quelle caratterizzate da sostituzioni puntiformi. Solo quest'ultimo tipo è importante per l'identificazione umana forense poiché l'informazione dei polimorfismi di lunghezza non ha alcun impatto nella definizione degli aplogruppi. Nelle eteroplasmie causate da mutazioni puntiformi una singola base viene sostituita, mostrando due picchi coincidenti che rappresentano le basi A (in verde) e G (in nero).

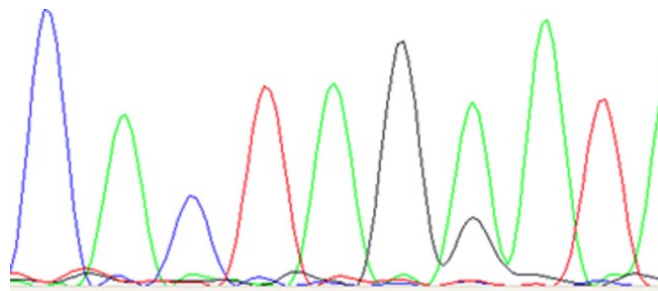


Fig. 5 Elettroferogramma del sequenziamento Sanger di un campione che mostra eteroplasmia puntiforme nella 7^a base da sinistra e mostra due picchi coincidenti che rappresentano le basi A (in verde) e G (in nero).

I meccanismi attraverso i quali una mutazione che colpisce una singola molecola di mtDNA si stabilisce in una popolazione di mitocondri o di cellule sono stati discussi da Stewart e colleghi, essi sono:

- La *segregazione vegetativa*: durante la divisione cellulare il mtDNA passa alle cellule figlie, se è presente un variante eteroplasmica, ogni cellula figlia riceverà una porzione della molecola mutata;
- La *replicazione rilassata*: il mtDNA viene continuamente distrutto e replicato senza uno specifico meccanismo cellulare che assicuri che venga replicato in un esatto numero di copie, spostando così il livello di eteroplasmia all'interno della cellula³.

Quando i livelli di eteroplasmia superano una certa soglia critica possono indurre delle compromissioni metaboliche e causare delle malattie. Questi mtDNA mutanti possono essere ereditati dalla madre o possono insorgere *de novo*. Sono stati fatti vari studi che dimostrano che i livelli di mtDNA mutante possono variare in modo non casuale tra le generazioni e tra i tessuti di un individuo. Tuttavia, la nostra mancanza di comprensione dei meccanismi cellulari e molecolari alla base delle dinamiche dell'eteroplasmia del mtDNA ha reso difficile prevedere chi erediterà o svilupperà malattie associate al mtDNA¹³.

Un decennio fa si pensava che la maggior parte degli individui fosse *omoplasmica* (caratterizzato da un solo tipo di mtDNA), tuttavia, con l'avvento delle nuove tecnologie i ricercatori hanno osservato molti casi di eteroplasmia nell'uomo. Ad esempio, è stata osservata eteroplasmia nella posizione 16169 della regione di

controllo del mtDNA nei presunti resti dello zar Nicola II di Russia e nei resti del fratello dello zar, il granduca di Russia Georgij Romanov³.

1.6 Eteroplasmie mitocondriali in vari tessuti

I livelli di eteroplasmia variano nei diversi tessuti e dunque, è necessario per definire correttamente le mutazioni in ciascun campione di tessuto analizzato, raccogliere più campioni di uno stesso individuo, gli studi si sono concentrati sull'analisi del sangue, delle cellule buccali e capelli³. Una delle prime scoperte in quest'area di ricerca è che i livelli del tipo di mutazione sono più elevati nei tessuti ad alto fabbisogno energetico, ad esempio nelle cellule del cervello, muscolo scheletrico, fegato e cuore. L'elevata richiesta energetica di questi tessuti può portare a una replicazione incontrollata dei mitocondri maggiormente soggetta ad errori, il che porta a una maggiore probabilità di mutagenesi⁴. Diversi studi si sono concentrati sulla variazione dei livelli di eteroplasmia in diversi tessuti in ambito forense, essi sono incentrati per lo più su eventi mutazionali trovati nella regione di controllo del mtDNA. Il primo studio sull'eteroplasmia della regione di controllo umana è stato condotto da Wilson e colleghi nel 1997 su una famiglia composta da madre, una figlia e un figlio. I campioni di DNA sono stati estratti da macchie di sangue, tamponi buccali e capelli di questi individui, essi sono stati amplificati tramite PCR e sequenziati. Sia nella figlia che nella madre, l'eteroplasmia è stata osservata nel DNA estratto da macchie di sangue, tamponi buccali e capelli, ma i livelli di nucleotidi variavano a seconda del tessuto analizzato. Nel DNA estratto da sangue, la

proporzione di citosina era maggiore della timina, gli estratti di tampone buccale hanno mostrato un contributo più equilibrato dei due nucleotidi, mentre nei campioni di capelli, alcuni conservavano lo stesso contributo di ciascun nucleotide, e alcuni sembravano avere una distribuzione opposta di T e C. Da qui la conclusione che esaminando diversi tessuti dal punto di vista forense se i campioni di riferimento (sangue e tampone buccale) mostrano omoplasmia del mtDNA e i campioni di capelli mostrano, invece, eteroplasmia nello stesso sito nucleotidico, non è possibile escludere la possibilità che le due tracce, apparentemente diverse, possano appartenere allo stesso individuo⁵. Uno studio più recente ha condotto un'indagine approfondita sull'eteroplasmia nell'uomo della regione di controllo mitocondriale, analizzando nove diversi campioni di tessuto provenienti da cento individui. In questo studio hanno valutato i livelli di eteroplasmia mediante il sequenziamento di Sanger e solo il 12% di tutti gli individui non ha mostrato eteroplasmia, mentre l'88% ha mostrato almeno una posizione eteroplasmica all'interno dei tessuti studiati. Inoltre, hanno valutato le varie differenze tessuto-specifiche, dimostrando che: il più alto numero relativo di eteroplasmie è a livello dei muscoli e del fegato (79%, 69%), seguito da cervello, capelli e cuore (36,7%-30,2%). Percentuali più basse sono state osservate nelle cellule ossee, del sangue, polmonari e buccali (19,8%-16,2%)⁶.

1.6.2 Eteroplasmie mitocondriali nei campioni di capelli

L'eteroplasmia è più frequentemente osservata nei campioni di capelli, principalmente a causa della deriva genetica e dei colli di bottiglia che si verificano a causa della natura semiclonale del follicolo pilifero⁷. Per capire meglio il meccanismo

alla base della distribuzione di tale variabilità, la conoscenza dell'istogenesi dei capelli è essenziale. I capelli hanno principalmente origini ectodermiche, con un certo contributo dello strato germinale mesodermico. Alla nona-dodicesima settimana di gestazione, un gruppo di cellule dell'ectoderma inizia a crescere verso il basso e incontra un gruppo di cellule mesodermiche che in seguito formeranno la papilla dermica e la guaina del follicolo. Entro la sedicesima-ventesima settimana della gestazione, i follicoli piliferi iniziano a produrre peli, ciascuno di questi avrà cellule staminali che alimentano le cellule nei fusti dei capelli in crescita, con conseguente limitazione del pool di molecole di mtDNA che forniscono la prova del primo collo di bottiglia somatico che il capello subisce durante il suo sviluppo. Un'ulteriore variazione genetica è fornita dalla rapida divisione cellulare che porta a un'ulteriore segregazione delle molecole del mtDNA mentre i capelli continuano a crescere. Dopo l'embriogenesi, si stima che una persona media abbia circa 5 milioni di follicoli piliferi del corpo. Sebbene il numero dei follicoli rimanga costante, le dimensioni dei singoli follicoli e di tutti i peli risultanti possono cambiare drasticamente nel tempo⁸.

Nel bulbo della radice del capello si trovano i melanociti, essi producono quantità e tipi variabili di melanosomi (granuli di pigmento) che alla fine determinano il colore dei capelli, poiché trasportati dalle cellule corticali dalla radice al fusto del capello.

Le cellule corticali potranno avere due popolazioni mitocondriali, una dalle cellule germinali del bulbo e una dai melanociti, con conseguente presenza di diversi pool di mtDNA. Dal momento che le cellule della corticale hanno vita breve, nuove cellule vengono continuamente prodotte, e queste possono portare nuove molecole di

mtDNA, il che può comportare una variazione dei livelli di eteroplasmia lungo il capello ma anche tra capelli diversi⁹.

1.7 Sequenze mitocondriali nucleari umane (NUMT)

Diversi studi negli ultimi anni si sono incentrati sullo studio dei *NUMT*, ossia segmenti di mtDNA inseriti nel genoma nucleare. Il termine NUMT è stato coniato da Lopez et al che ha riportato la trasposizione di un frammento di mtDNA a 7,9 kbp in Ndna del gatto domestico. Il meccanismo con cui avviene tale trasferimento non è ancora ben compreso ma sono stati ipotizzati due possibili meccanismi: il meccanismo di giunzione dell'estremità non omologa (NHEJ) nella riparazione della rottura del doppio filamento e la ricombinazione omologa (che richiede invece, omologia di sequenza). Si ritiene che l'inserimento di queste sequenze sia un processo evolutivo ancora in corso e che la maggior parte dei NUMT sono stati inseriti negli ultimi 100.000 anni. Oggi sono stati identificati più di 1077 NUMT umani, anche grazie all'emergenza delle tecnologie basate su MPS. Queste piattaforme hanno migliorato le nostre conoscenze della posizione di NUMT e ci ha aiutato a scoprire che tali frammenti possono variare da 39 bp a intere sequenze mitogenomiche. Questi sono noti come mega-NUMTS, e hanno riscosso interesse dalla genetica forense, in quanto tali inserti possono portare a delle interpretazioni errate come la possibilità di trasmissione mitocondriale paterna¹⁰. Recentemente, sono state rivendicate le prove di ereditarietà biparentale del mtDNA, infatti, uno studio ha indagato sulla presenza di mega-NUMT in 8 familiari sani, ed è emerso che

questi inserti mitogenomici completi possono causare la presenza di diversi mitotipi. Le cellule senza DNA nucleare, cioè trombociti e fusti del capello, hanno mostrato solo il mitotipo dell'aplogruppo V, mentre in cellule con nDNA presentavano il mitotipo hg U4c1. Poiché i NUMT possono complicare l'analisi della sequenza del mtDNA a causa della presenza di miscele che potrebbero essere erroneamente interpretate come prove di ereditarietà paterna gli esperti suggeriscono l'analisi dei fusti dei capelli e dei trombociti oltre a quella delle cellule del sangue e della bocca, dove tale fenomeno è più comune.

1.8 DNA mitocondriale nell'indagine forense

L'importanza dell'analisi del mtDNA in ambito forense è dovuta in primo luogo al fatto che in caso di campioni altamente degradati e quindi con scarsa presenza di DNA nucleare, esso è un valido sostituto in quanto presente in numerose copie rispetto all'nDNA, tuttavia, il potere discriminatorio del mtDNA è inferiore. Il limite maggiore del mtDNA è che ha un'eredità matrilineare e non è sottoposto a ricombinazione genetica, pertanto, in assenza di eventi mutazionali ogni individuo imparentato per via materna avrà la stessa sequenza (aplotipo) di mtDNA. Dunque, non sarà possibile mediante mtDNA discriminare individui che condividono la stessa discendenza materna.

Un altro limite dell'utilizzo del mtDNA in ambito forense, come già accennato, è la presenza di eteroplasmie mitocondriali, poiché potremmo avere dei risultati fuorvianti la reale appartenenza di un campione biologico. Data la presenza di

eteroplasmie nei diversi tessuti di un unico individuo non sarà possibile escludere l'appartenenza di due diverse sequenze allo stesso soggetto. Da qui l'importanza dello studio delle eteroplasmie del mtDNA per evitare eventuali errori di interpretazione in ambito forense.

1.9 Database del DNA mitocondriale

La significatività di corrispondenza di due sequenze di mtDNA può essere valutata contando quante volte la sequenza del mtDNA, nota come mitotipo, è presente in una popolazione rilevante. Per fare questo vengono utilizzati dei database, il più importante contenente aplotipi di mtDNA è l'EDNAP Mitochondrial DNA Population Database (EMPOP). EMPOP oltre ad essere un database di sequenze di riferimento di popolazione, è anche utilizzato come strumento di controllo della qualità dei dati di mtDNA. Questo database utilizza un algoritmo, il SAM, che permette di individuare aplotipi mitocondriali indipendentemente dall'allineamento tra la query e la sequenza di riferimento. La versione più aggiornata è il SAM-E ed è quello attualmente utilizzato. Questo database include tutti gli aplogruppi Phylotree in un comodo formato di ricerca e fornisce il numero di sequenze EMPOP assegnate ai rispettivi aplogruppi stimando gli aplogruppi del DNA mitocondriale utilizzando l'approccio della massima verosimiglianza EMMA. PhyloTree fornisce una filogenesi completa aggiornata della variazione globale del mtDNA umano, include sia aplogruppi precedentemente pubblicati ma anche quelli identificati di recente. Un

altro importante database di mtDNA umano è *Mitomap* contenente variazioni del mtDNA umano insieme a varianti geografiche e specifiche di alcune malattie³.

L’FBI ha anche creato un database di popolazione mtDNA, il CODISmt, che raccoglie dati sulle regioni ipervariabili da 14 popolazioni¹². In Italia il Ge.FI (Genetisti Forensi Italiani), ha analizzato 593 campioni per creare un database di alta qualità del mtDNA della popolazione italiana¹.

SCOPO DELLO STUDIO

Con la crescente importanza dello studio del mtDNA in ambito forense cresce anche la necessità di sviluppare nuove e più efficienti procedure di analisi e interpretazione dei dati. Nel caso di campioni di riferimento, come ad esempio sangue e tamponi buccali, definire un profilo di mtDNA è relativamente semplice con le tecniche di sequenziamento a disposizione. Al contrario, campioni come i capelli mostrano maggiori difficoltà di interpretazione a causa dell'elevata incidenza di varianti eteroplasmiche sia lungo il capello che tra capelli di uno stesso individuo. È noto che i fusti dei capelli mostrano un collo di bottiglia nella segregazione delle molecole del mtDNA durante lo sviluppo dei capelli stessi. Per questo motivo un gruppo di esperti nel campo della genetica forense ha messo a punto il MitoMetrics Hair Heteroplasmy Study, un progetto collaborativo che pone la sua attenzione verso la rilevazione e interpretazione di eteroplasmie nei campioni di capelli e nei campioni di riferimento (sangue e tampone buccale) attraverso l'utilizzo della tecnologia Massive Parallel Sequencing. Dunque, il progetto MitoMetrics ha come obiettivo quello di raggiungere una gamma più ampia di dati sui livelli di varianti eteroplasmiche di mtDNA nei capelli. Questi dati saranno alla base di una valutazione dei tassi di mutazione somatica nei capelli umani, al fine di migliorare le prestazioni nel campo della genetica forense.

Il mio progetto di tesi si inserisce in questo studio concentrandosi sulla raccolta, selezione, quantificazione e sequenziamento dei campioni per l'analisi dell'eteroplasmia del mtDNA. In particolare, abbiamo analizzato l'eteroplasmia del

mtDNA nei campioni di capelli, e sono stati confrontati con i valori di eteroplasmia riportati per campioni di sangue e tamponi buccali utilizzati come riferimenti. Quindi, abbiamo proceduto con tecniche di estrazione automatizzate, analisi di quantificazione in tempo reale e applicazione del Massive Parallel Sequencing.

Capitolo 2

MATERIALI E METODI

2.1 Progettazione dello studio

Il progetto MitoMetrics ha richiesto un minimo di 5 partecipanti, da ogni individuo ho raccolto 1 tampone buccale, 1 campione di sangue e 3 fusti di capelli. Prima di aver prelevato i campioni, i partecipanti hanno firmato un modulo di consenso informato conforme ai requisiti della legge italiana (GDPR2016/679) per il trattamento dei dati personali. I campioni sono stati etichettati secondo il Mitometrics Hair Heteroplasmy Study Protocol. Il nome del campione è seguito da una lettera che indica il tipo di campione: `_S` per tamponi buccali (swab), `_B` per il sangue (blood), e `_H` (hair) per i fusti dei capelli. Solo per questi ultimi si forniscono ulteriori informazioni indicando il numero del fusto del capello e del frammento del capello (Fig. 6)¹⁵.

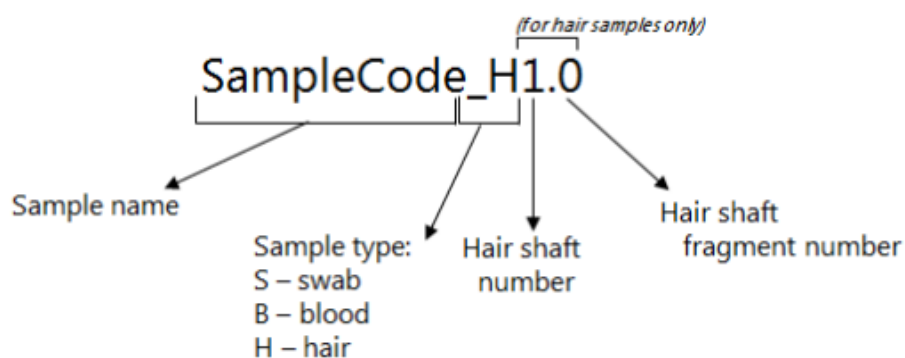


Fig. 6 Codifica dei campioni utilizzati

2.2 Raccolta dei campioni

La raccolta del tampone buccale è stata effettuata utilizzando dei tamponi sterili preconfezionati (CITOSWAB ® Collection swabs, LABOINDUSTRIA):

- Preparare cinque provette Falcon da 15 ml etichettandole con il codice appropriato e irradiarle con luce UV per 5 minuti.
- Tamponare l'interno della guancia di ogni volontario evitando il contatto con superfici contaminate come lingua e denti.
- Mettere il tampone umido nella provetta Falcon da 15 ml e conservare il campione a -20°C, per evitare la formazione di muffe che potrebbero potenzialmente degradare il materiale raccolto.

La raccolta del campione di sangue è stata effettuata da personale specializzato adoperando tutte le opportune misure di sicurezza. Anch'esso è stato conservato a -20°C in provette appositamente codificate.

Per quanto riguarda la raccolta dei campioni di capelli ho seguito la procedura qui riportata:

- Preparare cinque piastre Petri etichettandole con il codice appropriato e irradiarle con luce UV per 5 minuti.
- Chiedere a ciascuno dei partecipanti di pettinarsi i capelli con le dita per raccogliere i capelli.

- Indossare guanti raccogliere i capelli e metterli nella capsula di Petri. Queste possono essere conservate nella stanza a temperatura ambiente prima dell'elaborazione.
- Aprire con cura la capsula di Petri e contare i capelli raccolti.
- Selezionare 3 capelli in cui è visibile il bulbo pilifero, per consentire la distinzione tra parte apicale e distale dello stelo.
- Ogni capello viene misurato, dalla radice alla punta, nella capsula di Petri utilizzando un metro e un paio di pinze, la lunghezza e tutte le caratteristiche peculiari sono state registrate, come mostrato nella tabella 1
- Preparare delle provette da 1,5 ml necessari per raccogliere i frammenti di capelli ed etichettarli seguendo il protocollo: SampleCode_H (numero del fusto del capello). (numero del frammento)
- Le provette vengono poste nella cappa per l'irradiazione UV per 5 minuti per consentire la rimozione di potenziali contaminanti.
- I capelli vengono posti uno ad uno in una capsula di Petri pulita per procedere con il taglio con una lama chirurgica (Benefit Medical Devices Srl.)

Ogni fusto di capelli deve essere diviso in frammenti di 2 cm lungo la sua lunghezza.

La sezione della radice (anche se non visibile) va asportata di ca. 0,5 cm

dall'estremità prossimale dei capelli e trasferita in una provetta (etichettata

SampleCode_HX.0). Dovrebbe essere utilizzato un minimo di 3 frammenti per fusto

del capello, inclusa la prima sezione di 2 cm dopo la radice (SampleCode_HX.1), la

sezione centrale (SampleCode_HX.2) e la punta (SampleCode_HX.3). I frammenti

devono essere numerati dalla radice alla punta partendo da 0. Nel caso di capelli corti (<6 cm) si possono utilizzare i frammenti successivi (Fig.7), mentre per quelli più lunghi (>6cm) i frammenti saranno selezionati dopo aver esaminato l'intera lunghezza del capello (Fig. 8)¹⁵.

