

*Alla mia famiglia,
colonne portanti della mia vita.*

*A mia nonna Lucia,
la mia piccola farfalla bianca,
che da lassù è sempre con me.*



UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

**Corso di Laurea Magistrale
Biologia Molecolare e Applicata**

**L'eteroplasmia del DNA mitocondriale in genetica forense: Analisi e strategie
interpretative delle varianti eteroplasmiche mediante Massive Parallel
Sequencing (MPS)**

**Mitochondrial DNA heteroplasmy in forensic genetics: Analysis and
interpretative strategies of heteroplasmic variants using Massive Parallel
Sequencing (MPS)**

Tesi di Laurea Magistrale di

Lucia Torre

Relatore Chiar.mo Prof.

Chiara Turchi

Sessione Autunnale

Anno Accademico 2022/2023

Sommario

CAPITOLO 1	6
INTRODUZIONE	6
1.1 LA GENETICA FORENSE	6
1.2 TIPIZZAZIONE DEL DNA NELL'IDENTIFICAZIONE FORENSE	7
1.3 INTRODUZIONE MITOCONDRIO	9
1.4 IMPORTANZA mtDNA IN GENETICA FORENSE	13
1.5 ETEROPLASMIA.	18
1.6 ETEROPLASMIA E MALATTIA	20
1.8 SEQUENZE MITOCONDRIALI NUCLEARI (NUMT)	23
1.9 ETEROPLASMIA RILEVATA CON MPS	25
1.11 DATABASE MTDNA	26
CAPITOLO 2	30
MATERIALI E METODI	30
2.1 RACCOLTA CAMPIONI	30
2.2 ESTRAZIONE	33
2.2.3 ESTRAZIONE DA CAMPIONI DI SANGUE	33
2.2.5 ESTRAZIONE DA CAPELLI	38
2.3 QUANTIFICAZIONE	39
2.3.1 FUSTI DI CAPELLI	39

2.3.2 QUANTIFICAZIONE QUBIT™ per campioni di riferimento	41
2.4. PREPARAZIONE MANUALE DELLA LIBRERIE	42
2.4.1 REAZIONE DI AMPLIFICAZIONE DEL mtDNA TARGET	43
2.4.3 LEGAME DEI BARCODE E PURIFICAZIONE DELLA LIBRERIA	47
2.4.4 QUANTIFICAZIONE LIBRERIA TRAMITE qPCR	49
2.4.5 LIBRERIE PREPARATE MEDIANTE Ion Chef™ SYSTEM	51
2.4.6 PREPARAZIONE TEMPLATE SU HID Ion Chef™ SYSTEM	51
2.5 ANALISI DEI RISULTATI	52
CAPITOLO 3	54
RISULTATI	54
3.1 QUANTIFICAZIONE ESTRATTI	54
3.2 RISULTATI QUANTIFICAZIONE LIBRERIE	61
3.3 RISULTATI SEQUENZIAMENTO	63
3