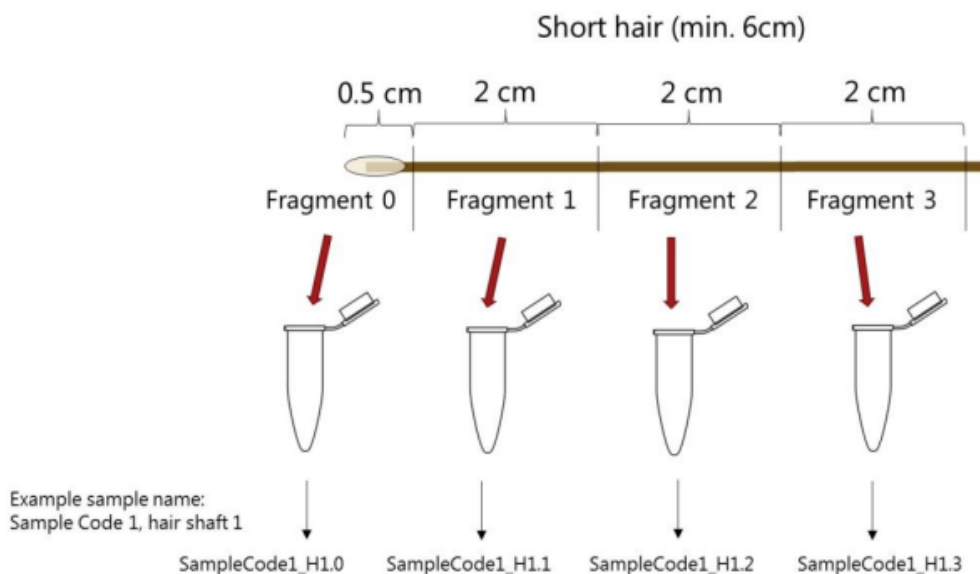


Fig. 7 Preparazione e identificazione del campione di capelli corti

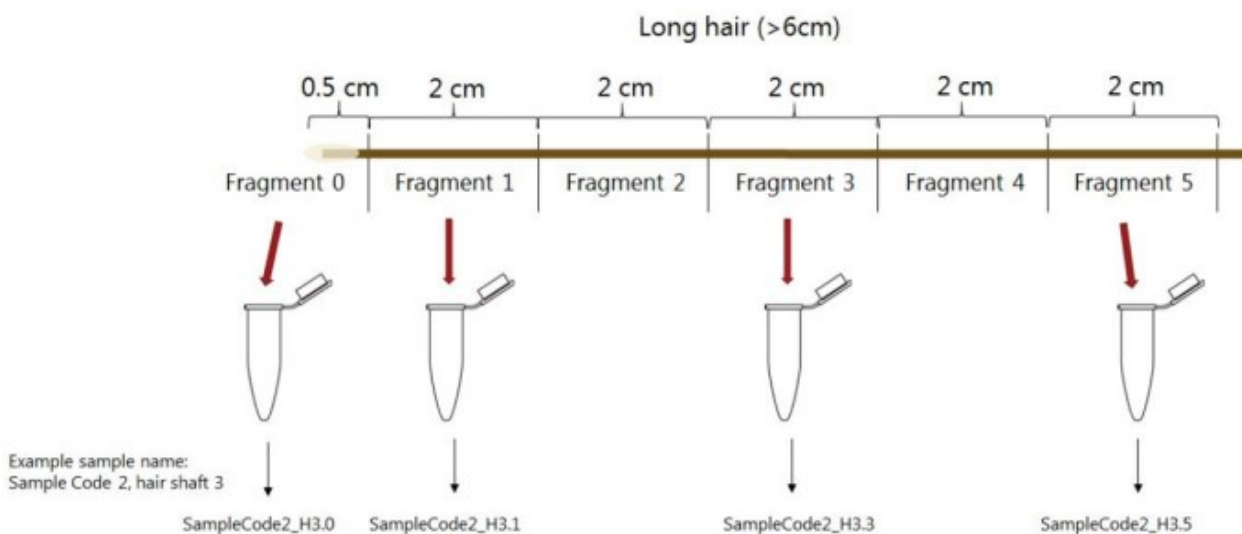


Fig. 8 Preparazione e identificazione del campione di capelli lunghi

Codice del donatore		2	4	5	7	8
<i>Sesso</i>		F	F	F	M	F
<i>Origine</i>		Europea	Europea	Europea	Europea	Europea (con padre iraniano)
<i>Colore naturale dei capelli</i>		Marrone scuro	Marrone scuro	Marrone scuro-nero	Marrone scuro-nero	Marrone scuro
<i>Trattamenti</i>		Nessuno	Punte decolorate	Tinta	Nessuno	Nessuno
<i>Tipo di capello</i>		Liscio	Ondulato	Liscio	Ondulato	Liscio
<i>Numero di capelli raccolti</i>		9	14 (1 bianco)	5	8	5
Capello 1	<i>Codice</i>	2_H1	4_H1	5_H1	7_H1	8_H1
	<i>Lunghezza</i>	25.5 cm	44 cm	23 cm	8.5 cm	40 cm
	<i>Numero del frammento</i>	13	20	12	5	21
	<i>Frammento selezionato</i>	2_H1.1 2_H1.6 2_H1.12	4_H1.1 4_H1.6 4_H1.12 4_H1.19	5_H1.1 5_H1.6 5_H1.10	7_H1.1 7_H1.2 7_H1.3	8_H1.1 8_H1.10 8_H1.19
Capello 2	<i>Codice</i>	2_H2	4_H2	5_H2	7_H2	8_H2
	<i>Lunghezza</i>	30 cm	35 cm	20 cm	8.2 cm	29 cm
	<i>Numero del frammento</i>	16	18	11	5	14
	<i>Frammento selezionato</i>	2_H2.1 2_H2.7 2_H2.14	4_H2.1 4_H2.6 4_H2.11 4_H2.17	5_H2.1 5_H2.5 5_H2.9	7_H2.1 7_H2.2 7_H2.3	8_H2.1 8_H2.6 8_H2.12
Capello 3	<i>Codice</i>	2_H3	4_H3	5_H3	7_H3	8_H3
	<i>Lunghezza</i>	26 cm	22 cm	32 cm	8.3 cm	30 cm
	<i>Numero del frammento</i>	14	11	16	5	15
	<i>Frammento selezionato</i>	2_H3.1 2_H3.6 2_H3.12	4_H3.1 4_H3.5 4_H3.9	5_H3.1 5_H3.7 5_H3.14	7_H3.1 7_H3.2 7_H3.3	8_H3.1 8_H3.7 8_H3.13

Tabella 1. Informazioni sui donatori e caratteristiche dei capelli.

2.3 Estrazione

Le procedure di estrazione sono state condotte utilizzando il Maxwell FSC DNA IQ Casework Kit (Promega) caricato sullo strumento Maxwell 16 (Promega).

2.3.1 Estrazione dai tamponi buccali

I tamponi buccali sono stati estratti seguendo il protocollo di estrazione Solid Support con DNA IQ™ Spin Basket (Fig. 8):

- La soluzione di proteinasi K e l'1-tioglicerolo devono essere scongelati a temperatura ambiente. Accendere il termomixer (Eppendorf) e lasciarlo scaldare a 56°C
- Preparare la miscela di estrazione aggiungendo per ogni campione:

Buffer di estrazione casework	386 µl
Proteinasi K	10 µl
1-tioglicerolo	4 µl
Volume totale	400 µl

- Posizionare la testina del tampone sul fondo di una provetta da 2 ml etichettata e aggiungere la miscela di estrazione.
- Chiudere il coperchio della provetta, vortexare il campione ad alta velocità per 5 secondi, e incubare il campione a 56°C per 30 minuti.
- Centrifugare il campione alla massima velocità per rimuoverne eventuali gocce presenti sul coperchio.
- Mettere il DNA IQ™ Spin Basket in una provetta pulita da 2 ml e trasferire il campione al DNA IQ™ Spin Basket con una

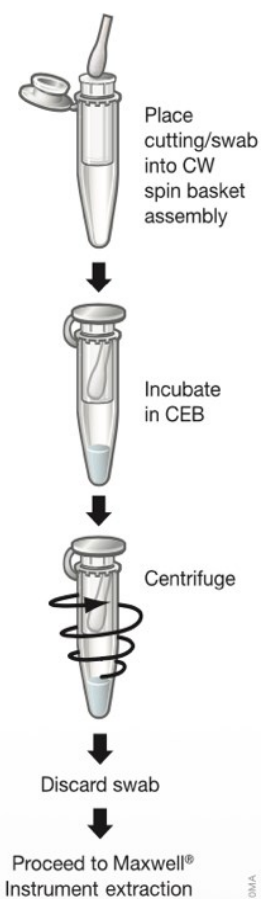


Fig.8 Elaborazione di un campione solido

- pinza, orientando il tampone verso il fondo dello Spin Basket. Trasferire il lisato dalla provetta di incubazione alla nuova provetta.
- Centrifugare alla massima velocità per 2 minuti. Rimuovere con attenzione e scartare il DNA IQ™ Spin Basket.
 - Aggiungere 200 µl di Lysis Buffer alla provetta contenente l'estratto.
 - Chiudere il coperchio della provetta e agitare il campione per 5-10 secondi.
 - Posizionare le cartucce Maxwell 16 LEV da utilizzare e premere verso il basso sulla cartuccia per inserirla sulla postazione.
 - Posizionare una provetta di eluizione etichettata vuota nella postazione dedicata all'eluizione per ciascuna cartuccia e aggiungere 50 µl di tampone di eluizione a ciascuna provetta di eluizione.
 - Staccare la guarnizione superiore delle cartucce e posizionare uno stantuffo nel pozzetto n. 8 di ciascuna cartuccia.
 - Trasferire il campione nel pozzetto n. 1 di ciascuna cartuccia.
 - Accendere lo strumento Maxwell 16 e verificare che l'impostazione LEV sia selezionata.
 - Aprire lo sportello dello strumento e premere "Run/Stop", posizionare il deck tray con le cartucce e le provette di eluizione sull'apposito supporto.
 - Premere "Run/Stop" e chiudere lo sportello per iniziare l'estrazione automatizzata del DNA.
 - Al termine del processo di estrazione, aprire lo sportello e premere "Run/Stop".

- Chiudere i coperchi delle provette e conservare gli estratti a 4°C per brevi periodi o a -20°C per periodi più lunghi (Promega).

2.3.2 Estrazione da sangue

I campioni di sangue sono stati estratti seguendo il protocollo di estrazione “DNA IQ REFERENCE SAMPLE KIT FOR MAXWELL 16” (Promega). Seguendo tale procedura:

- Posizionare le cartucce di tale kit sul supporto con il lato in rilievo rivolto verso il numero stampato sul supporto. Rimuovere il tappo dalle cartucce.
- Posizionare un plunger nel pozzetto 7 di ogni cartuccia.
- Aggiungere due gocce di sangue nel pozzetto 1 delle cartucce, secondo l’ordine precedentemente stabilito.
- Posizionare gli elution tubes nell’apposito supporto magnetico e mettere in ognuno 300 µl di elution buffer.
- Accendere lo strumento e accertarsi che sia in modalità SEV.
- Aprire lo sportello della macchina e trasferire le cartucce precedentemente allestite sull’apposito supporto.
- Posizionare gli elution tubes nei pozzetti di fronte le cartucce.
- Premere il pulsante “Run/Stop” della macchina e chiudere lo sportello. Inizierà il processo di estrazione automatizzata.
- Quando il processo è concluso aprire lo sportello e premere nuovamente il pulsante “Run/Stop”.

- Rimuovere le provette di eluizione dallo strumento e posizionarle sul porta provette magnetico, trasferire la soluzione al loro interno (DNA) nelle corrispondenti provette di tipo Eppendorf da 0,5 ml precedentemente codificate col codice del campione.
- Gettare le cartucce e le elution tubes utilizzate. Pulire lo strumento e le superfici utilizzate.
- Il campione può essere conservato a -20°C fino alla successiva fase di amplificazione.

2.3.3 Estrazione da capelli

I frammenti di capelli selezionati per l'analisi del mtDNA sono stati sottoposti ad ulteriore lavaggio e procedure di decontaminazione:

- Preparare 4 provette Eppendorf da 2 ml per ogni frammento selezionato.
- Etichettare ciascuna provetta e irradiarla con luce UV per 5 minuti.
- Preparare due provette Falcon da 50 ml, una per una soluzione di EtOH al 70% e una per H₂O ultrapura, posizionarli entrambi sotto la luce UV.
- Mettere 1500 µl di EtOH al 70% nella prima provetta e 1500 µl di H₂O ultrapura nella seconda e terza provetta.
- Utilizzando un paio di pinze, posizionare il frammento di capelli nella prima provetta con l'EtOH al 70% e attendere cinque minuti.
- Vortexare il campione a velocità moderata.

- Utilizzando un nuovo paio di pinze sterili, trasferire il frammento nella seconda provetta con H₂O ultrapura.
- Vortexare il campione a velocità moderata.
- Utilizzando un nuovo paio di pinze sterili, trasferire il frammento nella terza provetta con H₂O ultrapura.
- Vortexare il campione a velocità moderata.
- Preparare una capsula Petri e con una nuova pinza sterile trasferire i capelli di 2 cm frammento alla capsula Petri.
- Tagliare il frammento di 2 cm in segmenti di 0,5 cm, in modo che i reagenti di estrazione possano agire pienamente sulla totalità del frammento.
- Trasferire i frammenti da 0,5 cm in una provetta Eppendorf LoBind da 2 ml.
- Conservare i frammenti a -20°C, perché la maggior parte di essi sarà bagnata quindi bisognerà prevenire lo sviluppo di muffe nel campione.

I frammenti di capelli sono stati estratti con il Maxwell FSC DNA IQ Casework Kit (Promega):

- La soluzione di proteinasi K e l'1-tioglicerolo devono essere scongelati a temperatura ambiente. Accendere il Thermomixer (Eppendorf) e lasciarlo scaldare a 56°C.
- Preparare la miscela di estrazione aggiungendo per ogni campione:

Buffer di estrazione	386 μ l
Proteinasi K (18 mg/ml)	10 μ l
1-tioglicerolo	4 μ l
Volume totale	400 μl

- Agitare sul vortex la miscela di estrazione e aggiungere 400 μ l della mix a ciascun Eppendorf LoBind 2 ml tubo contenente i frammenti di capelli.
- Chiudere il coperchio del tubo e sigillarlo con parafilm.
- Vortexare il campione ad alta velocità per 5 secondi e incubare il campione a 56°C per 1 ora
- Centrifugare il campione alla massima velocità per rimuovere eventuali gocce presenti sul coperchio.
- Aggiungere 200 μ l di Lysis Buffer e 100 μ l di DTT alla provetta contenente l'estratto.
- Chiudere il coperchio della provetta e agitare il campione per 5-10 secondi.
- Posizionare le cartucce Maxwell 16 LEV da utilizzare sul supporto e premere verso il basso sulla cartuccia per farla inserire in posizione.
- Posizionare una provetta di eluizione etichettata vuota nella postazione per l'eluizione per ciascuna cartuccia inserita e aggiungere 50 μ l di tampone di eluizione a ciascuna provetta di eluizione.
- Staccare la guarnizione superiore delle cartucce e posizionare uno stantuffo nel pozzetto n. 8 di ciascuna cartuccia.

- Trasferire i frammenti di capelli nel pozzetto n. 1 di ciascuna cartuccia.
- Accendere lo strumento Maxwell 16 e verificare che l'impostazione LEV sia selezionata.
- Aprire lo sportello dello strumento e premere “Run/Stop”, posizionare il deck tray con le cartucce e le provette di eluizione sull'apposito supporto.
- Premere "Run/Stop" e chiudere lo sportello per iniziare l'estrazione automatizzata del DNA (Fig. 9)
- Al termine del processo di estrazione, aprire lo sportello e premere “Run/Stop”.
- Verificare la presenza di capelli nel pozzetto n. 1, se il processo di estrazione automatizzato lo era con successo i frammenti di capelli dovrebbero diventare polvere fine.
- Chiudere i coperchi delle provette e conservare gli estratti a 4°C per brevi periodi o a -20°C per periodi più lunghi.

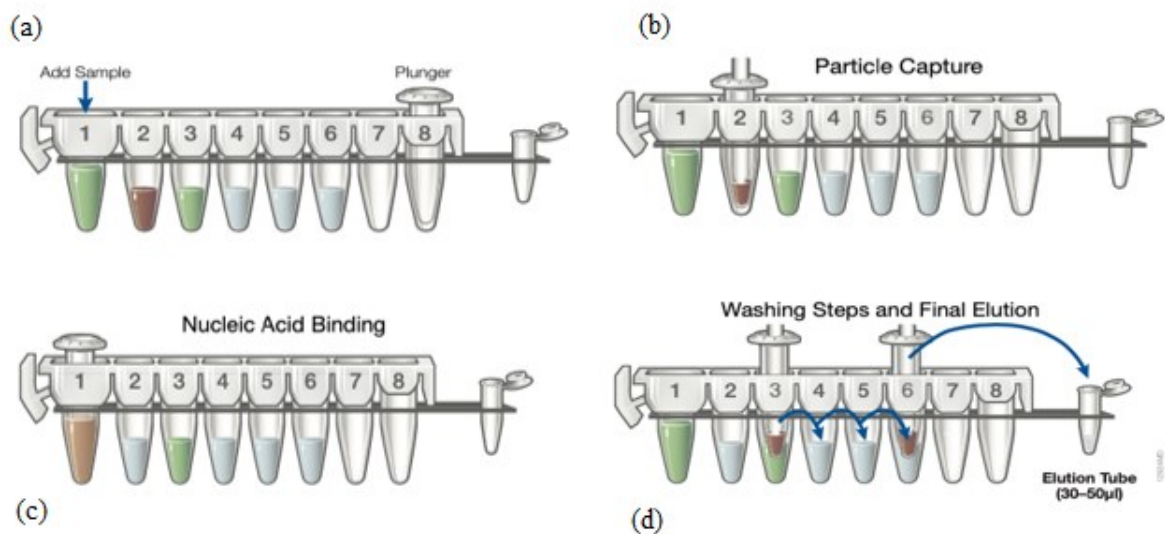


Fig. 9 Processo di estrazione automatizzato utilizzando lo strumento Maxwell16, una barra magnetica coperta da uno stantuffo sposta le particelle paramagnetiche da un pozzo all'altro.

2.4 Quantificazione del mtDNA nei frammenti di capelli

La quantificazione del mtDNA estratto dai campioni è stata ottenuta utilizzando una tecnica recentemente sviluppata, la qPCR, che utilizza un target di nDNA (181 bp), due target di mtDNA (69bp e 143 bp) e un controllo positivo interno (IPC) per monitorare la potenziale inibizione della PCR. Attraverso l'uso di un oligonucleotide sintetico, il test consente di calcolare il numero di copie del mtDNA. Lo standard DNA - G1521 (Promega) è stato utilizzato per costruire la curva standard, attraverso una serie di 7 diluizioni a partire da una concentrazione iniziale di 62,5 ng/μl fino a 0,004 ng/μl, analizzate in triplicato. Per calcolare gli equivalenti del genoma mitocondriale (mtGE) presenti in tali standard, l'oligo sintetico, mirato al sistema mtND1, e il suo complementare, sono stati combinati per simulare dsDNA e utilizzati come standard di mtDNA. Il mtND1 oligonucleotide e il complementare avevano una concentrazione iniziale di 100 μM corrispondente a $6,0 \times 10^{13}$ copie/μl, una serie di cinque diluizioni da una concentrazione iniziale di 600.000 copie “a doppio filamento”/μl fino a 960 copie/μl è stata utilizzata per costruire la curva standard per calcolare il mtGE presente nello standard G1521 (Promega)¹⁴.

Per ogni campione testato, la multiplex è stata fatta in duplicato in reazioni da 20 μl, i volumi e le concentrazioni dei componenti sono riassunti nella tabella 2:

4x QuantiNova Multiplex PCR MM	5 µl
IPC – primer Fw	0.5 µl
IPC – primer Rv	0.5 µl
IPC – probe	0.5 µl
IPC – oligo	1 µl
nuRNU2_181bp – primer Fw	0.9 µl
nuRNU2_181bp – primer Rv	0.9 µl
nuRNU2_181bp – probe	1 µl
mtND1 – primer Fw	1 µl
mtND1 – primer Rv	1 µl
mtND1 – probe	1 µl
mt143bp – primer Fw	0.5 µl
mt143bp – primer Rv	0.5 µl
mt143bp – probe	0.5 µl
Standard or Sample	2 µl
H ₂ O	3,2 µl

Tabella 2. Saggio di Real Time PCR

Procedura:

- Scongellare a temperatura ambiente la 4x QuantiNova Multiplex PCR Master Mix, campioni, primer, sonde e acqua priva di RNasi.
- Accendere il Rotor-Gene 6000 Corbett (Qiagen).
- Vortexare e centrifugare reagenti e campioni.
- Preparare una mix di reazione secondo la Tabella 2 per ciascun campione.
- Programmare il termociclatore per la Real-time PCR come descritto nella Tabella 3.

Stage	Step	Temperature	Time
Hold	Activate the enzyme	95°C	2 minutes
40 Cycles	Denature	95°C	3 seconds
	Anneal and extend	60°C	30 seconds
Hold		10°C	Hold

Tabella 3. Programma per la Real Time PCR

- Posizionare le provette nel termociclatore per la Real-time PCR e avviare il programma.
- Eseguire l'analisi dei dati.

2.5 Quantificazione dei campioni di riferimento

Per quantificare il DNA presente nei campioni di riferimenti dei volontari abbiamo utilizzato il Qubit Kit. Nel kit sono presenti il reagente di dosaggio concentrato, tampone di diluizione e standard di DNA prediluiti. Il test è selettivo per il doppio filamento DNA e fornisce concentrazioni accurate da 10 pg/μl a 100 ng/μl. La quantificazione viene effettuata in un volume totale di 200 μl come segue:

- Preparare le provette da 0,5 ml necessarie per l'analisi ed etichettarle con i rispettivi codici per ogni campione
- Preparare la soluzione di lavoro Qubit diluendo il reagente Qubit dsDNA HS fornito 1:200 in Qubit dsDNA HS Buffer.
- Aggiungere 190 μl di Qubit working solution a ciascuna delle provette utilizzate per gli standard. Aggiungere 10 μl di ogni standard Qubit alle provette appropriate, vortexare per 2-3 secondi.

- Aggiungere 198 μ l di Qubit working solution a ciascuna delle provette, poi aggiungere 2 μ l di campione nella rispettiva provetta per il dosaggio fino a raggiungere un volume di 200 μ l. Vortexare per 2-3 secondi.
- Tutte le provette devono incubare per 2 minuti a temperatura ambiente
- Accendere il Qubit 2.0 Fluorometer, premere DNA, selezionare dsDNA High Sensibility come tipo di analisi. Selezionare Sì nella schermata standard per leggere il nuovo standard.
- Posizionare la provetta Standard n. 1 nella camera del campione, chiudere il coperchio e premere Leggi. Una volta che la lettura è completata rimuovere la provetta.
- Posizionare la provetta Standard n. 2 nella camera del campione, chiudere il coperchio e premere Leggi. Una volta che la lettura è completata rimuovere la provetta.
- Iniziare ora la lettura dei campioni posizionando ciascuna provetta nella camera apposita, la concentrazione viene calcolata dallo strumento una volta che l'utente ha impostato la misura unità richieste.
- Ogni campione viene analizzato due volte e quindi viene fatta una media del valore come concentrazione finale espressa in ng/ μ l.

2.6 Sequenziamento con HID Ion GeneStudio S5 System

Per valutare la presenza e l'incidenza di varianti eteroplasmatiche nei mitogenomi nei fusti dei capelli, è stato utilizzato l'Applied Biosystems Precision ID mtDNA Whole Genome Panel (Thermo Fisher Scientific). Questo pannello è un approccio di sequenziamento di nuova generazione (NGS) per l'analisi del DNA mitocondriale appositamente progettato per applicazioni forensi. La procedura utilizza il sistema Ion Chef, il sequenziamento e l'analisi con lo Ion ChefTM System e il ConvergeTM Software. Il pannello per l'intero genoma del mtDNA è caratterizzato da 162 ampliconi composto da due pool multiplex (con 81 coppie di primer ciascuno) mirato all'intero genoma mitocondriale umano (16.569 bp) (Fig.10).

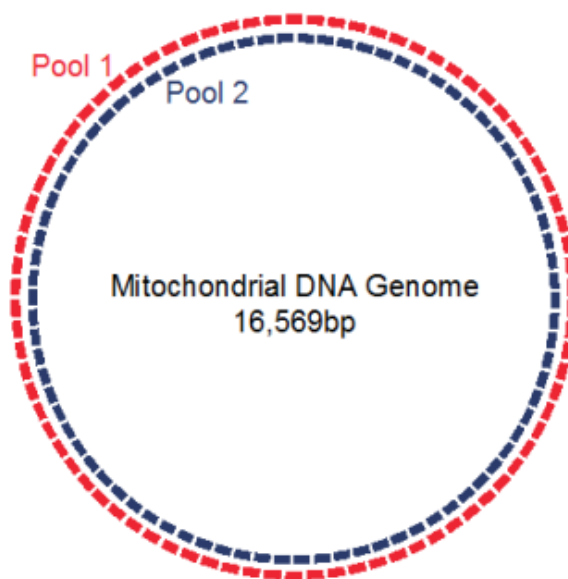


Fig. 10. Rappresentazione visiva del sistema a doppio pannello. Non mostra il numero reale di ampliconi.

2.6.1 Reazione di amplificazione del mtDNA target

Il Precision ID library kit (Thermo Fisher Scientific) comprende tutti i reagenti per la preparazione delle librerie per i pannelli Precision ID e richiede un minimo di 0,1 ng di DNA genomico, equivalente a 2.900 copie del mtDNA per la reazione di amplificazione del bersaglio. Dopo le procedure di quantificazione, i campioni di frammenti di capelli sono stati preparati per la reazione di amplificazione utilizzando il *metodo 2-in-1*, come riportato nel manuale (Precision ID mtDNA Panel with the HID Ion S5/HID Ion GeneStudio S5 System, MAN0017770, rev. C.0):

- Preparare due master mix, uno per ogni pool.
- Preparare la master mix per il pannello Precision ID mtDNA Pool 1 (Tabella 4), per poi aggiungerla ad ogni provetta di reazione:

5X Ion AmpliSeq Mix HiFi	4 μ l
Pannello Precision ID mtDNA Pool 1	10 μ l
gDNA, 0,1 ng	6 μ l
Volume totale	20 μ l

Tabella 4. Pool 1 del pannello Precision ID mtDNA

- Preparare la master mix per il pannello Precision ID mtDNA Pool 2 (Tabella 5), per poi aggiungerla ad ogni provetta di reazione:

5X Ion AmpliSeq Mix HiFi	4µl
Pannello Precision ID mtDNA Pool 2	10µl
gDNA, 0,1 ng	6µl
Volume totale	20µl

Tabella 5. Pool 2 del pannello Precision ID mtDNA

- Vortexare le provette di reazione, e poi centrifugare per raccogliere le goccioline. Per amplificare le regioni target, eseguire il programma descritto nella Tabella 6:

Stage	Step	Temperature	Time
Hold	Activate the enzyme	99°C	2 minutes
26 Cycles	Denature	99°C	15 seconds
	Anneal and extended	60°C	4 minutes
Hold		10°C	Hold

Tabella 6. Programma di amplificazione della regione target del mtDNA con Precision ID

- Al termine dell'amplificazione trasferire 10 µl da ciascuna pool in una nuova provetta, per un totale di 20 µl.

Al contrario i campioni di riferimento sono stati preparati per la reazione di amplificazione utilizzando il *metodo conservativo*:

- Preparare due master mix, una per ogni pool.
- Preparare la master mix per il pannello Precision ID mtDNA Pool 1, per poi aggiungere ad ogni provetta di reazione:

5X Ion AmpliSeq™ HiFi Mix	2μ
Precision ID mtDNA panel Pool 1	5μ
gDNA, 0.1 ng	1μ
Nuclease-free water	2μ
Volume totale	10μ

Tabella 7. Pool 1 del pannello Precision ID mtDNA

- Preparare la master mix per il pannello Precision ID mtDNA Pool 2 (Tabella 8), e poi aggiungerla ad ogni provetta di reazione:

5X Ion AmpliSeq™ HiFi Mix	2μ
Precision ID mtDNA panel Pool 2	5μ
gDNA, 0.1 ng	1μ
Nuclease-free water	2μ
Volume totale	10μ

Tabella 8. Pool 2 del pannello Precision ID mtDNA

- Per amplificare le regioni target, eseguire il programma descritto nella Tabella 9:

Stage	Step	Temperature	Time
Hold	Activate the enzyme	99°C	2 minutes
21 Cycles	Denature	99°C	15 seconds
	Anneal and extended	60°C	4 minutes
Hold		10°C	Hold

Tabella 9. Programma di amplificazione della regione target del mtDNA con l'ID precision.

- Al termine dell'amplificazione trasferire 10 μl dalla provetta del Pool 2 alla provetta del Pool 1, per un totale di 20 μl.

2.6.2 Digestione degli ampliconi

- Aggiungere 2 μ l di FuPa Reagent a ciascun campione amplificato. Il volume totale sarà di 22 μ l.
- Vortexare le provette e centrifugare per raccogliere le goccioline.
- Accendere il termociclatore e impostare le condizioni descritte nella Tabella 10.

Temperatura	Tempo
50°C	2 minuti
55°C	15 secondi
60°C	4 minuti
10°C	Tenere fino a 1 ora

Tabella 10. Condizioni per la digestione degli ampliconi sul termociclatore

2.6.3 Legame degli adattatori e purificazione della libreria

Poiché più librerie verranno sequenziate sullo stesso chip, sarà necessario aggiungere dei codici a barre diversi per ciascuna libreria. Gli adattatori per i Barcode sono forniti alla concentrazione appropriata nel Precision ID IonCode 1-96 kit. Procedere come segue:

- Aggiungere i componenti visualizzati nella Tabella 11 a ciascuna provetta di reazione contenente il ampliconi parzialmente digeriti. Seguire l'ordine elencato e assicurarsi di aggiungere la DNA ligasi per ultima.

Order of addition	Component	Volume
1	Switch Solution (yellow cap)	4 μ l
2	Precision ID IonCode Barcode Adapter	2 μ l
3	DNA ligase (blue cap)	2 μ l
-	Total volume (including 22 μ l of digested amplicon)	30 μ l

Tabella 11. Componenti per la reazione di ligazione

- Vortexare le provette e centrifugare per raccogliere le goccioline.
- Caricare i campioni nel termociclatore per eseguire la reazione di ligazione utilizzando il programma riportato nella tabella 12.

Temperatura	Tempo
22	30 minuti
68	10 minuti
10	Tenere fino a 1 ora

Tabella 12. Condizioni per la reazione di ligazione nel termociclatore

- Preparare i reagenti per le fasi di purificazione, scongelando l'Agencourt AMPure XP Reagent a temperatura ambiente, e agitare sul vortex per disperdere le sfere prima di utilizzarlo.
- Preparare l'etanolo al 70% per i passaggi successivi combinando 230 μ l di etanolo con 100 μ l di acqua priva di nucleasi per ogni campione.
- Aggiungere 45 μ l di Agencourt AMPure XP Reagent a ciascuna libreria.

- Pipettare su e giù per 5 volte per miscelare accuratamente la sospensione di microsfele con il DNA, incubare la miscela per 5 minuti a temperatura ambiente.
- Mettere le provette su uno stendino magnetico e incubare per 2 minuti o fino a che la soluzione non diventa trasparente.
- Rimuovere con cautela, ed eliminare il surnatante senza disturbare il pellet.
- Aggiungere 150 μ l di etanolo al 70% appena preparato, quindi ruotare la provetta sulla griglia magnetica per lavare le biglie. Rimuovere il surnatante senza disturbare il pellet. Ripetere questo passaggio per un secondo lavaggio.
- Assicurarsi che tutte le goccioline di etanolo vengano rimosse, asciugando all'aria le sfere a temperatura ambiente per 5 minuti tenendo le provette sulla griglia magnetica. L'etanolo residuo inibisce la reazione di PCR quindi sarà necessario eliminarlo.
- Successivamente, rimuovere le provette dallo stendino e aggiungere 50 μ l di low TE al pellet per disperdere le perline.
- Chiudere bene i coperchi delle provette e agitare su vortex.
- Incubare per 5 minuti a temperatura ambiente, centrifugare per raccogliere le goccioline.
- Riposizionare le provette sulla griglia magnetica per 2 minuti.
- Recuperare la soluzione in una nuova provetta etichettata.

2.6.4 Quantificazione della libreria mediante qPCR

Dopo l'eluizione, le librerie Precision ID vengono quantificate mediante qPCR con il kit di quantificazione Ion Library TaqMan. La quantificazione viene eseguita come segue:

- Vortexare e centrifugare i campioni per raccogliere le goccioline.
- Preparare diluizioni 1:100 aggiungendo 2 µl di libreria e 198 µl di Nuclease-free water.
- Conservare i campioni rimanenti a 4 C°.
- Preparare tre diluizioni seriali della libreria di controllo E. coli DH10B (definiti anche Standard) seguendo le concentrazioni elencate nella tabella 13.

Standards	Control Library Volume	Nuclease-free water volume	Concentration
1	5 µl (undiluted)	45 µl	6,8 pM
2	51 µl Std 1	45 µl	6,8 pM
3	51 µl Std 2	45 µl	6,8 pM

Tabella 13. Concentrazioni per la diluizione seriale della libreria di controllo

- Preparare la mix di reazione per ciascun campione, per il controllo negativo, e le librerie di controllo. Per ogni reazione, combinare 10 µl di Ion Library qPCR Master Mix e 1 µl di Ion Library TaqMan Quantitation Assay 20x in una provetta, e mescolare accuratamente.
- Aliquotare 11 µl in ciascuna provetta di reazione.
- Aggiungere 9 µl della libreria diluita (1:100) alle provette di reazione, per un volume totale di 20 µl.

- Impostare il programma per la Real-time PCR sul Rotor-Gene 6000 Corbett (Qiagen) come descritto nella tabella 14.

Stage	Temperature	Time
Hold	50 °C	2 minutes
Hold	95 °C	20 seconds
40 Cycles	95 °C	3 seconds
	60 °C	32 seconds

Tabella 14. Programma per la Real-time PCR per la quantificazione delle librerie

- Eseguire la reazione e raccogliere i dati in tempo reale.

2.6.5 Diluizione, raggruppamento e archiviazione delle librerie

- Al completamento della qPCR, calcolare la concentrazione media di ciascuna libreria non diluita utilizzando la seguente equazione:

$$\text{Concentrazione media della libreria non diluita} = (\text{quantità media qPCR}) \times (\text{diluizione della libreria})$$

- Diluire le librerie come descritto nella tabella 15.

Diluite to	Minimum volume	Templating size in Planned Run setup
40 pM	25 µl	200 BP

Tabella 15. Diluizione della libreria

Dopo aver diluito la libreria di campioni alla sua concentrazione target (pM), raggruppare volumi uguali di più librerie. Utilizzare le librerie raggruppate nelle reazioni di preparazione dei modelli sullo Ion Chef Instrument (ThermoFisher Scientific).

2.6.6 Preparazione del template su *HID Ion Chef*

La preparazione del template sullo strumento Ion Chef consiste nell'amplificazione clonale di sequenze target in una libreria su Ion Sphere Particles (ISP). Queste sono caricate dallo strumento su un chip di sequenziamento Ion 520. Per ogni chip Ion 520, un totale di 25 campioni possono essere sequenziati contemporaneamente ed è necessario creare una corsa pianificata seguendo le istruzioni (Precision ID mtDNA Panels, Application Guide, TFS) utilizzando il software Torrent Suite. Sono state eseguite due corse poiché non tutti i campioni potevano essere caricati su un unico chip. È stato utilizzato per entrambe le corse il Chip Ion 520™.

La configurazione dello strumento per la corsa richiede le seguenti procedure:

- Disimballare le cartucce dei reagenti Ion S5 Precision ID Chef 45 minuti prima dell'uso, per consentire la stabilizzazione a temperatura ambiente, pipettare 25 µl di pool di librerie diluite sul fondo dell'appropriata Ion Chef Library Sample Tube (Fig. 10).

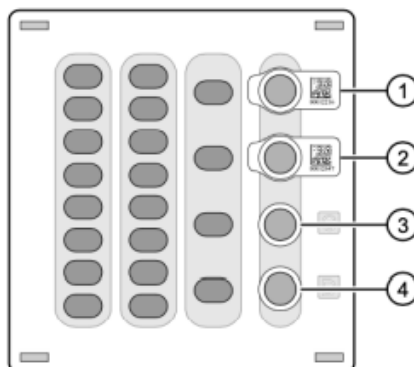


Fig. 10 Cartucce per i reagenti dello Ion S5 Precision ID Chef. Posizione 1 (pool di librerie di DNA), Posizione 2 (pool di librerie di DNA), Posizione 3 (NaOH) e Posizione 4 (provetta vuota)

- Coprire e conservare le due librerie di DNA diluite in ghiaccio fino a quando non si è pronti per caricarle sullo strumento.
- Rimuovere tutte le cartucce e i materiali da utilizzare dalla confezione, poi metterli sulla postazione accanto allo Ion Chef.
- Caricare i puntali delle pipette e la piastra PCR.
- Caricare le cartucce dei reagenti e le soluzioni.
- Caricare le provette di ripristino e le cartucce di arricchimento v2.
- Caricare il Chip-loading sulla centrifuga.
- Verificare che i materiali di consumo siano installati correttamente (Fig. 11).

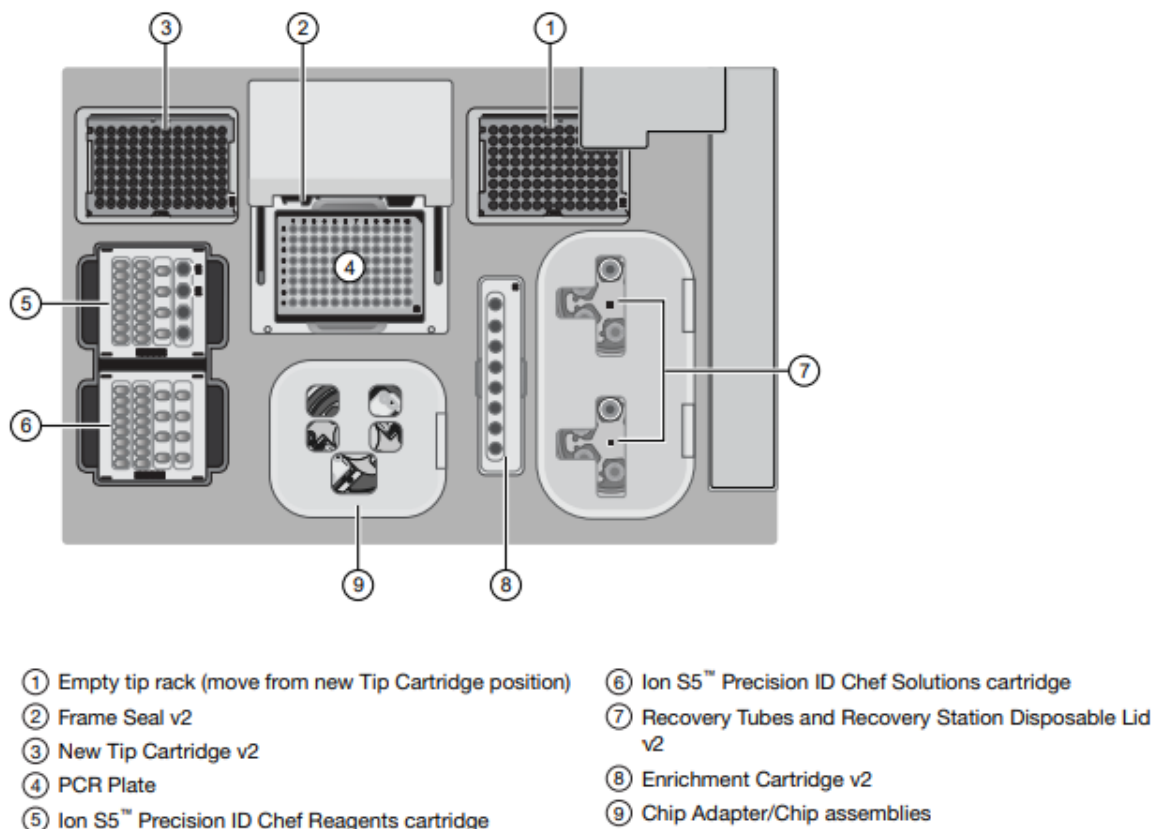


Fig. 11 Strumento Ion Chef caricato di tutti i materiali di reazione

- Avviare la corsa Ion Chef selezionando l'opzione sullo schermo touch nella home.
- Seguire passo dopo passo per fare in modo che lo strumento vada verso la configurazione seguendo le istruzioni sullo schermo.
- Quando richiesto, chiudere lo sportello dello strumento. Il sistema si attiverà.
- Toccare avvia controlli per avviare Deck Scan, in cui lo strumento esegue la scansione dei codici a barre di tutti materiali di utilizzo e dei reagenti.
- Quando la Deck Scan sarà completata, toccare sul display Data Destination, durante il quale lo strumento visualizza il tipo di kit, il tipo di chip, i codici a barre del chip e le corse pianificate. Se tutte le opzioni sono corrette, tocca Next.
- Il menu Run Option consente di impostare l'ora di completamento dell'esecuzione, se necessario. Successivamente premere Start Run.
- Inizializzare lo Ion S5™ almeno 40 minuti prima che Ion Chef™ System termini il caricamento del chip.
- Al termine della corsa, scaricare lo strumento e sequenziare il chip.

2.6.7 Sequenziamento con HID Ion GeneStudio S5™ System

Il sequenziatore HID Ion S5™ è equipaggiato per verificare la compatibilità di ogni chip e il consumo dei materiali che vengono caricati durante l'inizializzazione e il sequenziamento. Prima dell'inizio della reazione di sequenziamento:

- Disimballare le cartucce Ion S5™ Precision ID Sequencing Reagents 45 minuti prima utilizzare e lasciarli equilibrare a temperatura ambiente.
- Disimballare il flacone della soluzione di lavaggio Precision ID. Capovolgere la bottiglia 5 volte, per mescolare bene. Togliere dal suo sacchetto sottovuoto, togliere il tappo rosso dalla soluzione di lavaggio e dai detergenti immediatamente prima di installare lo strumento.
- Inizializzare toccando Initialize nel menu principale della schermata touch.
- Rimuovere il flacone della soluzione di lavaggio Precision ID per accedere e svuotare i rifiuti serbatoio, e reinserire il serbatoio dei rifiuti vuoto.
- Sostituire la cartuccia Ion S5 Precision ID Sequencing Reagents esaurita con una cartuccia nuova equilibrata a temperatura ambiente.
- Installare il nuovo flacone della soluzione di lavaggio Precision ID.
- Lasciare il chip che è stato sequenziamento nel cumulo dei chip, chiudere lo sportello e avviare l'inizializzazione.
- Al termine, toccare Run nel menu principale.
- Rimuovere il chip di sequenziamento usato e inserire il nuovo chip di sequenziamento

- Verificare che sia selezionata la corsa pianificata corretta, premere Review, confermare le impostazioni sono corretti o apportare modifiche.
- Chiudere lo sportello dello strumento, premere Start-Run per iniziare la corsa di sequenziamento.

2.7 Analisi dei dati

Al completamento dell'MPS, tutti i raw data sono stati processati con il Torrent Suite Software (v5.12.3) e le reads allineate con la sequenza di riferimento del DNA mitocondriale umano rCRS (Revised Cambridge Reference Sequence). L'analisi primaria dei dati è stata eseguita con i plug-in Coverage Analysis (v5.12.0.0) e il Variant Caller (v5.12.0.4, TVC). Successivamente per confermare e revisionare le chiamate ottenute dal TVC, è stata effettuata un'analisi secondaria con il software Integrative Genomics Viewer IGV (ref). Il valore minimo di copertura è stato impostato a 20x, e la soglia per il rilevamento dell'eteroplasmia è stata fissata al 10% della copertura totale della posizione.

Capitolo 3

RISULTATI

3.1 Risultati della quantificazione del mtDNA

Per valutare l'efficienza e l'affidabilità della reazione qPCR sono stati presi in considerazione diversi parametri, come la linearità (R^2), l'efficienza di amplificazione e la pendenza. I dati sono stati analizzati utilizzando il software Rotor Gene Q v. 2.1.0 (Qiagen), il *cycle threshold* (il ciclo soglia) è stato calcolato automaticamente dal software ed è stato applicato a tutti i campioni (Fig. 13,14). La curva di calibrazione è stata costruita utilizzando lo standard Human Genomic DNA - G1521 (Promega) disponibile in commercio, attraverso una serie di 7 diluizioni a partire da una concentrazione iniziale di 62,5 ng/ μ l verso il basso a 0,004 ng/ μ l. La linearità e l'efficienza della qPCR hanno dato risultati soddisfacenti, confermando la robustezza del saggio. Anche l'IPC incluso nel test non è riuscito a rilevare la presenza di eventuali inibitori della PCR (Fig. 15).

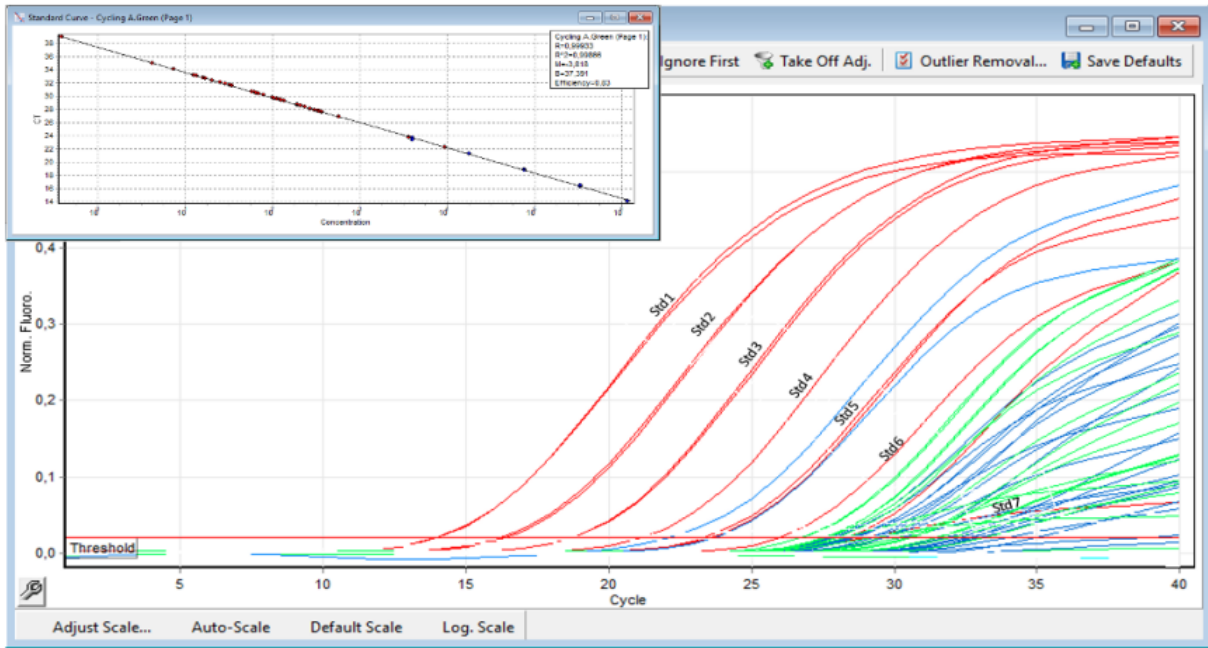


Fig. 13. Curva della qPCR per il target mtDNA di 143bp. Il numero del ciclo è rappresentato sull'asse x mentre l'intensità della fluorescenza è sull'asse y.

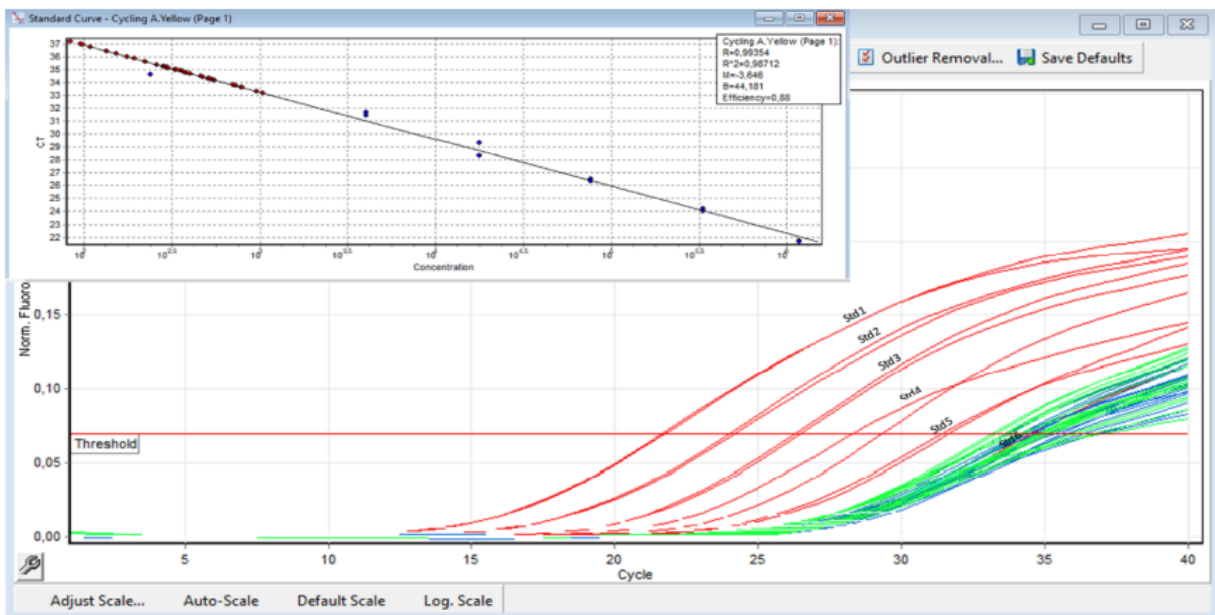


Fig. 14. Curva della qPCR per il target mtDNA ND1 di 69bp. Il numero del ciclo è rappresentato sull'asse x mentre l'intensità della fluorescenza è sull'asse y.

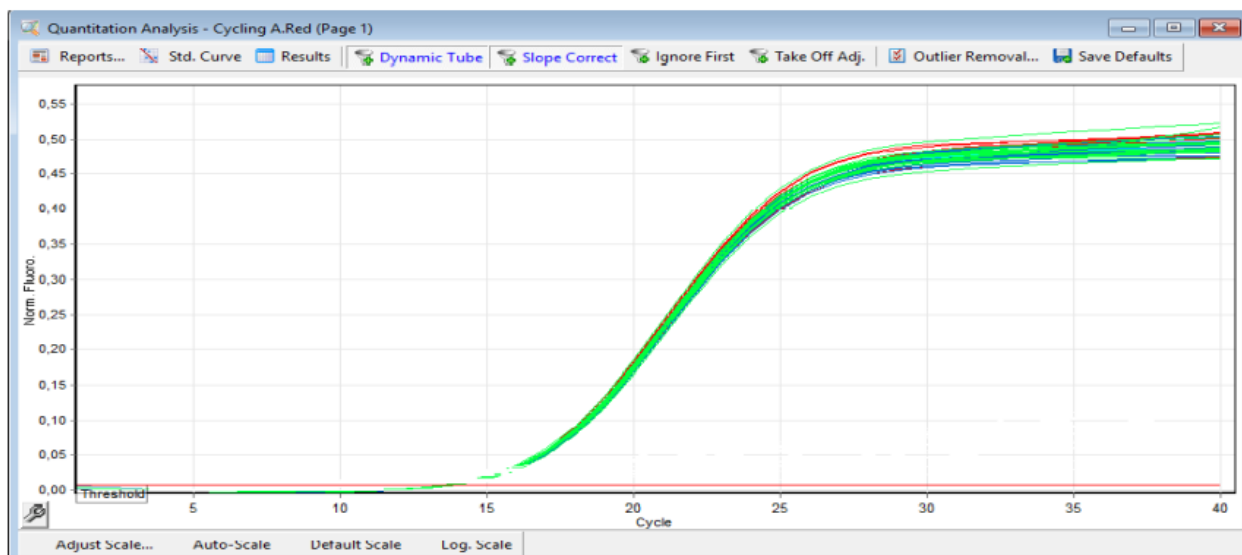


Fig. 15. Curva qPCR IPC. Il numero del ciclo è rappresentato sull'asse x mentre l'intensità della fluorescenza è sull'asse y

È noto che la quantità di mtDNA diminuisce lungo la lunghezza del fusto del capello¹⁶, per valutare se tutti i frammenti soddisfacevano il range quantitativo richiesto per l'analisi con il Massive Parallel Sequencing, è stato utilizzato una tetraplex qPCR per quantificare il mtDNA. Il test qPCR includeva anche un target di nDNA, per valutare la presenza di potenziali contaminazioni, ma non ha rilevato alcun nDNA nei frammenti di capelli analizzati, escludendo l'ipotesi di contaminazione durante il processo di estrazione. I risultati della quantificazione sono visualizzati nella tabella 17, dove abbiamo incluso anche l'indice di degradazione (DI) in base al rapporto tra i frammenti di mtDNA grandi (143 bp) e piccoli (69 bp). Più grande è l'indice di degradazione, minore è il numero di frammenti di mtDNA recuperati dai campioni. Quando il valore del DI è 1 o inferiore il mtDNA è definito non degradato, mentre un $DI > 1$ indica che negli estratti sono presenti meno

frammenti di 143 bp o più grandi. Un DI indeterminato suggerisce che tutto il mtDNA di dimensione inferiore a 143 bp presente in un campione è degradato. Il numero di copie del mtDNA per il target mtND1 (69bp) varia da 55,9 a 1004,9 (mtGE/ μ l), mentre per il target più lungo 143bp i valori vanno da 0,4 a 573,5 (mtGE/ μ l). L'unico frammento che non è stato rilevato dal target mtDNA più lungo era 2_H1.12 risultante in un DI indeterminato. È interessante notare che questa è l'estremità distale o punta del capello 2_H1 suggerendo che il mtDNA in effetti diminuisce all'aumentare della distanza dal bulbo pilifero. Dai dati raccolti l'indice di degradazione sembra aumentare in alcuni dei fusti di capelli con l'aumentare della distanza dal bulbo, ma l'andamento non è costante sia lungo lo stesso capello che tra i capelli raccolti. Ad esempio, prendendo in considerazione i segmenti più prossimali alla radice in ciascuno dei capelli analizzati, l'indice di degradazione varia da 1,2 (2_H2.1) a 4,1 (8_H3.1). D'altra parte, quando osserviamo il frammento medio di ciascun capello, il DI varia da 1,1 (8_H3.7) a 133,5 (8_H1.10). Infine, le punte o le estremità distali mostrano valori DI che variano da 2,4 (4_H3.9) a 364,7 (2_H2.14).

Campioni	Target mtND1 (mtGE/μl) 69bp	Target mt143 (mtGE/μl) 143bp	Indice di degradazione mtND1/mt143bp
2_H1.1	168,9	116,2	1,5
2_H1.6	109,3	2,4	45,5
2_H1.12	73,1	Indeterminato	Indeterminato
2_H2.1	617,2	497	1,2
2_H2.7	144,1	6,4	22,5
2_H2.14	145,9	0,4	364,8
2_H3.1	194,5	86,3	2,3
2_H3.6	146,2	14,8	9,9
2_H3.12	55,9	4,4	12,7
4_H1.1	179,8	65,9	2,7
4_H1.6	279,5	20,42	13,7
4_H1.12	343,1	16,2	21,2
4_H1.19	219,1	4,2	52,2
4_H2.1	422,9	113,1	3,7
4_H2.6	544,5	190,4	2,9
4_H2.11	347,6	43	8,1
4_H2.17	248,6	15,1	16,5
4_H3.1	566,2	313,5	1,8
4_H3.5	97,3	86,4	1,1
4_H3.9	405,8	170,9	2,4
5_H1.1	749,8	355,5	2,1
5_H1.6	252,8	12,5	20,2
5_H1.10	252,2	49,4	5,1
5_H2.1	547,7	176,8	3,1
5_H2.5	653,2	304,6	2,1
5_H2.9	323,7	59,8	5,4
5_H3.1	1004,9	573,5	1,8
5_H3.7	246,6	95,7	2,6
5_H3.14	311,7	28,8	10,8
7_H1.1	117	32,1	3,6
7_H1.2	213,3	132,5	1,6
7_H1.3	165,7	7,1	23,3
7_H2.1	288,8	98	2,9
7_H2.2	190,7	27,5	6,9
7_H2.3	227,9	33,3	6,8
7_H3.1	158	55,9	2,8
7_H3.2	211,5	42,4	5
7_H3.3	177,3	11	16,1
8_H1.1	197,9	153,3	1,3
8_H1.10	200,3	1,5	133,5
8_H1.19	57,5	2,2	26,1
8_H2.1	493,3	298,5	1,7
8_H2.6	135,2	6,1	22,2
8_H2.12	113,4	2,6	43,6
8_H3.1	129,9	31,9	4,1
8_H3.7	228,9	212,3	1,1
8_H3.13	107,7	4,7	22,9

Tabella 17. Valori della quantificazione del mtDNA e valori dell'indice di degradazione per tutti i campioni analizzati.

La variazione dell'indice di degradazione è stata esaminata anche su base individuale rappresentando graficamente i dati nei grafici 1-5. Scegliamo di raggruppare i frammenti della sezione prossimale, mediana, e distale per valutare la variazione del DI tra capelli di un individuo:

I dati sui capelli dell'individuo 2 sono rappresentati nel Grafico 1, come discusso in precedenza il frammento distale 2_H2.12 non è stato quantificato correttamente a causa della mancanza di 143 bp o frammenti più grandi di mtDNA, portando a un indice di degradazione indeterminato. D'altro canto, si può apprezzare l'aumento del DI lungo l'intera lunghezza dell'H2 e dell'H3, mostrati in arancione e verde rispettivamente.

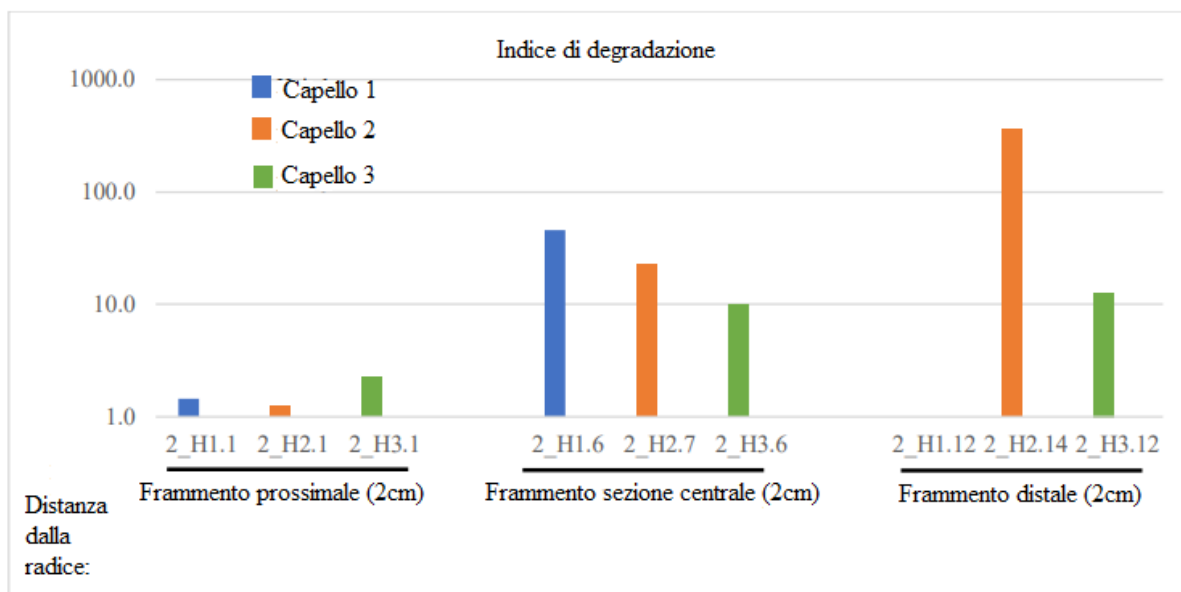


Grafico 1. Variazione dell'indice di degradazione dei campioni di capelli dell'individuo 2

I dati sui capelli del volontario 4 sono illustrati nel Grafico 2. A parte H3, che non mostra un graduale incremento del DI, il primo e il secondo capello esaminati mostrano un relativo aumento del DI. In questo caso per H1 e H2, abbiamo analizzato due frammenti della sezione centrale per ogni capello poiché i capelli erano piuttosto lunghi e avevano subito la decolorazione.

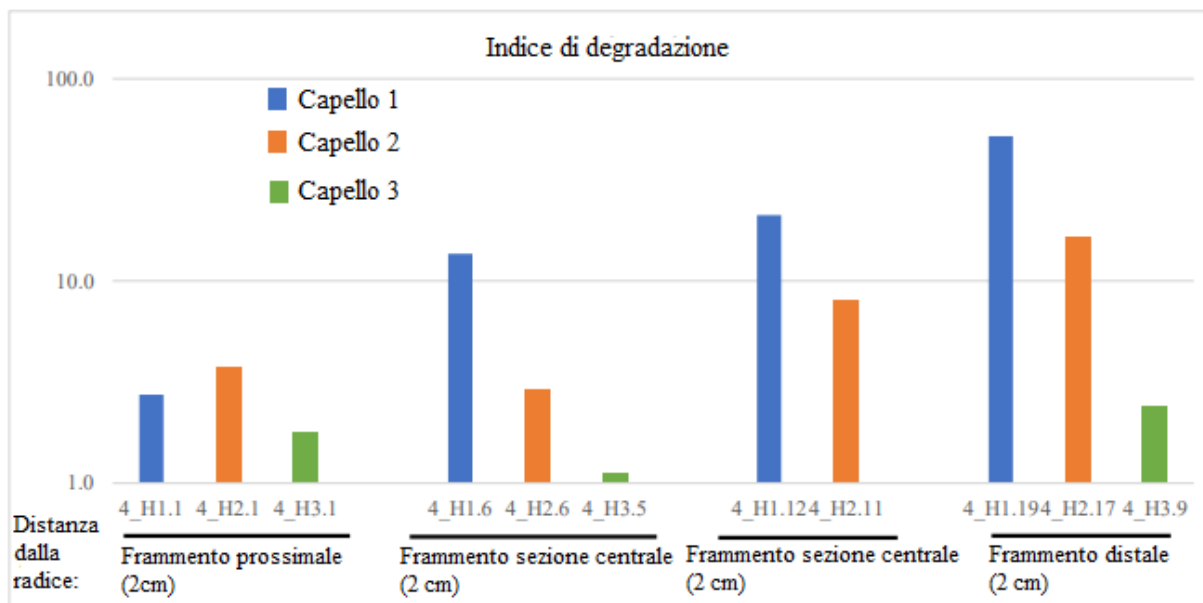


Grafico 2. Variazione dell'indice di degradazione dei campioni di capelli dell'individuo 4

I frammenti di capelli dell'individuo 5 sono gli unici selezionati per l'analisi MPS a causa del loro DI basso, aventi solo due frammenti con valori superiori a 10. Il graduale aumento del DI atteso può essere osservato in tutti i capelli, anche se 5_H1.6 ha un DI più alto della punta dello stesso capello (Graf. 3).

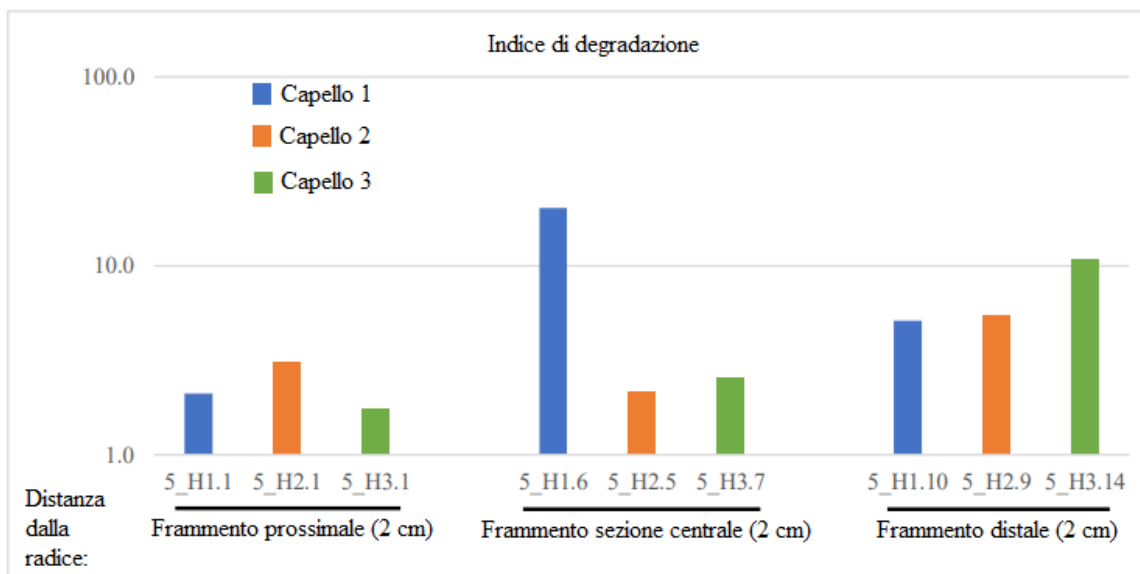


Grafico 3. Variazione dell'indice di degradazione dei campioni di capelli dell'individuo 5

L'individuo 7 era l'unico volontario maschio nello studio, i suoi capelli erano corti, quindi, i segmenti raccolti erano tutti adiacenti l'uno all'altro. In questo caso, i frammenti più vicini alla radice hanno valori DI molto simili (grafico 4).

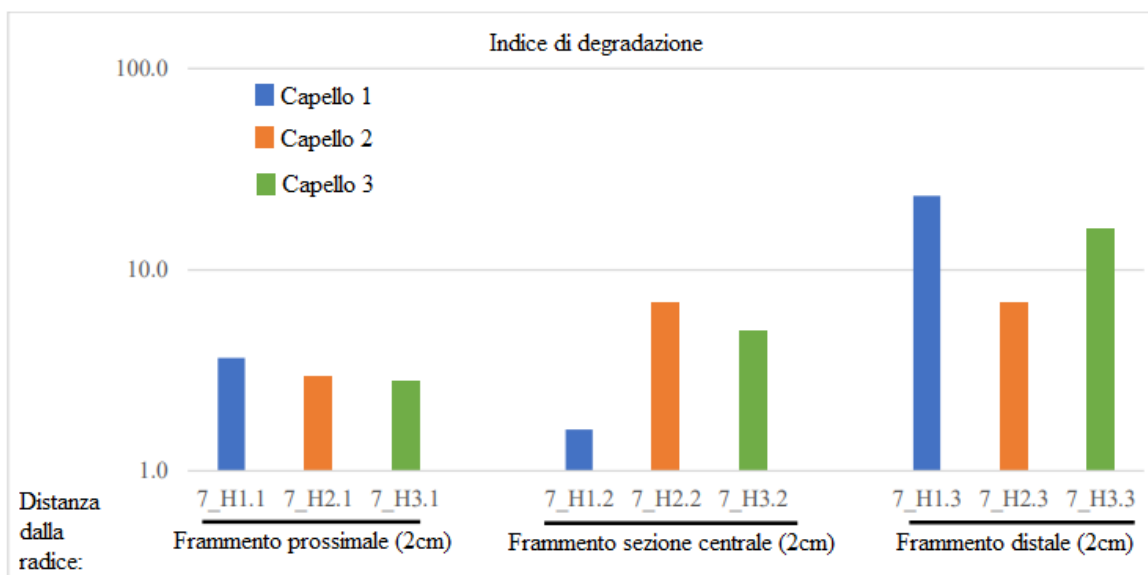


Grafico 4. Variazione dell'indice di degradazione dei campioni di capelli dell'individuo 7

Nel caso dell'individuo 8, sia H1 che H2 mostrano un graduale aumento di DI lungo il loro lunghezza, mentre H3 ha il valore DI più basso nel segmento medio (grafico 5).

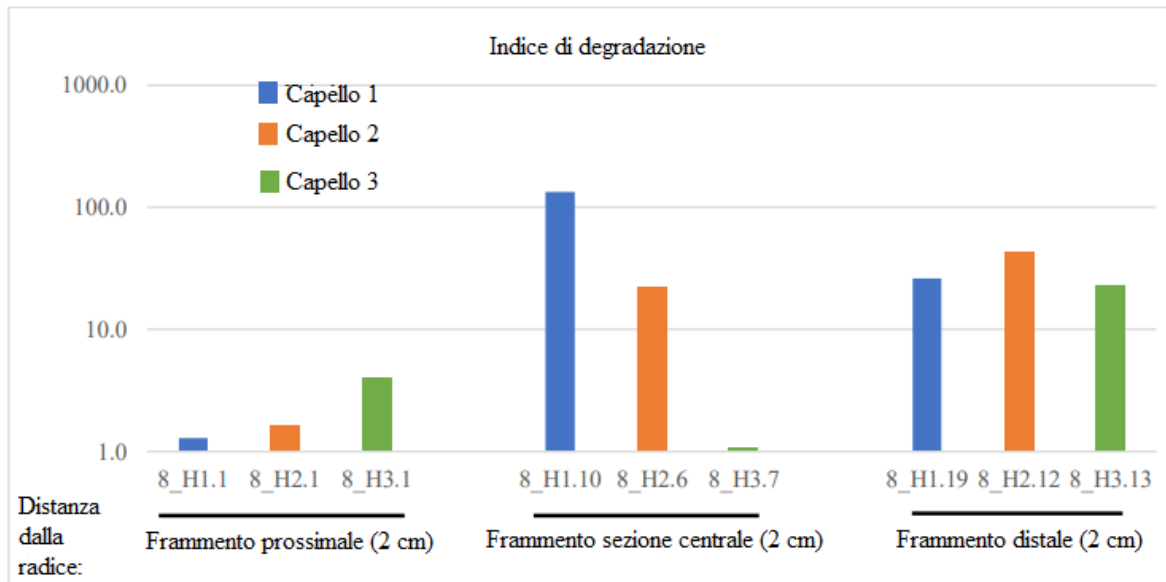


Grafico 5. Variazione dell'indice di degradazione dei campioni di capelli dell'individuo 8

In base ai risultati ottenuti della quantizzazione del DNA mitocondriale sono stati selezionati, per il sequenziamento, i campioni appartenenti al donatore 4, donatore 5 e donatore 7. Per tali campioni si è proceduto alla quantizzazione del nDNA utilizzando lo strumento Qubit 2.0 Fluorometer. Per il sangue è stata rilevata una quantità maggiore di DNA nucleare per tutti i donatori (4_B: 22 ng/ul, 5_B:24,2 ng/ul, 7_B:14 ng/ul) rispetto alla quantità rilevata per i tamponi buccali (4_S: 3,68 ng/ul, 5_S:2,1 ng/ul, 7_S:3,88 ng/ul).

Al termine della quantificazione del DNA mitocondriale e DNA nucleare, si è proceduto all'allestimento delle librerie per MPS utilizzando il kit Precision ID

library kit (Thermo Fisher Scientific) e Precision ID mtDNA Whole Genome Panel, un pannello sviluppato specificamente per applicazioni forensi che permette l'analisi dell'intera sequenza del DNA mitocondriale tramite l'amplificazione di due pool di primer. Ogni pool contiene 81 coppie di primer, con una minima sovrapposizione di primer tra i pool così da coprire l'intero genoma. I 162 ampliconi permettono di ottenere dati anche da campioni altamente compromessi e degradati come fusti di capelli, denti e ossa. Il processo di preparazione di una libreria MPS è suddiviso in diverse fasi, dopo l'amplificazione dell'estratto mediante PCR, ad ogni singolo campione viene legato un Barcode e un adattatore specifico che permetterà di identificare il campione dopo il sequenziamento.

3.2 Risultati della quantificazione delle librerie

Dopo l'amplificazione, i campioni sono stati quantificati utilizzando lo Ion Library TaqMan Quantitation Kit, ciascuno è stato analizzato in duplicato in reazioni da 20 μ . In conformità con le linee guida MIQE per la valutazione dei saggi qPCR, i dati in tempo reale sono stati analizzati con il software Rotor Gene Q v. 2.1.0 di Qiagen (Fig. 16).

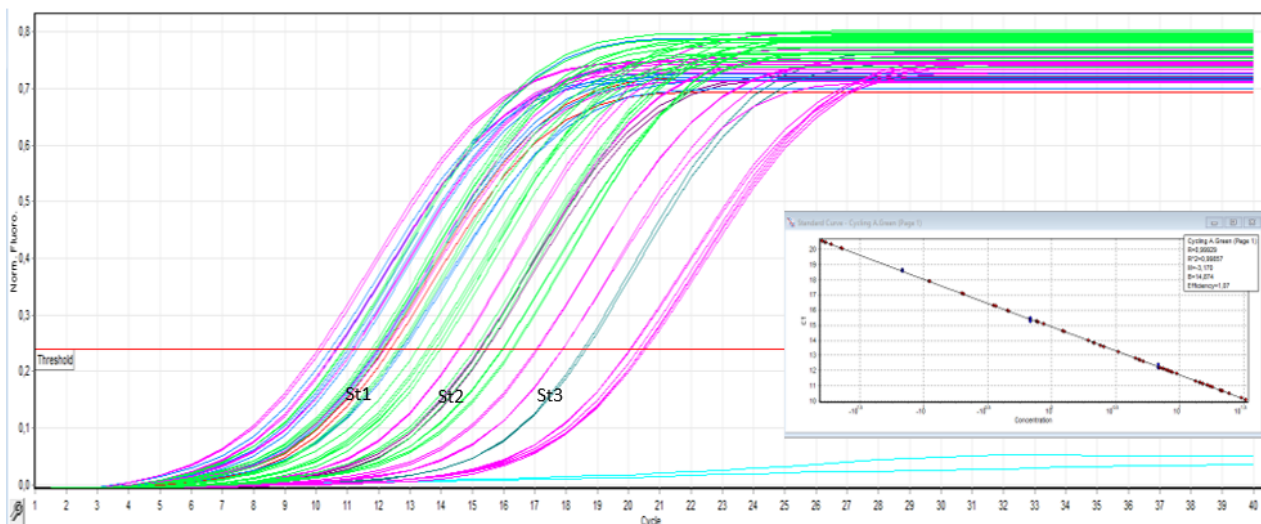


Fig 16. Curva qPCR per la quantificazione delle librerie MPS utilizzando il TaqMan Quantitation Kit. Il numero dei cicli è rappresentato sull'asse x mentre l'intensità della fluorescenza è visualizzata sull'asse y

Al termine della reazione è stata inoltre calcolata una media dei valori della quantificazione dei duplicati, inoltre, poiché i campioni sono stati diluiti 1:100 prima della reazione di quantificazione, la media risultante è stata moltiplicata per 100. I risultati sono visualizzati nella Tabella 16. Alcuni dei campioni sono risultati al di sotto della soglia prestabilita di 30 pM, dunque, sono stati scartati e abbiamo proseguito l'analisi con i campioni idonei.

Campioni	qPCR 1	qPCR 2 (pM)	Media	Media*100
4_H1.1	1,22	1,26	1,24	124,33
4_H1.6	0,11	0,11	0,11	11,11
4_H1.12	0,20	0,20	0,20	20,20
4_H2.1	0,02	0,02	0,02	1,65
4_H2.6	0,02	0,02	0,02	1,65
4_H2.11	0,02	0,02	0,02	2,30
4_H3.1	30,40	32,88	31,64	3164,02
4_H3.5	7,06	8,07	7,57	756,74
4_H3.9	0,79	0,78	0,78	78,27
4_S	20,82	21,58	21,20	2119,95
4_B	14,39	14,49	14,44	1443,93
5_H1.1	29,6	30,99	30,295	3029,50
5_H1.6	1,48	1,36	1,42	142,00
5_H1.10	3,2	3,2	3,2	320,00
5_H2.1	0,00	13,41	6,70	670,27
5_H2.5	21,10	24,35	22,73	2272,91
5_H2.9	4,53	4,82	4,68	467,58
5_H3.1	28,1	28,7	28,4	2840,00
5_H3.7	8,25	7,80	8,03	802,76
5_H3.14	1,8	1,99	1,895	189,50
5_S	27,59	29,56	28,575	2857,50
5_B	15,16	16,42	16,42	1641,70
7_H1.1	4,92	5,23	5,07	507,44
7_H1.2	2,41	2,57	2,49	248,72
7_H1.3	0,45	0,46	0,46	45,73
7_H2.1	8,80	8,59	8,69	869,44
7_H2.2	0,76	0,77	0,77	76,51
7_H2.3	0,35	0,37	0,36	35,96
7_H3.1	1,94	3,31	2,63	262,56
7_H3.2	2,16	2,15	2,15	215,12
7_H3.3	0,87	0,68	0,77	77,25
7_S	17,94	17,40	17,67	1767,09

7_B	7,42	9,49	8,45	845,48
-----	------	------	------	--------

Tabella 16, Risultati della quantificazione delle librerie con qPCR,

Seguendo le istruzioni del produttore (Precision ID mtDNA Panels, Application Guide, TFS) ogni libreria è stata diluita a 40 pM, aggiungendo Low TE.

3.3 Risultati analisi delle prestazioni del sequenziamento

Le metriche recuperate alla fine della corsa aiutano a valutare le prestazioni del sequenziamento e la qualità dei dati prima dell'effettiva analisi dei dati del mitogenoma. Queste metriche di corsa sono visualizzate nella Data tab del TSS e sono illustrati in Fig. 17-18.

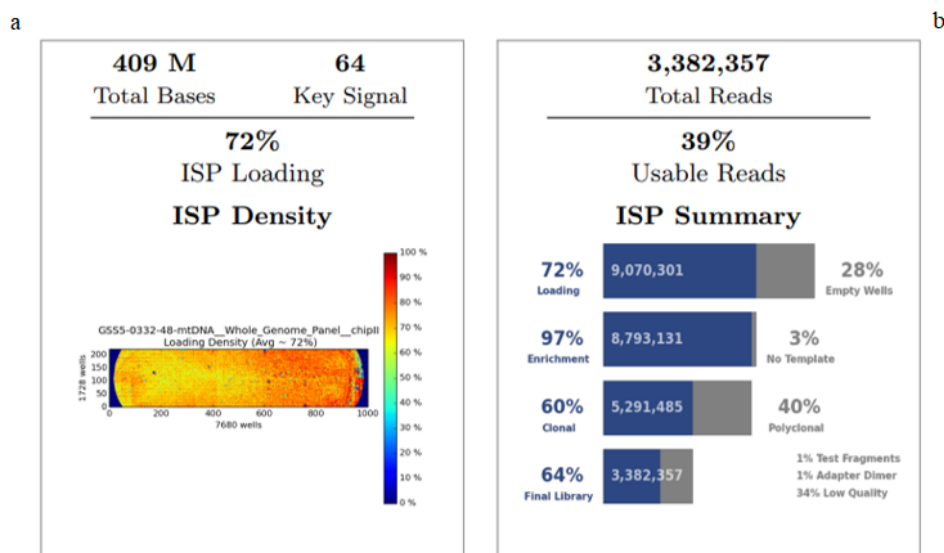


Fig. 17. Riepilogo della corsa per i campioni sequenziati

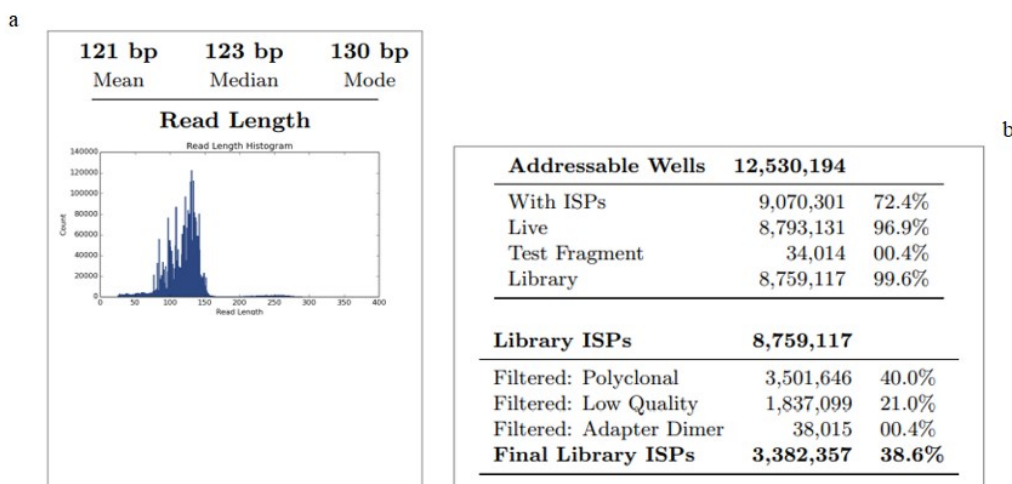


Fig. 18. Riepilogo della corsa per i campioni sequenziati

Il Chip Ion 520™ mostrato nelle figure sopra riportate è stato utilizzato per sequenziare 19 frammenti di capelli e i campioni di riferimento, ossia 2 tamponi buccali e 3 campioni di sangue. Nella figura 17 (a) vediamo rappresentato il carico effettivo delle Ion Sphere Particles (ISPs) sul chip con un valore del 72%, le aree raffigurate in rosso sono quelle con un caricamento ottimale, le aree con densità di caricamento inferiori variano dal giallo al verde, mentre le aree blu indicano un caricamento non riuscito. In questo caso la maggior parte del chip presenta aree gialle in cui la densità di caricamento non è stata del tutto ottimale. Il numero di letture totali è visualizzato in Fig. 17 (b), ed è 3,382,357, di queste il 39%, secondo lo strumento Ion S5, sono letture utilizzabili poiché superano delle verifiche, tra cui, l'avvenuto caricamento, l'arricchimento e fanno parte delle ISP clonali.

L'istogramma in Fig. 18 (a) mostra la lunghezza di tutte le letture presenti nei file di output, sopra troviamo la Media, Mediana e Moda della lunghezza delle letture riportate. La Fig. 18 (b) invece, riporta la classificazione e le tabelle riassuntive delle ISP delle librerie, Lo strumento è stato impostato per l'allineamento con l'rCRS,

quindi, sono state analizzate e illustrate anche le metriche post-allineamento (Fig. 19). Le basi totali incluse nell'allineamento sono visualizzate in Fig. 19 (a), dove vengono riportate sia quelle allineate che quelle non allineate. Il grafico di allineamento riporta in blu le basi allineate, in questo caso 96% e in viola quelle non allineate, 4%. Nella Fig. 19 (b) viene rappresentata l'accuratezza media calcolata per ciascuna posizione in ogni lettura. Infine, i parametri di qualità dell'allineamento vengono visualizzati nella Fig. 19 (c).

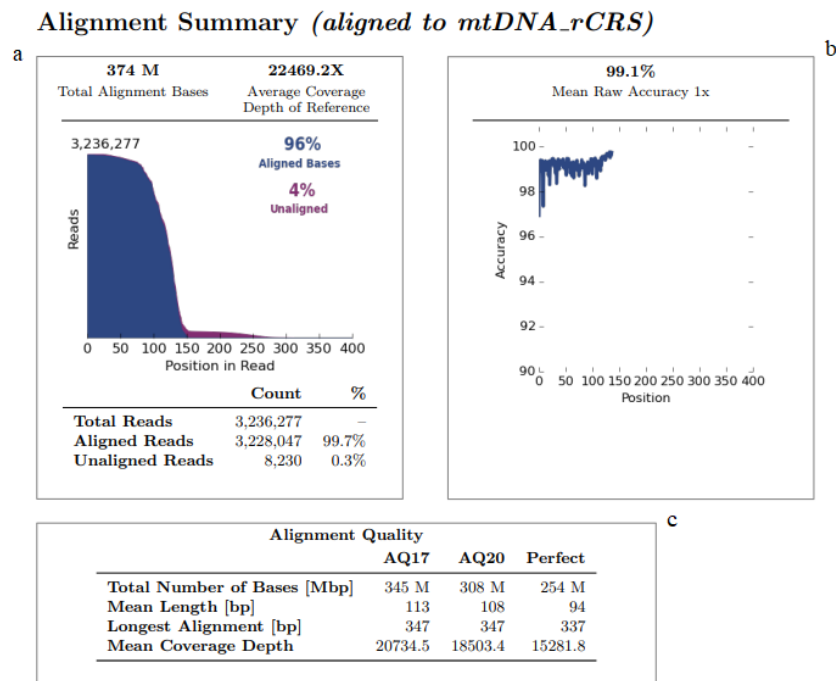


Fig. 19. Metriche post-allineamento incluse nel report di esecuzione

Tutti i valori presenti nelle metriche di corsa raggiungono e superano le soglie fissate dalle istruzioni dello studio, quindi, abbiamo considerato l'esecuzione del sequenziamento riuscita e abbiamo proceduto con l'analisi del coverage.

Le metriche di corsa del secondo chip sono rappresentate nella Fig. 20,21.

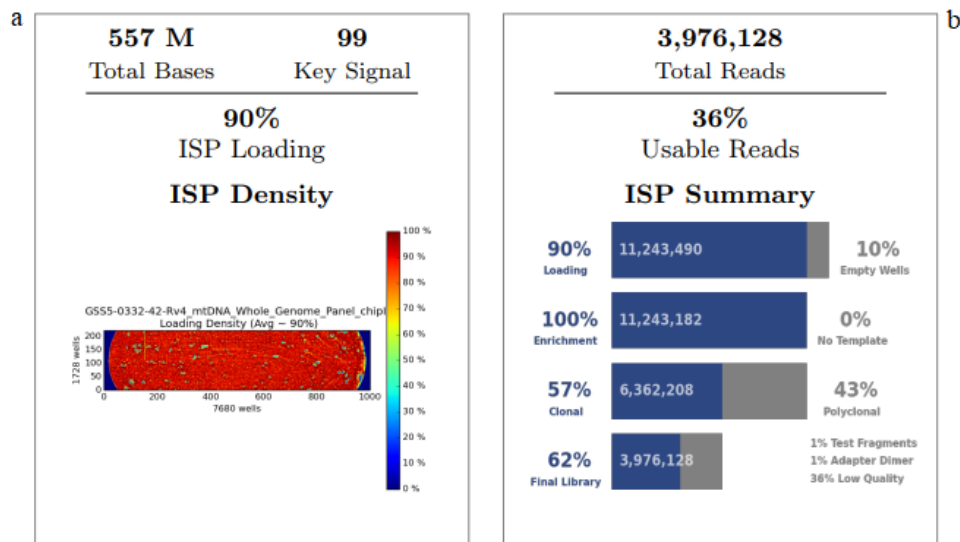


Fig. 20. Riepilogo della corsa per i campioni sequenziati

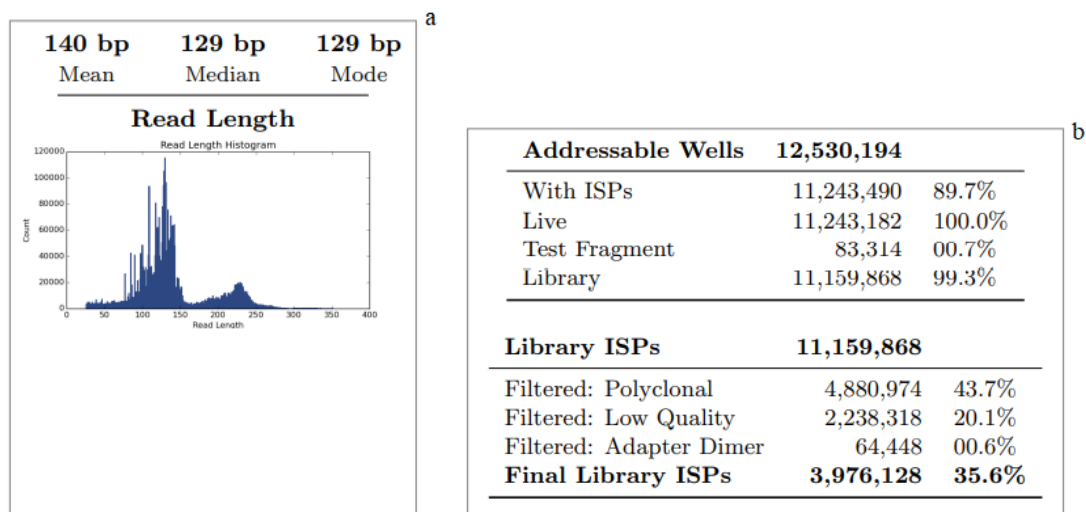


Fig. 21 Riepilogo della corsa per i campioni sequenziati

Nella figura 20 (a) vediamo il carico effettivo delle Ion Sphere Particles (ISPs) sul chip con un valore del 90%. In questo caso la maggior parte del chip presenta aree rosse in cui la densità di caricamento risulta ottimale. Il numero di letture totali è visualizzato in Fig. 20 (b), ed è di 3.976,128, di queste il 36%, secondo lo strumento, sono letture utilizzabili. Nella Fig. 21 vediamo rappresentate le metriche post-allineamento. In Fig. 21 (a) troviamo, come prima, la Media, Mediana e Moda della lunghezza delle letture riportate. La Fig. 21 (b) invece, riporta la classificazione e le

tabelle riassuntive delle ISP delle librerie. Le basi totali incluse nell'allineamento sono visualizzate in Fig. 22 (a), e anche qui vengono riportate sia quelle allineate che quelle non allineate. Il grafico di allineamento riporta in blu le basi allineate, in questo caso 96% e in viola quelle non allineate, 4%. Nella Fig. 22 (b) viene rappresentata l'accuratezza media calcolata per ciascuna posizione in ogni lettura. Nella Fig. 19 (c) abbiamo i parametri di qualità dell'allineamento.

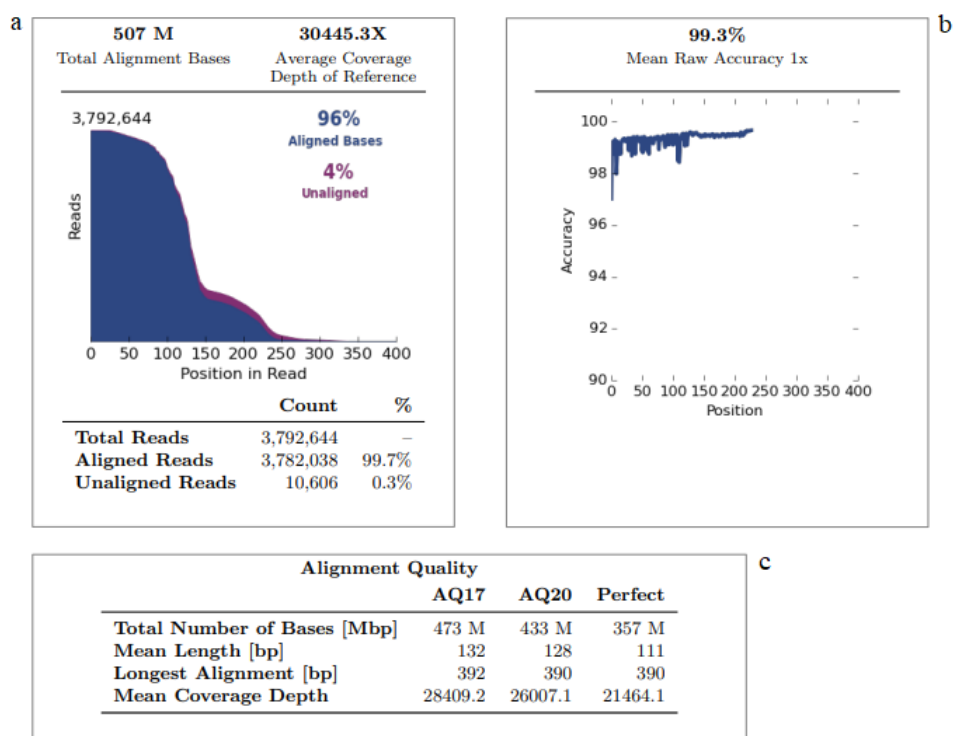


Fig. 22 Metriche post-allineamento incluse nel report di esecuzione

3.4 Analisi del coverage e performance del pannello

Il plug-in del coverage analysis genera un Coverage Analysis Report, il report riassume i codici a barre utilizzati, i campioni, il numero di letture mappate sull'rCRS, la percentuale di letture sul target, la profondità media del coverage e l'uniformità del coverage per base (Tabella 17).

Barcode ID	Sample Name	Mapped Reads	On Target	Mean Depth	Uniformity
IonXpress_001	5_H1.1	305493	100,00%	2149	88,44%
IonXpress_002	5_H1.6	208055	100,00%	1449	88,04%
IonXpress_003	5_H1.10	244738	100,00%	1705	89,09%
IonXpress_004	5_H2.1	140861	100,00%	990,3	61,01%
IonXpress_005	5_H2.5	141100	100,00%	1008	96,06%
IonXpress_006	5_H2.9	150532	100,00%	1048	97,66%
IonXpress_007	5_H3.1	333413	100,00%	2377	90,38%
IonXpress_008	5_H3.7	146061	100,00%	1035	96,48%
IonXpress_009	5_H3.14	328867	100,00%	2297	88,92%
IonXpress_010	5_S	373121	100,00%	2691	88,11%
IonXpress_012	4_H1.1	138355	100,00%	973,1	93,87%
IonXpress_013	4_H1.6	39261	100,00%	273,2	96,23%
IonXpress_014	4_H1.12	67548	100,00%	466,5	95,98%
IonXpress_018	4_H3.1	101748	100,00%	689,3	73,17%
IonXpress_019	4_H3.5	128413	100,00%	901,3	98,34%
IonXpress_020	4_H3.9	129730	100,00%	745,6	52,30%
IonXpress_021	7_H1.1	142355	100,00%	1002	96,64%
IonXpress_022	7_H1.2	155460	100,00%	1002	60,63%
IonXpress_023	7_H1.3	160855	100,00%	1112	64,25%
IonXpress_024	7_H2.1	149420	100,00%	1064	98,19%
IonXpress_025	7_H2.2	145008	100,00%	1017	60,76%
IonXpress_026	7_H2.3	96583	100,00%	564,8	56,87%
IonXpress_027	7_H3.1	172000	100,00%	1188	75,86%
IonXpress_028	7_H3.2	126764	100,00%	904,1	61,86%
IonXpress_029	7_H3.3	131545	100,00%	926,6	96,57%
IonXpress_030	4_S	180060	100,00%	1296	61,33%
IonXpress_031	7_S	143781	100,00%	1038	60,47%
IonXpress_032	5_B	141233	100,00%	1019	98,24%
IonXpress_033	4_B	156301	100,00%	1126	98,97%
IonXpress_034	7_B	143073	100,00%	1037	99,98%

Tabella 17, Risultati Coverage plugin

Per valutare ulteriormente le prestazioni del Precision ID mtDNA Whole Genome Panel è stata calcolata il Relative Depth of Coverage (rDoC) relativo a ciascuno dei 162 ampliconi (Grafico 6).

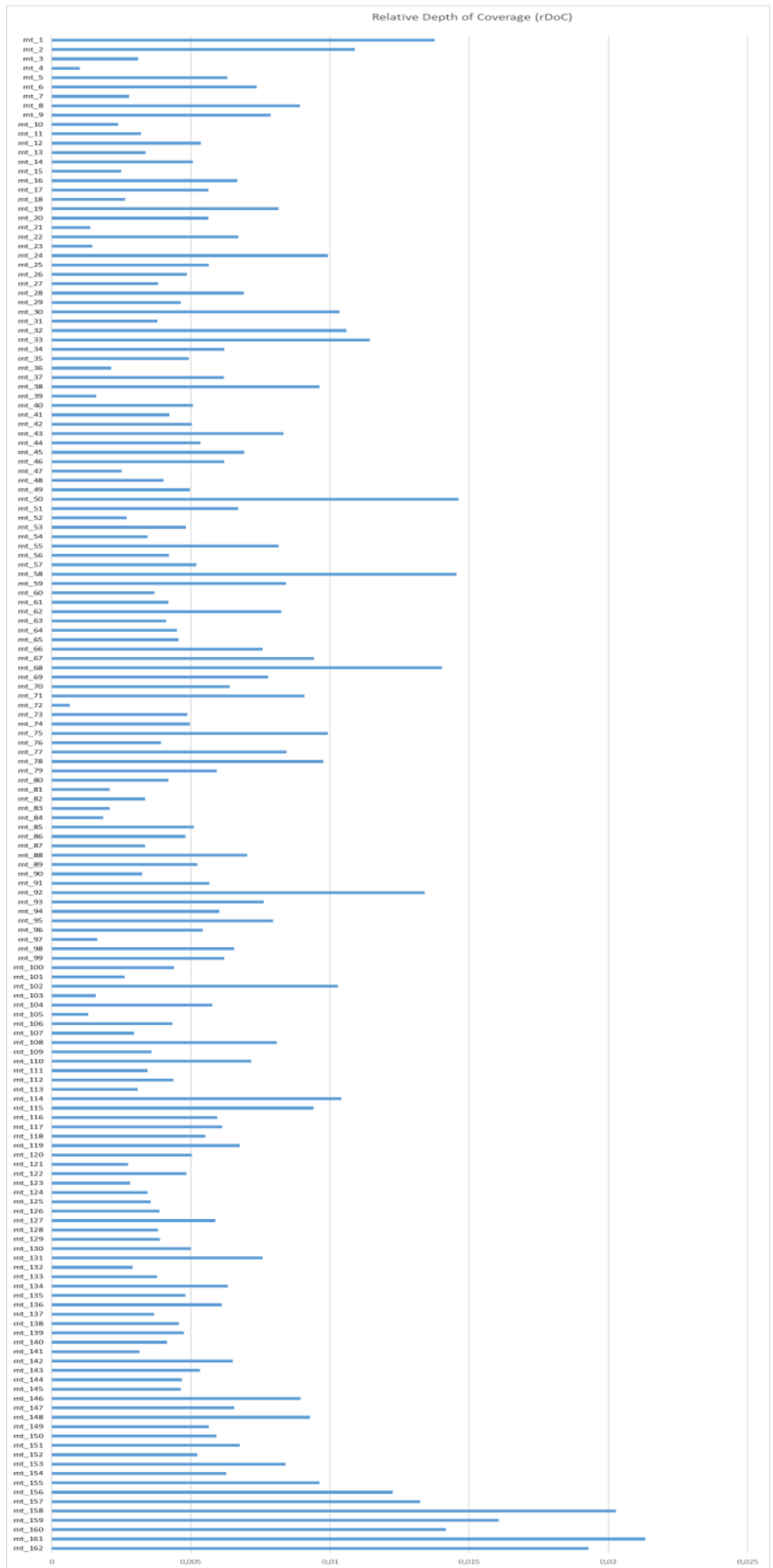


Grafico 6. Relative Depth of Coverage per tutti gli ampliconi inclusi nel pannello. Il codice dell'amplicone si trova sull'asse x, mentre il valore medio è rappresentato sull'asse y.

Questo è un importante criterio di qualità che ci permette di definire le prestazioni di ciascun amplicone. Infatti, alcuni possono essere considerati under performing, tra questi abbiamo: mt_4, mt_21, mt_23, mt_39, mt_72, mt_97, mt_84 e il mt_103. Sfortunatamente per alcuni dei campioni sequenziati è stato osservato uno sbilanciamento nella amplificazione dei due pool di primer (Grafico 7).

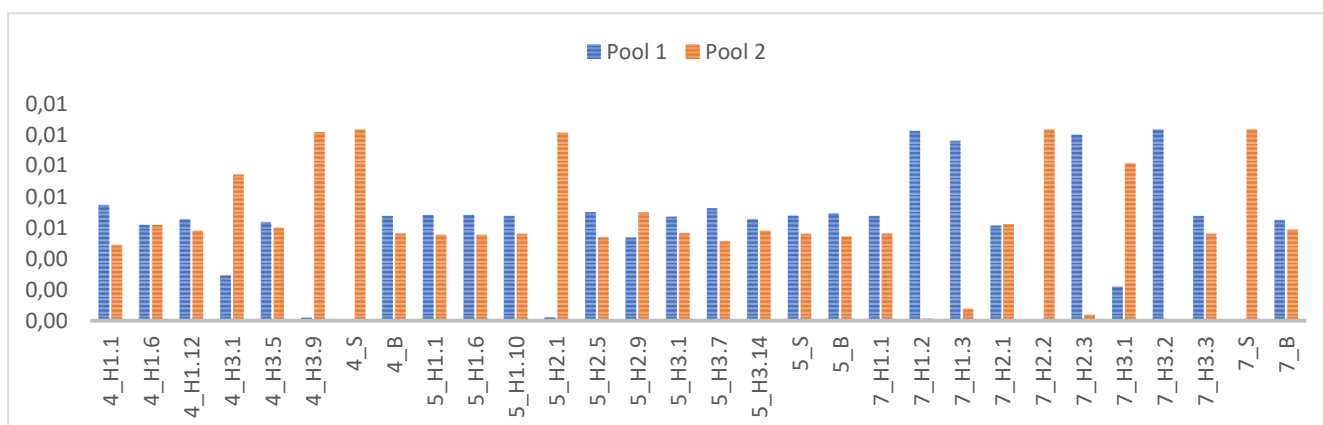


Grafico 7. Rappresentazione dello sbilanciamento osservato tra il pool 1 e il pool 2 per ogni campione analizzato

Possiamo dire che alcuni pool risultano abbastanza sbilanciati per diversi campioni, ad esempio nel 4_H1.1 e 4_H3.1, mentre risultano nettamente sbilanciati per i campioni 4_H3.9 e 7_H3.1. Inoltre, il pool 1 è quasi completamente assente nei campioni 4_H3.9, 5_H2.1, mentre il pool 2 è quasi assente nei campioni 7_H1.2, 7_H1.3 e 7_H2.3. Nei campioni 4_S e 7_S possiamo affermare la totale assenza del pool 1. Il motivo della non omogenea copertura dei pannelli per i diversi campioni non è del tutto chiara, potrebbe essere dovuto alla distribuzione disuguale dei pool durante la preparazione manuale delle librerie, causato dall'elevata viscosità dei reagenti. In quanto la frequenza con cui i pool risultano meno efficienti sembra essere del tutto casuale e non ci permette di definire la relativa inefficienza di questi ultimi.

3.5 Analisi delle varianti eteroplasmatiche rilevate tramite MPS

L'analisi primaria delle varianti eteroplasmatiche è stata eseguita con il software Torrent Suite (TSS, v5.12.3) utilizzando il plug-in Variant Caller (TVC, v5.0.4.0). La soglia di profondità di copertura minima di 20X è stata utilizzata per filtrare le chiamate delle varianti, mentre la soglia di frequenza delle varianti di sequenza per l'eteroplasmia è stata fissata al 10%. L'analisi consisteva nell'applicazione dei parametri "Germline-Low Stringency" e nell'allineamento delle letture all'rCRS (riferimento NCBI NC_012920), con il nucleotide 1-80 ripetuto dopo la posizione 16.569 per accomodare l'allineamento circolare. Al termine dell'analisi del plug-in, è disponibile il report riepilogativo di ciascun campione insieme ai file BAM (binary Allineamento Index) e BAI (binary Allineamento Index) per le letture mappate ed elaborate TVC (TSS UG). L'analisi secondaria è stata eseguita in modo indipendente da parte di due analisti, utilizzando l'Integrative Genomics Viewer (IGV, 2.12.3). Le varianti che hanno superato il filtraggio TVC e sono state confermate dall'analisi IGV sono riassunte nelle tabelle 18, 19 e 20.

Per l'analisi dell'aplogruppo sono stati utilizzati gli aplotipi rilevati nel campione di sangue dei tre donatori selezionati per il sequenziamento. Il donatore 4 presenta un aplotipo costituito da 31 varianti, di cui 8 sono presenti nella regione di controllo e 23 nella regione codificante. L'aplotipo rilevato è stato inserito all'interno del software EMPOP mtDNA Database v4/R13, impostando l'intervallo di query per includere l'intero genoma mitocondriale che ha identificato l'aplogruppo J2b1 (Fig.23)

Haplogrouping							
Rank 1. MRCA: J2b1 4.14							
Source	Haplogroup	Range	Diagnostic Mutations	Cost	Transcript	Missing Mutations	Extra Mutations
PT17	J2b1	ALL	16089T 16126C 16193T 16519Y 73G 150Y 152C 263G 295T 3151C 489C 750G 1439G 2706G 4236C 4769G 5633T 7028T 7476T 8860G 10172A 10398G 11251G 11719A 12612G 13708A 14766T 15257A 15326G 15452A 15812A	4.14	C16168T(0.948) T16193C(0.917) [Y16519T(0.000)] [Y150T(0.000)] [C152T(0.380) 3151delC(0.400) [G5460R(0.000)] [T14302Y(0.000)] A15452C(1.493)]	16193T 152C 3151C 15452A	16168T
Rank 2. MRCA: J2b1a 4.83-5.27							
Source	Haplogroup	Range	Diagnostic Mutations	Cost	Transcript	Missing Mutations	Extra Mutations
PT17	J2b1a	ALL	16089T 16126C 16193T 16278T 16519Y 73G 150Y 152C 263G 295T 3151C 489C 750G 1438G 2706G 4216C 4769G 5633T 7028T 7476T 8860G 10172A 10398G 11251G 11719A 12612G 13708A 14766T 15257A 15326G 15452A 15812A	4.83	C16168T(0.948) T16193C(0.917) T16278C(0.689) [Y16519T(0.000)] [Y150T(0.000)] C152T(0.380) 3151delC(0.400) [G5460R(0.000)] [T14302Y(0.000)] A15452C(1.493)]	16193T 16278T 152C 3151C 15452A	16168T
PT17	J2b1a-16311	ALL	16089T 16126C 16193T 16278T 16311C 16519Y 73G 150Y 152C 263G 295T 3151C 489C 750G 1438G 2706G 4216C 4769G 5633T 7028T 7476T 8860G 10172A 10398G 11251G 11719A 12612G 13708A 14766T 15257A 15326G 15452A 15812A	5.27	C16168T(0.948) T16193C(0.917) T16278C(0.689) C16311T(0.447) [Y16519T(0.000)] [Y150T(0.000)] C152T(0.380) 3151delC(0.400) [G5460R(0.000)] [T14302Y(0.000)] A15452C(1.493)]	16193T 16278T 16311C 152C 3151C 15452A	16168T

Figura 23. Aplotipo identificato dal software Empop per il donatore 4.

Il donatore 5 presenta un aplotipo costituito da 34 varianti, di cui 11 presenti nella regione di controllo e 23 nella regione codificante. La sottomissione dell'aplotipo al software EMPOP mtDNA Database v4/R13 ha identificato l'aplotipo U1b3 come mostrato nella figura riportata qui di seguito (Fig.24).

Haplogrouping							
Rank 1. MRCA: U1b3 3.00							
Source	Haplogroup	Range	Diagnostic Mutations	Cost	Transcript	Missing Mutations	Extra Mutations
PT17	U1b3	ALL	16111Y 16249C 16311C 16327T 16519Y 73G 146C 195C 263G 285T 3151C 750G 1438G 2387C 2706G 4769G 7028T 8395T 8860G 10895C 11467G 11566G 11719A 12386G 12372A 12879C 13104G 14070G 14766T 15148A 15172A 15326G 15954C	3.00	[Y16111C(0.000)] T16172C(0.629) [C16267Y(0.000)] [Y16519T(0.000)] T152C(0.380) 3151delC(0.400) [G12308R(0.000)] A15763G(1.593)]	3151C	16172C 152C 15763G
Rank 2. MRCA: U1b 3.93							
Source	Haplogroup	Range	Diagnostic Mutations	Cost	Transcript	Missing Mutations	Extra Mutations
PT17	U1b	ALL	16111Y 16249C 16327T 16519Y 73G 146C 263G 285T 3151C 750G 1438G 2387C 2706G 4769G 7028T 8395T 8860G 10895C 11467G 11566G 11719A 12386G 12372A 12879C 13104G 14070G 14766T 15148A 15172A 15326G 15954C	3.93	[Y16111C(0.000)] T16172C(0.629) [C16267Y(0.000)] T16311T(0.447) [Y16519T(0.000)] T152C(0.380) T195C(0.484) 3151delC(0.400) [G12308R(0.000)] A15763G(1.593)]	3151C	16172C 16311C 152C 195C 15763G

Figura 24. Aplotipo identificato dal software Empop per il donatore 5.

Infine, il donatore 7 presenta un aplotipo costituito da 31 varianti, di cui 6 sono presenti nella regione di controllo e 25 nella regione codificante. Dal software EMPOP è stato identificato appartenete all'aplotipo T2b (Fig.25).

Haplogrouping							
Rank 1. MRCA: T2b 1.16-1.54							
Source	Haplogroup	Range	Diagnostic Mutations	Cost	Transcript	Missing Mutations	Extra Mutations
PT17	T2b	ALL	16126C 16294T 16296Y 16304C 16519Y 73G 263G 3151C 709A 750G 930A 1438G 1888A 2706G 4236C 4769G 4970G 5147A 7028T 8699A 8860G 10463C 11251G 11719A 11813G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A	116	[Y16296C(0.000) C16304T(0.761) [Y16519T(0.000) 3151delC(0.400)]	16304C 3151C	(none)
PT17	T2b-152	ALL	16126C 16294T 16296Y 16304C 16519Y 73G 263G 3151C 709A 750G 930A 1438G 1888A 2706G 4236C 4769G 4970G 5147A 7028T 8699A 8860G 10463C 11251G 11719A 11813G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A	154	[Y16296C(0.000) C16304T(0.761) [Y16519T(0.000) C152T(0.380) 3151delC(0.400)]	16304C 152C 3151C	(none)
Rank 2. MRCA: T2b 1.71-1.99							
Source	Haplogroup	Range	Diagnostic Mutations	Cost	Transcript	Missing Mutations	Extra Mutations
PT17	T2b-16362	ALL	16126C 16294T 16296Y 16304C 16362C 16519Y 73G 263G 3151C 709A 750G 930A 1438G 1888A 2706G 4236C 4769G 4970G 5147A 7028T 8699A 8860G 10463C 11251G 11719A 11813G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A	1.71	[Y16296C(0.000) C16304T(0.761) C16362T(0.550) [Y16519T(0.000) 3151delC(0.400)]	16304C 16362C 3151C	(none)
PT17	T2b-150	ALL	16126C 16294T 16296Y 16304C 16519Y 73G 150T 263G 3151C 709A 750G 930A 1438G 1888A 2706G 4236C 4769G 4970G 5147A 7028T 8699A 8860G 10463C 11251G 11719A 11813G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A	1.76	[Y16296C(0.000) C16304T(0.761) [Y16519T(0.000) T1500(0.596) 3151delC(0.400)]	16304C 150T 3151C	(none)
PT17	T2b-2b	ALL	16126C 16294T 16296Y 16519Y 73G 152Y 263G 3151C 709A 750G 930A 1438G 1888A 2706G 4236C 4769G 4970G 5147A 7028T 8699A 8860G 10463C 11251G 11719A 11813G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A	1.99	[Y16296C(0.000) [Y16519T(0.000) [Y152T(0.000) 3151delC(0.400) G12422T(1.593)]	3151C 11242G	(none)

Figura 25. Aplogruppo identificato dal software Empop per il donatore 7.

Nella tabella 18 sono riportati gli aplotipi delle diverse matrici testate nello studio relative al donatore 4.

4_S	4_B	4_H1.1	4_H1.6	4_H1.12	4_H3.1	4_H3.5	4_H3.9
A73G	A73G	A73G	A73G	73R	A73G	A73G	A73G
cov<20X	C150T	C150T	C150T	C150T	C150T	C150T	cov<20X
cov<20X	rCRS	rCRS	rCRS	185R	rCRS	rCRS	cov<20X
A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G
C295T	C295T	C295T	295Y	295Y	C295T	C295T	295Y
cov<20X	T489C	T489C	T489C	489Y	T489C	T489C	cov<20X
A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G
cov<20X	rCRS	rCRS	1043Y	rCRS	rCRS	rCRS	cov<20X
A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G
cov<20X	rCRS	rCRS	rCRS	2080Y	rCRS	rCRS	cov<20X
rCRS	rCRS	2550R	2550R	2550R	rCRS	rCRS	Rcrs
cov<20X	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	cov<20X
rCRS	rCRS	rCRS	rCRS	2817R	rCRS	rCRS	rCRS
rCRS	rCRS	rCRS	rCRS	3880R	rCRS	rCRS	rCRS
T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C
cov<20X	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	cov<20X
cov<20X	rCRS	rCRS	rCRS	5417R	rCRS	rCRS	cov<20X
5460R	5460R	rCRS	5460R	5460R	G5460A	rCRS	rCRS
cov<20X	C5633T	C5633T	C5633T	5633Y	C5633T	C5633T	cov<20X
cov<20X	rCRS	rCRS	rCRS	6888R	rCRS	rCRS	cov<20X
cov<20X	rCRS	rCRS	rCRS	6958R	rCRS	rCRS	cov<20X
C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T
C7476T	C7476T	C7476T	7476Y	7476Y	cov<20X	C7476T	C7476T
A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G
rCRS	rCRS	rCRS	rCRS	9565R	rCRS	rCRS	rCRS
cov<20X	G10172A	G10172A	10172R	G10172A	G10172A	G10172A	cov<20X
A10398G	A10398G	A10398G	10398R	10398Y	A10398G	A10398G	A10398G
A11251G	A11251G	A11251G	11251R	A11251G	A11251G	A11251G	A11251G
cov<20X	G11719A	G11719A	G11719A	11719R	G11719A	G11719A	cov<20X
A12612G	A12612G	A12612G	12612R	12612R	A12612G	A12612G	A12612G
G13708A	G13708A	G13708A	13708R	G13708A	G13708A	G13708A	G13708A
14302Y	14302Y	rCRS	14302Y	rCRS	T14302C	rCRS	14302Y
cov<20X	C14766T	C14766T	C14766T	14766Y	C14766T	C14766T	cov<20X
G15257A	G15257A	G15257A	15257R	15257R	G15257A	G15257A	G15257A
A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G
C15452A	C15452A	C15452A	15452	15452	C15452A	C15452A	C15452A
T15453C	T15453C	T15453C	15453Y	15453Y	T15453C	T15453C	T15453C
cov<20X	G15812A	G15812A	15812R	15812R	G15812A	G15812A	cov<20X
C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T
T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C
cov<20X	C16168T	C16168T	C16168T	C16168T	C16168T	C16168T	cov<20X

Tabella 18. Aplotipi identificati per il tampone buccale, il sangue e fusti dei capelli del donatore 4.

Per quanto riguarda il donatore 4 possiamo dire in primo luogo che i frammenti

4_H1.6 e il 4_H1.12 presentano un numero eccessivamente alto di eteroplasmie, il

che fa pensare a una possibile contaminazione di questi campioni, per questo motivo l'aplotipo rilevato per questi due frammenti saranno oggetto di rivalutazioni future. In entrambi i campioni di riferimento (4_S e 4_B) sono state rilevate due eteroplasmie, la 5460R (Fig. 26) e la 14302Y. La prima è risultata un'eteroplasmia anche nel 4_H1.6 e nel 4_H1.12, mentre è completamente mutata nel 4_H3.1, nei frammenti 4_H1.1, 4_H3.5, 4_H3.9 risulta come parte della rCRS. La seconda eteroplasmia rilevata, 14302Y, è presente anche nel 4_H1.6 e nel 4_H3.9, mentre è parte della rCRS nel 4_H1.1, 4_H1.12 e 4_H3.5, invece nel 4_H3.1 risulta completamente mutata. Da notare che, nel campione 4_S il pool 1 non è stato adeguatamente amplificato (reads con coverage inferiore al threshold di 20x), di conseguenza possiamo fare un confronto diretto tra i due riferimenti esaminati.

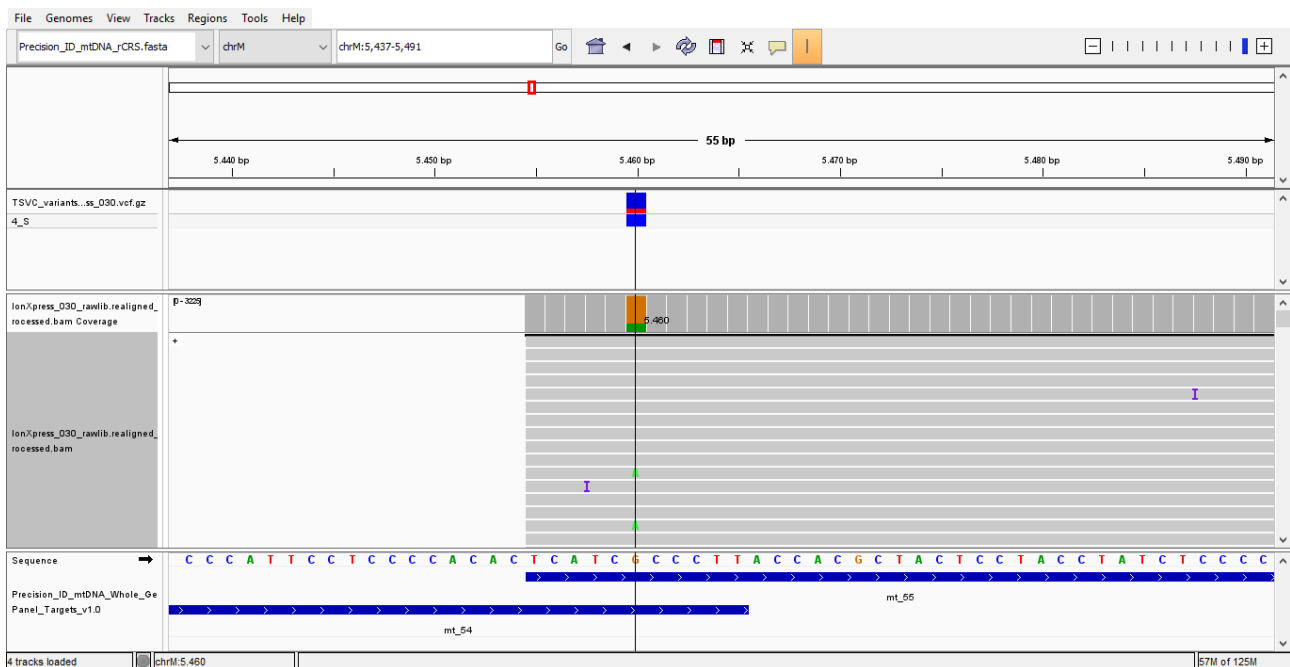


Fig. 26 Eteroplasmia in posizione 5460 nel campione 4_S, confermata tramite analisi con IGV

L'intero capello 4_H1 presenta l'eteroplasmia 2550R (Fig. 27) mentre questa posizione risulta non essere mutata in tutti gli altri campioni. Tuttavia, dato l'elevato

numero di eteroplasmie riscontrate per i restanti frammenti di questo capello non possiamo confrontare ulteriori varianti riscontrate in questi aplotipi anomali.

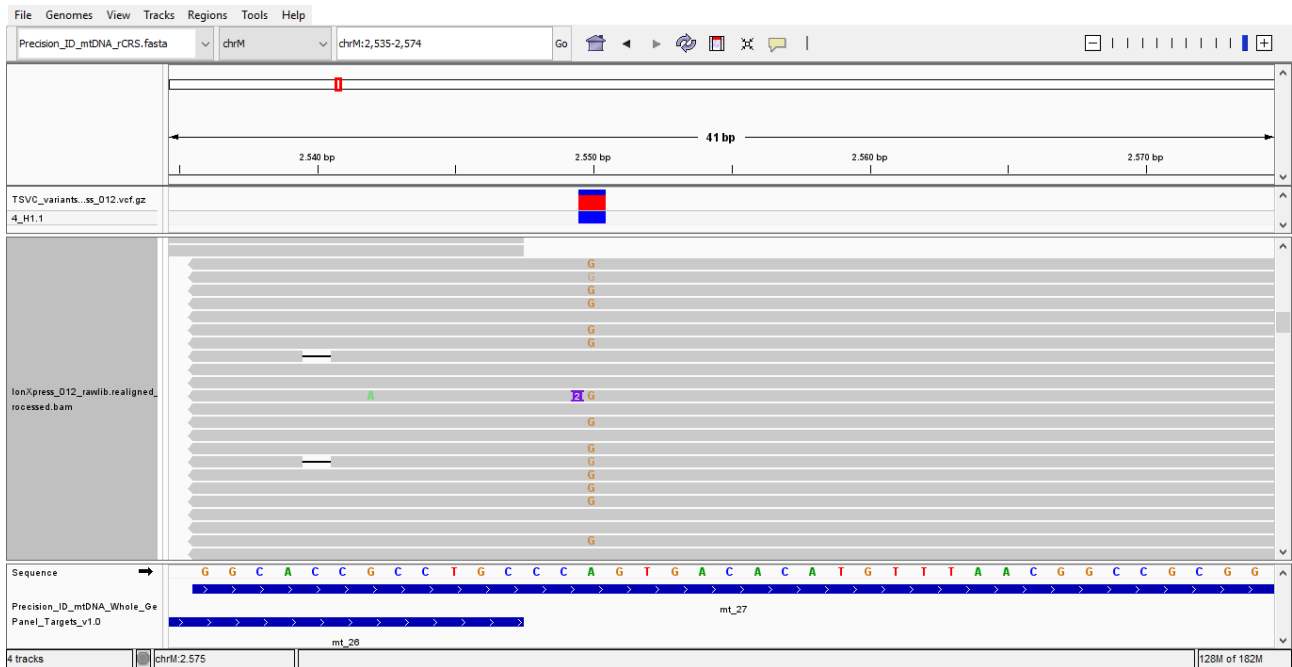


Fig. 27 Eteroplasmia in posizione 2550 nel campione 4_H1.1, confermata tramite analisi con IGV

Infine, il capello 4_H3 mostra aplotipi simili in tutti i tre frammenti ad eccezione della posizione 295, mutata completamente nel primo e secondo frammenti ma eteroplasmica dell'ultimo e della posizione 14302 (tabella 18).

Nella tabella 19 sono riportati gli aplotipi del donatore 5.

5_S	5_B	5_H1.1	5_H1.6	5_H1.10	5_H2.1	5_H2.5	5_H2.9	5_H3.1	5_H3.7	5_H3.14
A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G
T146C	T146C	T146C	T146C	T146C	T146C	T146C	T146C	T146C	T146C	T146C
T152C	T152C	T152C	T152C	T152C	T152C	T152C	T152C	T152C	T152C	T152C
T195C	T195C	T195C	T195C	T195C	T195C	T195C	T195C	T195C	T195C	T195C
A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G
C285T	C285T	C285T	C285T	C285T	C285T	C285T	C285T	C285T	C285T	C285T
A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G
rCRS	rCRS	rCRS	rCRS	rCRS	G1097A	G1097A	G1097A	rCRS	rCRS	rCRS
A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G
T2387C	T2387C	T2387C	T2387C	T2387C	T2387C	T2387C	T2387C	T2387C	T2387C	T2387C
A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G
A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G
C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T
C8395T	C8395T	C8395T	C8395T	C8395T	C8395T	C8395T	C8395T	C8395T	C8395T	C8395T
A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G
rCRS	rCRS	9925R	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS
T10885C	T10885C	T10885C	T10885C	T10885C	A11467G	T10885C	T10885C	T10885C	T10885C	T10885C
A11467G	A11467G	A11467G	A11467G	A11467G	A11467G	A11467G	A11467G	A11467G	A11467G	A11467G
A11566G	A11566G	A11566G	A11566G	A11566G	A11566G	A11566G	A11566G	A11566G	A11566G	A11566G
G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A
12308R	A12308G	A12308G	A12308G	12308R	A12308G	A12308G	A12308G	A12308G	A12308G	A12308G
G12372A	G12372A	G12372A	G12372A	G12372A	G12372A	G12372A	G12372A	G12372A	G12372A	G12372A
T12879C	T12879C	T12879C	T12879C	T12879C	T12879C	T12879C	T12879C	T12879C	T12879C	T12879C
A13104G	A13104G	A13104G	A13104G	A13104G	A13104G	A13104G	A13104G	A13104G	A13104G	A13104G
A14070G	A14070G	A14070G	A14070G	A14070G	A14070G	A14070G	A14070G	A14070G	A14070G	A14070G
C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T
G15148A	G15148A	G15148A	15148R	G15148A	G15148A	G15148A	G15148A	G15148A	G15148A	G15148A
G15172A	G15172A	G15172A	15172R	G15172A	G15172A	G15172A	G15172A	G15172A	G15172A	G15172A
A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G
A15763G	A15763G	A15763G	A15763G	A15763G	A15763G	A15763G	A15763G	A15763G	A15763G	A15763G
A15954C	A15954C	A15954C	A15954C	A15954C	A15954C	A15954C	A15954C	A15954C	A15954C	A15954C
T16172C	T16172C	T16172C	T16172C	T16172C	T16172C	T16172C	T16172C	T16172C	T16172C	T16172C
T16249C	T16249C	T16249C	T16249C	T16249C	T16249C	T16249C	T16249C	T16249C	T16249C	T16249C
16261Y	16261Y	16261Y	16261Y	16261Y	C16261T	C16261T	C16261T	C16261T	rCRS	rCRS
T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C
C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T

Tabella 19. Aplotipi identificati per il tampone buccale, il sangue e fusti dei capelli del donatore 5.

Per il donatore 5 è stata osservata l'eteroplasmia 16261Y (Fig. 28) in entrambi i campioni di riferimento. Questa eteroplasmia successivamente è stata rilevata anche

in tutti i frammenti del primo capello: 5_H1.1, 5_H1.6, 5_H1.10. Viceversa, nei restanti frammenti si presenta come una mutazione omoplasmica, tranne nell' 5_H3.7 e nel 5_H3.14.

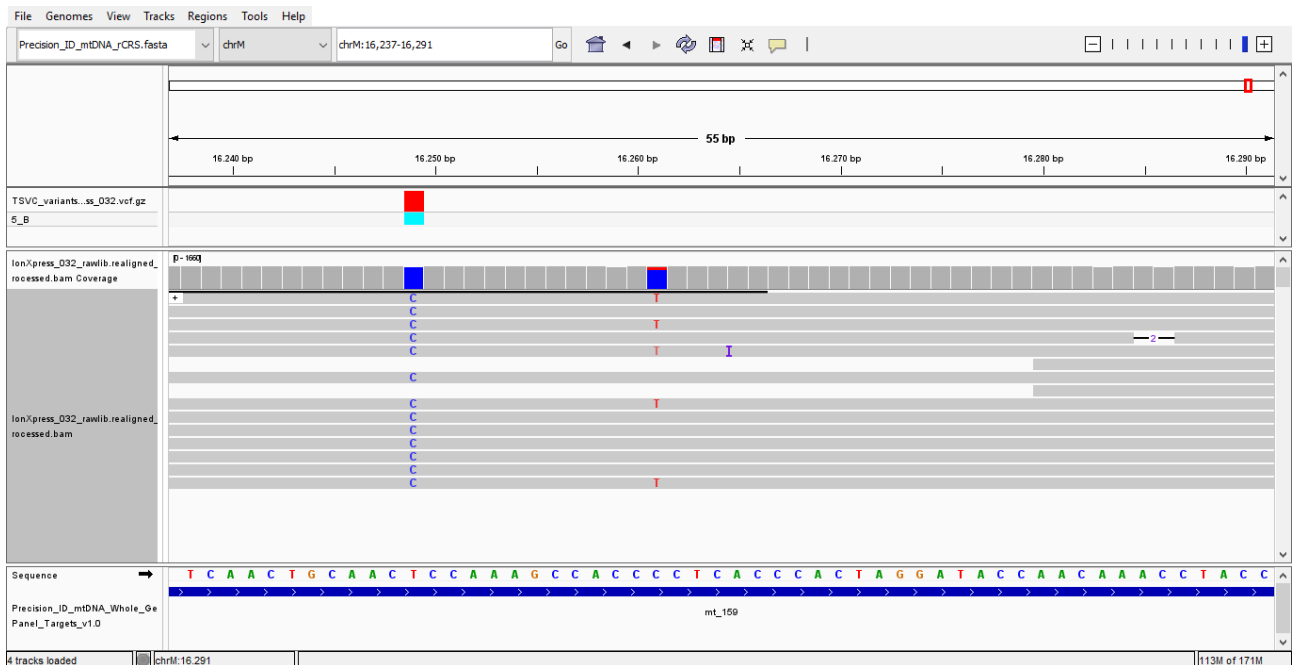


Fig 28. Un eteroplasmia rilevata mediante ispezione IGV in posizione 16261 nel campione 5_B

Da notare che l'eteroplasmia 9925R è presente solo sul frammento 5_H1.1 mentre per tutti gli altri campioni del donatore 5 questa posizione non mostra alcuna mutazione.

Nel 5_H1.6 sono state riscontrate altre due posizioni eteroplasmiche: 15148R e 15172R, che negli altri campioni risultano totalmente mutate.

Nella tabella 20 sono riportati gli aplotipi del donatore 7.

7_S	7_B	7_H1.1	7_H1.2	7_H1.3	7_H2.1	7_H2.2	7_H2.3	7_H3.1	7_H3.2	7_H3.3
A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G
cov<20X	rCRS	215R	215R	215R	rCRS	cov<20X	Rcrs	rCRS	rCRS	rCRS
A263G	A263G	A263G	cov<20X	cov<20X	A263G	A263G	cov<20X	A263G	cov<20X	A263G
G709A	G709A	G709A	G709A	G709A	G709A	G709A	G709A	G709A	G709A	709R (
A750G	A750G	A750G	cov<20X	A750G	A750G	A750G	cov<20X	A750G	cov<20X	A750G
G930A	G930A	G930A	cov<20X	cov<20X	G930A	G930A	cov<20X	G930A	cov<20X	G930A
A1438G	A1438G	A1438G	cov<20X	cov<20X	A1438G	A1438G	cov<20X	A1438G	cov<20X	A1438G
cov<20X	G1888A	G1888A	G1888A	G1888A	G1888A	cov<20X	G1888A	G1888A	G1888A	1888R
cov<20X	A2706G	A2706G	A2706G	A2706G	A2706G	cov<20X	A2706G	A2706G	A2706G	A2706G
rCRS	rCRS	rCRS	rCRS	3633Y	rCRS	rCRS	cov<20X	rCRS	cov<20X	rCRS
T4216C	T4216C	T4216C	T4216C	4216Y	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C
cov<20X	A4769G	A4769G	A4769G	A4769G	A4769G	cov<20X	A4769G	A4769G	A4769G	A4769G
A4917G	A4917G	A4917G	cov<20X	4917R	A4917G	A4917G	cov<20X	A4917G	cov<20X	A4917G
cov<20X	G5147A	G5147A	5147R	5147R	G5147A	cov<20X	G5147A	G5147A	G5147A	G5147A
cov<20X	rCRS	rCRS	6261R	rCRS	rCRS	cov<20X	rCRS	rCRS	rCRS	rCRS
C7028T	C7028T	C7028T	cov<20X	C7028T	C7028T	C7028T	cov<20X	C7028T	cov<20X	C7028T
cov<20X	rCRS	rCRS	rCRS	rCRS	rCRS	cov<20X	rCRS	rCRS	rCRS	7206R
cov<20X	rCRS	rCRS	rCRS	7645Y	rCRS	cov<20X	rCRS	rCRS	rCRS	rCRS
cov<20X	rCRS	rCRS	8455Y	rCRS	rCRS	cov<20X	rCRS	rCRS	rCRS	rCRS
cov<20X	rCRS	rCRS	rCRS	8513Y	rCRS	cov<20X	rCRS	rCRS	rCRS	rCRS
cov<20X	G8697A	G8697A	G8697A	G8697A	G8697A	cov<20X	G8697A	G8697A	G8697A	G8697A
cov<20X	rCRS	rCRS	rCRS	rCRS	rCRS	cov<20X	rCRS	rCRS	rCRS	8765Y
A8860G	A8860G	A8860G	cov<20X	cov<20X	A8860G	A8860G	cov<20X	A8860G	cov<20X	A8860G
cov<20X	rCRS	rCRS	rCRS	9189Y	rCRS	cov<20X	rCRS	rCRS	rCRS	rCRS
T10463C	T10463C	T10463C	cov<20X	cov<20X	T10463C	T10463C	cov<20X	T10463C	cov<20X	T10463C
A11251G	A11251G	A11251G	cov<20X	A11251G	A11251G	A11251G	cov<20X	A11251G	cov<20X	A11251G
cov<20X	G11719A	G11719A	G11719A	G11719A	G11719A	cov<20X	G11719A	G11719A	G11719A	G11719A
A11812G	A11812G	A11812G	cov<20X	11812R	A11812G	11812R	cov<20X	A11812G	cov<20X	A11812G
rCRS	rCRS	rCRS	cov<20X	rCRS	rCRS	rCRS	cov<20X	12413Y	cov<20X	rCRS
cov<20X	G13368A	G13368A	G13368A	13368R	G13368A	cov<20X	G13368A	G13368A	G13368A	G13368A
rCRS	rCRS	rCRS	cov<20X	13956R	rCRS	rCRS	cov<20X	rCRS	cov<20X	cov<20X
rCRS	rCRS	rCRS	cov<20X	13973	rCRS	rCRS	cov<20X	rCRS	cov<20X	rCRS
cov<20X	A14233G	A14233G	A14233G	A14233G	A14233G	cov<20X	A14233G	A14233G	A14233G	A14233G
cov<20X	C14766T	C14766T	C14766T	14766Y	C14766T	cov<20X	C14766T	C14766T	C14766T	C14766T
G14905A	G14905A	G14905A	cov<20X	cov<20X	G14905A	G14905A	cov<20X	G14905A	cov<20X	G14905A
A15326G	A15326G	A15326G	cov<20X	A15326G	A15326G	A15326G	cov<20X	A15326G	cov<20X	A15326G
C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A
cov<20X	A15607G	A15607G	A15607G	A15607G	A15607G	cov<20X	A15607G	A15607G	A15607G	A15607G
G15928A	G15928A	G15928A	cov<20X	15928R	G15928A	G15928A	cov<20X	G15928A	cov<20X	G15928A
T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C
cov<20X	rCRS	rCRS	rCRS	rCRS	rCRS	cov<20X	rCRS	rCRS	16169Y	16169Y
rCRS	rCRS	rCRS	rCRS	16292Y	rCRS	rCRS	cov<20X	rCRS	cov<20X	rCRS
C16294T	C16294T	C16294T	C16294T	C16294T	C16294T	C16294T	cov<20X	C16294T	cov<20X	C16294T
T16304C	T16304C	T16304C	T16304C	T16304C	T16304C	T16304C	cov<20X	T16304C	cov<20X	T16304C

T16519C T16519C T16519C cov<20X **16519Y** T16519C T16519C cov<20X T16519C cov<20X T16519C

Tabella 20. Aplotipi identificati per il tampone buccale, il sangue e fusti dei capelli del donatore 7.

Riguardo al donatore 7 nessuna eteroplasmia è stata ritrovata nei campioni di riferimento. Come per due frammenti del donatore 4, il frammento 7_H1.3 presenta un numero eccessivo di eteroplasmie, quindi, anche per questo campione potremmo ipotizzare una possibile contaminazione. L'eteroplasmia 215R (Fig. 29), è stata ritrovata su tutti i frammenti del primo capello, inoltre nel frammento 7_H1.3 viene rilevata la 14766Y, mentre in tutti gli altri frammenti risulta completamente mutata; ugualmente per la posizione 4216Y. Nell'ultimo frammento del terzo capello la 709R è diventata un'eteroplasmia, mentre negli altri frammenti è sempre mutata, stessa cosa vale per la 1888R.

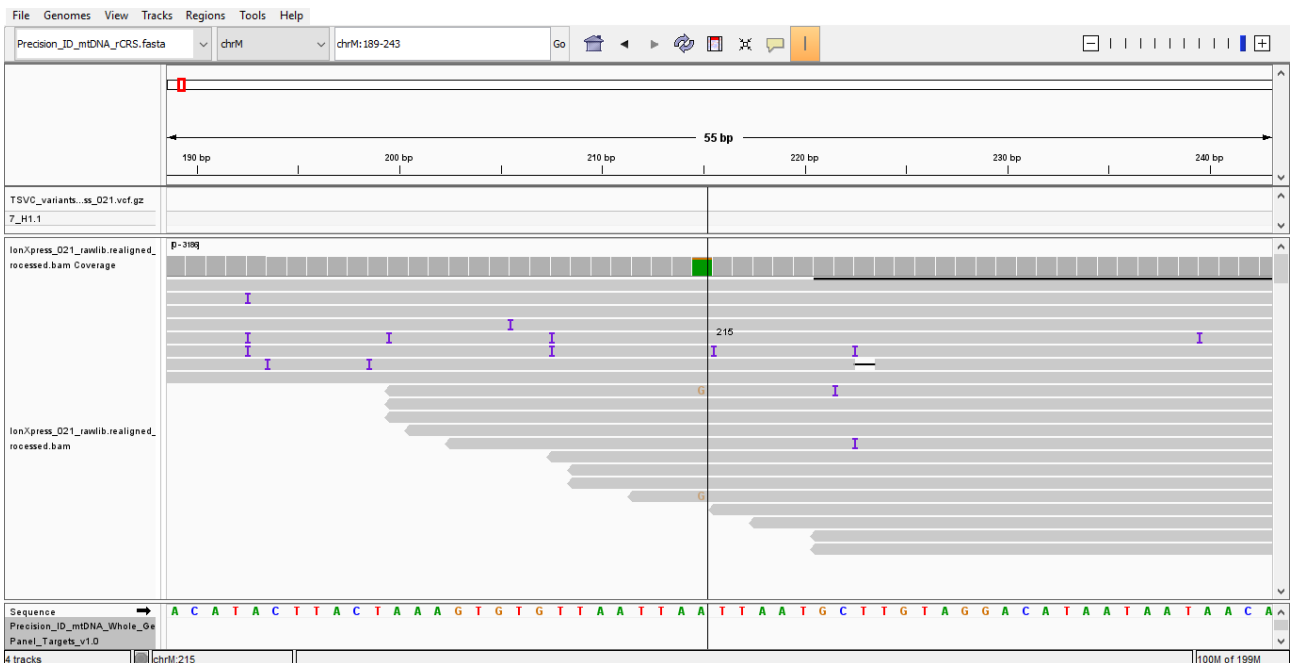


Fig. 29 Eteroplasmia in posizione 215, nel campione 7_H1.1 rilevata mediante analisi con IGV

Capitolo 4

DISCUSSIONE

L'analisi genetica del DNA mitocondriale (mtDNA) è diventata uno strumento importante nel campo della genetica forense, in gran parte a causa della sua abbondante presenza nelle cellule rispetto al DNA nucleare, fornendo un'alternativa di tipizzazione nel caso di campioni degradati, spesso rinvenuti sulle scene del crimine. Inoltre, poiché il mtDNA è ereditato esclusivamente per via matrilineare, consente l'identificazione di persone imparentate per via materna, di persone scomparse o vittime di disastri di massa (DVI). A causa della mancanza di ricombinazione, l'unica variazione osservata nel mtDNA è l'eteroplasmia, una miscela di diverse popolazioni di sequenze di mtDNA che si vengono a creare poiché i mitocondri contengono molte copie di mtDNA e le mutazioni puntiformi possono interessare solo una parte di queste, portando alla presenza di due o più sottopopolazioni in un singolo mitocondrio, o all'interno di una singola cellula o tra cellule diverse dello stesso individuo. Tradizionalmente, l'analisi del mtDNA era confinata alla Control Region (CR), che ospita i segmenti ipervariabili (HV1, HV2 e HV3), attraverso l'applicazione del Sanger Sequencing (STS). Tuttavia, l'emergere della tecnologia Massive Parallel Sequencing (MPS) ha permesso di sequenziare l'intero genoma mitocondriale. I capelli umani sono tra i reperti più comunemente riscontrati su una scena del crimine, ma sono noti per essere scarse fonti di DNA nucleare a meno che non siano presenti parti di tessuto adese o radici dei capelli¹⁷. Per questo motivo, la comunità forense si è concentrata sullo studio del mtDNA come

prezioso marker genetico nei campioni di capelli, monitorando principalmente la sua variazione lungo il fusto del capello. La variazione del mtDNA in tali campioni è causata dai colli di bottiglia che portano alla fissazione di varianti eteroplasmatiche. Numerosi studi si sono concentrati sullo studio del collo di bottiglia somatico e dei cambiamenti nelle frequenze alleliche dell'eteroplasma che si verificano nei capelli umani poiché la presenza di tali varianti può complicare l'analisi del mtDNA se i risultati devono essere confrontati con il materiale di riferimento. D'altra parte, una corretta interpretazione dell'eteroplasma può rafforzare il potere di identificazione del mtDNA¹⁸.

Per superare le difficoltà nell'interpretazione dei dati del mtDNA e con l'obiettivo di armonizzare la sua analisi in casi forensi, esperti nel campo della genetica forense hanno promosso un progetto collaborativo incentrato sulla rilevazione e interpretazione dell'eteroplasma nei capelli e nei campioni di riferimento attraverso l'applicazione della tecnologia MPS, il MitoMetrics Hair Heteroplasmy Study.

Considerando lo scopo di questo progetto e la necessità di ulteriori conoscenze sui livelli di varianti eteroplasmatiche dei capelli nel mitogenoma, abbiamo raccolto campioni di capelli da 5 diversi volontari unitamente a un tampone buccale e un campione di sangue, utilizzati come riferimenti per il confronto, per stabilire se le varianti rilevate sono diverse nei differenti tessuti. Siamo partiti dall'estrazione del mtDNA dalle diverse matrici. Dopo l'estrazione, i campioni di capelli sono stati quantificati utilizzando un test qPCR sviluppato recentemente che combina un target nDNA (181 bp), due target mtDNA (69 bp e 143 96 bp) e un IPC. Da questo test

siamo stati in grado di valutare la qualità e la quantità di mtDNA presente nei campioni, per determinare quali fossero adatti per l'analisi di varianti eteroplasmatiche attraverso la tecnologia MPS. I nostri risultati (Tabella 17) includono anche l'indice di degradazione per i campioni di capelli, che sembra aumentare lungo la lunghezza della maggior parte dei capelli analizzati, ma non per tutti. L'andamento decrescente della quantità di mtDNA lungo il fusto è stato studiato anche da Brandhagen e colleghi in un recente studio, dove hanno anche evidenziato come l'età, l'ambiente da cui viene recuperato e qualsiasi danno chimico/fisico possano influenzare il corretto recupero di un profilo mtDNA. È interessante notare che Brandhagen ha anche testato diversi protocolli di estrazione, mostrando come questi possono influire sul recupero del DNA. Non è ancora presente un protocollo convalidato e uniforme che consenta un elevato recupero di materiale genetico da campioni più complessi come i fusti dei capelli.

Sin dalla sua introduzione, il Massive Parallel Sequencing si è dimostrato uno strumento prezioso nel campo dell'analisi del sequenziamento. Diverse aziende hanno ora sviluppato dei pool di primer PCR utili per generare librerie MPS, consentendo riproducibilità e una maggiore qualità del test del mtDNA. Il Precision ID mtDNA Whole Genome Panel prodotto dalla Thermo Fisher Scientific è costituito da 162 ampliconi divisi in due pool multiplex (81 coppie di primer ciascuno) con una dimensione del frammento di 175 bp, l'utilizzo di ampliconi corti lo rende adatto anche a campioni degradati¹⁹. Abbiamo deciso di utilizzare questo pannello per analizzare i nostri campioni per studiare le varianti eteroplasmatiche sia nei capelli

che nei campioni di riferimento. La preparazione della libreria è stata eseguita manualmente, seguendo la Precision ID mtDNA Panels Application Guide (TFS), utilizzando il metodo 2-in-1 per i campioni di capelli e il metodo conservativo per i campioni di riferimento. Una volta che gli ampliconi sono stati digeriti utilizzando il FuPa reagent e i rispettivi Barcode Adapters, le librerie sono state purificate utilizzando il reagente Agencourt AMPure XP e lavaggi con etanolo. Dopo la purificazione, le librerie Precision ID sono state quantificate mediante qPCR utilizzando lo Ion Library TaqMan Quantitation Kit, con reazioni duplicate per ciascun campione e le diluizioni delle librerie di controllo. Le librerie sono state quindi raggruppate e caricate sullo strumento Ion Chef™ per la preparazione del template, e il chip 520™ risultante è stato sequenziato dal sequencer HID Ion S5™. I dati primari di sequenziamento sono stati ottenuti dal software Torrent Suite™ utilizzando i plug-in del Coverage Analysis e del Variant Caller (TVC). La soglia di profondità di copertura minima di 20X è stata utilizzata per filtrare le chiamate delle varianti, mentre la soglia per la chiamata delle eteroplasmie è stata fissata al 10%. Tali soglie sono state scelte sulla base di studi recenti sulla rilevazione di varianti eteroplasmatiche dei capelli utilizzando il Precision ID mtDNA Whole Genome Panel (Strobl, 2019¹⁹; Churchill Cihlar, 2020²⁰; Faccinnetto, 2021²¹). Tutte le letture sequenziate sono state allineate con la Precision_mtDNA_rCRS. Il Relative Depth of Coverage è stato calcolato per i 162 ampliconi inclusi nel pannello del genoma intero del mtDNA Precision ID Whole Genome Panel (Thermo Fisher Scientific). Tali risultati ci hanno permesso di valutare la performance del pannello e ci ha permesso

di individuare 8 (mt_4, mt_21, mt_23, mt_39, mt_72, mt_84, mt_97 e il mt_105) ampliconi che presentano un'efficienza di amplificazione più bassa rispetto gli altri. Studi recenti di altri gruppi che lavorano con lo stesso pannello hanno indicato la presenza di ampliconi a bassa profondità di lettura. Strobl et al¹⁹ hanno identificato 5 ampliconi con prestazioni inferiori agli altri: mt_3, mt_73, mt_97, mt_101 e mt_105. Nel 2018, Pereira e colleghi si sono imbattuti anche in una serie di 7 ampliconi che registravano valori di profondità di lettura al di sotto dei tassi previsti, tra cui: mt_4, mt_36, mt_73, mt_82, mt_84, mt_90 e mt_97. Il tratto poli-C intorno alla posizione 310 aveva una bassa profondità di lettura anche in altri studi e potrebbe in parte spiegare il basso numero di letture per il frammento mt_4.

Per i campioni valutati in questo studio è stato osservato uno sbilanciamento dei due pool per diversi campioni (Grafico 7). Da notare che non vi è né un pool che venga preferenzialmente amplificato rispetto all'altro né vi è una relazione con matrice biologica sottoposta ad analisi, di conseguenza lo sbilanciamento di questi appare del tutto casuale. Il motivo dello sbilanciamento dei pool nei nostri campioni non è del tutto chiara, potrebbe essere dovuto alla distribuzione disuguale dei pool durante la preparazione manuale delle librerie, causato dall'elevata viscosità dei reagenti.

A seguito dell'analisi di identificazione delle varianti da parte del plug-in TVC, i file BAM e BAI sono stati ispezionati con il software IGV.

Per il donatore 4 è stato possibile identificare un aplotipo costituito da 31 varianti, di cui 8 sono presenti nella regione di controllo e 23 nella regione codificante. Un

numero superiore di varianti, sono state rilevate per il donatore 5, 34 totali di cui 11 presenti nella regione di controllo e 23 nella regione codificante. Infine, il donatore 7 presenta un aplotipo costituito da 31 varianti, di cui 6 sono presenti nella regione di controllo e 25 nella regione codificante. Il software EMPOP mtDNA Database v4/R13 ha permesso l'identificazione degli aplogruppi J2b1, U1b3 e T2b rispettivamente per il donatore 4, donatore 5 e il donatore 7.

Dal confronto dei risultati ottenuti dall'analisi primaria tramite Variant Caller plug-in e la successiva ispezione manuale tramite IGV software, in totale sono state osservate 34 chiamate discordanti. Tuttavia, la maggior parte di esse sono state trovate sui campioni 4_H1.12 e 7_H1.3, i quali presentano un numero molto elevato di eteroplasmie (>4 posizioni eteroplasmiche per singolo campione) che ci inducono a pensare ad una contaminazione.

Il donatore 5 presenta aplotipi completi per tutte le matrici testate. I due riferimenti presentano stesso aplotipo tranne per la posizione 12308 eteroplasmica nel 5_S ma totalmente mutata nel 5_B.

Per quanto riguarda il primo capello del donatore 5, in tutti i frammenti è stata rilevata l'eteroplasmia 16261Y, riscontrata anche nei riferimenti. Nel primo frammento, il 5_H1.1 è comparsa l'eteroplasmia 9925R. Le posizioni 15148 e 15172 appaiono omoplasmiche in tutti i frammenti ad eccezione del 5_H1.6. Nell'ultimo frammento del primo capello, il 5_H1.10 è presente l'eteroplasmia 12308R rilevata anche nel tampone buccale.

Il secondo capello del donatore 5 (5_H2) presenta in tutti i frammenti lo stesso aplotipo, allo stesso modo il terzo capello (5_H3) presenta lo stesso aplotipo tranne per la posizione C16261T che risulta mutato solo per il 5_H3.1, mentre nei riferimenti questa posizione mostra un eteroplasmia.

Del donatore 4 abbiamo valutato solo due dei tre capelli richiesti poiché il capello 4_H2 alla quantizzazione delle librerie MPS non soddisfaceva la concentrazione richiesta per il sequenziamento. Andrà approfondito in futuro.

Per il 4_S si è riscontrato lo sbilanciamento dei due pool quindi per questo abbiamo solo un aplotipo parziale, tuttavia, le posizioni identificate sono le stesse per il tampone buccale del medesimo donatore. In entrambi sono presenti due eteroplasmie, la prima in posizione 5460R la seconda in posizione 14032Y (4_S: G75%, A 25%; 4_B: G70%, A30% per la prima e 4_S: T77%, C23%; 4_B: T 72%, C28% per la seconda). Dato il numero anomalo delle eteroplasmie riscontrate non è possibile fare un confronto attendibile per gli aplotipi riscontrati per il capello 2. Invece, per il capello 3 nell'ultimo frammento è presente l'eteroplasmia 295Y che nei primi due frammenti appare totalmente mutata e la 14302Y che è presente anche nei riferimenti e risulta totalmente mutata nel primo frammento del terzo capello. Bisogna però sottolineare che per il frammento 4_H3.9 è presente uno sbilanciamento dei pool per questo possiamo dire di aver ottenuto un risultato soltanto parziale.

Il donatore 7 mostra lo stesso aplotipo nei riferimenti per le posizioni interpretabili nel 7_S. Il primo capello del donatore 7, come per il 4_H1 presenta un numero

eccessivo di eteroplasmie, per questo abbiamo pensato anche qui ad una possibile contaminazione. Però, particolare è la presenza di un'eteroplasma, la 215R presente in tutti i frammenti del primo capello in percentuali diverse (7_H1.1: A 88%, G 12%; 7_H1.2: A 50%, G 50%; 7_H1.3: A 79%, G21%). I frammenti 7_H2.2, e 7_H2.3 presentano lo sbilanciamento rispettivamente per il pool 1 e per il pool 2 questo non ci permette di fare un confronto completo con l'aplotipo del frammento più prossimale, tuttavia, le varianti confrontabili concordano tutte tranne per la posizione 11812R.

Per il terzo capello i frammenti 7_H3.1 e 7_H3.2 risulterebbero confrontabili ma anche qui è stato riscontrato uno sbilanciamento dei due pool che non ci permette di definire un aplotipo completo. L'ultimo frammento del terzo capello presenta le eteroplasmie 709R e 1888R, mutate in tutti gli altri frammenti,

L'interpretazione della deriva eteroplasmica in diversi tessuti può essere complicata.

Lo scopo finale della tipizzazione del mtDNA è determinare se un individuo può o non può essere escluso come possibile fonte del mtDNA testato in studi forensi, rispetto ai campioni di riferimento. Seguendo sia le linee guida ISFG che SWAGDAM, se il campione in esame mostra la variante eteroplasmica e il riferimento no, è necessario selezionare un risultato inconcludente poiché i livelli di eteroplasma possono variare notevolmente tra i tessuti degli stessi individui o tra individui imparentati con la madre. In caso contrario, i campioni che differiscono per due o più posizioni nucleotidiche possono essere esclusi come provenienti dalla stessa fonte o linea materna. Pertanto, è essenziale identificare correttamente le

varianti eteroplasmatiche e la maggiore sensibilità delle piattaforme MPS può offrire un modo per rilevare queste molecole a bassa frequenza, che sono spesso essere impercettibili nell'elettroforesi capillare basata su Sanger. Nel nostro caso osservando il campione di riferimento 5_S, le due posizioni eteroplasmatiche non sono condivise da tutti i frammenti di capelli analizzati, il che potrebbe potenzialmente portare ad un'errata esclusione dei campioni in quanto provenienti dalla stessa linea materna.

Capitolo 5

CONCLUSIONE

Negli ultimi 25 anni, la tipizzazione del DNA mitocondriale è diventata un pilastro dell'identificazione umana e i progressi compiuti per migliorare il recupero di preziose informazioni genetiche da campioni forensi più complessi hanno il potenziale per aumentare il potere di discriminazione del mtDNA. Gli straordinari miglioramenti apportati sul campo sono stati possibili anche attraverso l'applicazione del Massive Parallel Sequencing, che analizza simultaneamente milioni di frammenti, consentendo l'analisi dei genomi mitocondriali completi nel giro di poche ore. La maggiore sensibilità e la natura quantitativa dei dati forniti dalle piattaforme MPS consentono il rilevamento e la caratterizzazione dell'eteroplasmia di livello inferiore rispetto alle consolidate tecnologie STS-CE.

Il flusso di lavoro MPS è stato testato su un ampio spettro di campioni forensi comunemente incontrati, ma relativamente pochi studi hanno affrontato l'uso di tali piattaforme per studiare i livelli di varianti eteroplasmatiche nel mitogenoma. Progetti di collaborazione come il Mitometrics Hair Heteroplasmic Study aiuteranno a costruire le basi per la valutazione delle mutazioni somatiche nei capelli umani, necessarie per migliorare l'interpretazione delle prove forensi. Considerando questo progetto e altri studi presenti in letteratura, la mia tesi si è concentrata sullo studio della variazione della sequenza del mtDNA nei fusti dei capelli. I capelli hanno stretti colli di bottiglia che portano alla segregazione delle molecole del mtDNA durante lo sviluppo dei capelli, con conseguente eteroplasmia o fissazione di una variante che

differisce dall'aplotipo osservato in altri tessuti, come i tamponi buccali di riferimento, comunemente raccolti nel contesto forense.

I capelli e i campioni di riferimento (sangue e tampone buccale) sono stati sottoposti a procedure di estrazione automatizzata, saggi di quantificazione in tempo reale e Massive Parallel Sequencing. I risultati qui discussi confermano che la quantità e la qualità del mtDNA diminuiscono lungo la lunghezza dei fusti del capello poiché abbiamo anche calcolato l'indice di degradazione dei diversi frammenti. Inoltre, su 24 campioni sequenziati con il pannello Precision ID dell'intero genoma mitocondriale, 14 hanno registrato una copertura del mitogenoma uguale o superiore al 95%, mentre i restanti mostravano una copertura più bassa (< 70%), a causa della bassa profondità di copertura relativa di alcuni degli ampliconi inclusi nel pannello. L'ulteriore analisi delle letture sequenziate ha confermato la capacità del pannello Precision ID dell'intero genoma di rilevare varianti eteroplasmatiche di basso livello in campioni complessi, che sarebbero state impercettibili nell'elettroforesi capillare basata sul Sanger. Un totale di 26 varianti PHP, sono state identificate in 13 diversi siti. Inoltre, sono state rilevate un elevato numero di eteroplasmie in due diversi donatori che fanno presupporre una contaminazione. Dei PHP rilevati la16261Y era presente in 4 diversi campioni e mostrava un'inversione del nucleotide maggiore in base al tipo di tessuto. I risultati ottenuti evidenziano il throughput e la sensibilità della tecnologia MPS, per il rilevamento di varianti eteroplasmatiche di basso livello. Anche se il test di routine del mtDNA sui campioni di riferimento può essere abbastanza semplice, la presenza di eteroplasmia, precedentemente considerata come l'eccezione piuttosto che

la regola, può complicare l'interpretazione dei risultati. Ciò è particolarmente vero, nel caso di campioni di capelli, a causa dell'elevata variabilità del mtDNA, che sembra essere il più alta nei capelli rispetto ad altri tessuti. La formulazione di linee guida per l'analisi e l'interpretazione delle varianti eteroplasmatiche rilevate tramite MPS faciliterebbe ovviamente l'armonizzazione dell'analisi del mtDNA nel contesto forense.

RINGRAZIAMENTI

Dedico questo elaborato di tesi a tutti quelli che mi sono stati vicini durante questo percorso universitario. Alla mia famiglia, in particolare, ai miei genitori che mi hanno permesso di coltivare il mio futuro, spero che possano raccogliere i frutti dei loro sacrifici. Vi sarò sempre grata.

Alle mie care amiche Giulia e Adriana, con voi ho condiviso questa esperienza, tra risate e qualche momento di disperazione ci siamo sempre aiutate l'un l'altra, grazie per aver reso questo percorso più leggero e spensierato.

Alla mia seconda famiglia, i miei amici, quelli che ci sono da sempre, che anche se distanti mi sono vicini col pensiero e col cuore, Lucia, Francesco, Alessia, Dalila, Francesca, Cristina.

Grazie alla mia relattrice, la prof.ssa Chiara Turchi per essersi resa sempre disponibile per ogni chiarimento o dubbio. Ringrazio anche la dottoressa Filomena Melchionda per avermi seguita in questo progetto di tesi, e di avermi trasmesso nuove conoscenze di cui farò tesoro.

In fine, un ringraziamento particolare a tutti quelli che non hanno creduto nelle mie capacità, nella mia tenacia e determinazione. Queste convinzioni sono state il mio carburante, che mi ha permesso di andare sempre avanti e di riuscire a superare ogni ostacolo. Termino con l'augurio più sincero per me stessa di reagire ad ogni difficoltà che incontrerò nella vita e di continuare a credere nei miei sogni, nonostante tutti gli impedimenti che vi possano essere.

Per vivere a pieno la vita è necessario accettarne la caducità.

Sigmund Freud, Caducità, 1915.

BIBLIOGRAFIA

- ¹ Introduzione alla genetica forense - Indagini di identificazione personale e di paternità di Adriano Tagliabracci; Springer Editore
- ² Biologia- cellula e tessuti; Edi- ermes
- ³ António Amorim, Teresa Fernandes, Nuno Taveira. Mitochondrial DNA in human identification: a review. PeerJ.2019 Aug 13;7:e7314. doi: 10.7717/peerj.7314. eCollection 2019. PMID: 31428537 PMCID: PMC6697116
- ⁴ M M Holland, T J Parsons. Mitochondrial DNA Sequence Analysis - Validation and Use for Forensic Casework. PMID: 26255820
- ⁵ Wilson MR, Polansky D, Replogle J, DiZinno JA, Budowle B. A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. Hum Genet. 1997 Aug;100(2):167-71. DOI: 10.1007/s004390050485. PMID: 9254844.
- ⁶ Jana Naue , Steffen Hörer , Timo Sängler , Christina Strobl , Petra Hatzer-Grubwieser , Walther Parson , Sabine Lutz-Bonengel PMID: 25526677
DOI: 10.1016/j.mito.2014.12.002
- ⁷ Bruce Budowle, Marc W Allard, Mark R Wilson, Ranajit Chakraborty. Forensics and Mitochondrial DNA: Applications, Debates, and Foundations.
- ⁸ Barrett A, Arbeithuber B, Zaidi A, Wilton P, Paul IM, Nielsen R, Makova KD. The pronounced somatic bottleneck in mitochondrial DNA of human hair. Philos Trans R

Soc Lond B Biol Sci. 2020 Jan 20;375(1790):20190175. DOI:

10.1098/rstb.2019.0175. Epub 2019 Dec 2. PMID: 31787049; PMCID: PMC6939377

⁹ C A Linch, D A Whiting, M M Holland. Human Hair Histogenesis for the Mitochondrial DNA Forensic Scientist. PMID: 11451065

¹⁰ Charla Marshall 1, Walther Parson. PMID: 33740708 DOI:

10.1016/j.fsigen.2021.102497

¹¹ Lutz-Bonengel S, Niederstätter H, Naue J, Koziel R, Yang F, Sängler T, Huber G, Berger C, Pflugradt R, Strobl C, Xavier C, Volleth M, Weiß SC, Irwin JA, Romsos EL, Vallone PM, Ratzinger G, Schmuth M, Jansen-Dürr P, Liehr T, Lichter P, Parsons TJ, Pollak S, Parson W. Evidence for multi-copy Mega-NUMTs in the human genome. *Nucleic Acids Res.* 2021 Feb 22;49(3):1517-1531. doi:

10.1093/nar/gkaa1271. PMID: 33450006; PMCID: PMC7897518.

¹² Gasparre G., Porcelli A.M., *The Human Mitochondrial Genome from Basic Biology to Disease*, Academic Press, 2020.

¹³ Claudia V Pereira, Bryan L Gitschlag, Maulik R Patel. Cellular mechanisms of mtDNA heteroplasmy dynamics. PMID: 34120542 DOI:

10.1080/10409238.2021.1934812

¹⁴ Xavier C, Eduardoff M, Strobl C, Parson W. SD quants-Sensitive detection tetraplex- system for nuclear and mitochondrial DNA quantification and degradation inference. *Forensic Sci Int Genet.* 2019 Sep; 42:39-44. DOI:

10.1016/j.fsigen.2019.06.004. Epub 2019 Jun 10. PMID: 31216503

¹⁵ MitoMetrics Hair Heteroplasmy Study – Protocol

¹⁶Brandhagen MD, Loreille O, Irwin JA, Fragmented Nuclear DNA is the Predominant Genetic Material in Human Hair Shafts, *Genes (Basel)*, 2018 Dec 18;9(12):640, DOI: 10.3390/genes9120640, PMID: 30567392; PMCID: PMC6316335.

¹⁷Desmyter S, Bodner M, Huber G, Dognaux S, Berger C, Noël F, Parson W. Hairy matters: MtDNA quantity and sequence variation along and among human head hairs. *Forensic Sci Int Genet.* 2016 Nov; 25:1-9. DOI: 10.1016/j.fsigen.2016.07.012. Epub 2016 Jul 25. PMID: 27484846.

¹⁸Just RS, Irwin JA, Parson W. Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. *Forensic Sci Int Genet.* 2015 Sep; 18:131-9. DOI: 10.1016/j.fsigen.2015.05.003. Epub 2015 May 6. PMID: 26009256; PMCID: PMC4550493.

¹⁹Strobl C, Eduardoff M, Bus MM, Allen M, Parson W. Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. *Forensic Sci Int Genet.* 2018 Jul; 35:21-25. DOI: 10.1016/j.fsigen.2018.03.013. Epub 2018 Mar 29. PMID: 29626805.

²⁰Cihlar JC, Amory C, Lagacé R, Roth C, Parson W, Budowle B. Developmental Validation of an MPS Workflow with a PCR-Based Short Amplicon Whole Mitochondrial Genome Panel. *Genes (Basel)*. 2020 Nov 13;11(11):1345. DOI: 10.3390/genes11111345. PMID: 33202822; PMCID: PMC7709034.

²¹Faccinnetto C, Sabbatini D, Serventi P, Rigato M, Salvoro C, Casamassima G, Margiotta G, De Fanti S, Sarno S, Staiti N, Luiselli D, Marino A, Vazza G. Internal validation and improvement of mitochondrial genome sequencing using the Precision ID mtDNA Whole Genome Panel. *Int J Legal Med.* 2021 108 Nov;135(6):2295-2306. doi: 10.1007/s00414-021-02686-w. Epub 2021 Sep 7. PMID: 34491421; PMCID: PMC8523450.

SITOGRAFIA IMMAGINI

Fig.1,4 Introduzione alla genetica forense - Indagini di identificazione personale e di paternità di Adriano Tagliabracci; Springer Editore

Fig. 2 www.unthsc.edu

Fig. 3 www.chimica-online.it

Fig. 4 Mark D Shriver , Rick A Kittles Genetic ancestry and the search for personalized genetic histories. PMID: 15266343 DOI: 10.1038/nrg1405

Fig. 5 Denise Syndercombe Court. Mitochondrial DNA in forensic use. PMID: 34374411 PMCID: PMC8457767 DOI: 10.1042/ETLS20210204

Fig. 6,7,8,9. MitoMetrics Hair Heteroplasmy Study – Protocol

Fig, 10,11, MAN0017770_PrecisionID_mtDNA_Panels_S5_UG

Fig, 12, 13, 14,15, SeqScape Software 3

Fig, 16 software Rotor Gene Q v, 2,1,0 di Qiagen

Fig. 17,18, 19, 20, 21, 22 Torrent Suite Software (v5.12.3)

Fig. 23, 24, 25, 26, 27 IGV Software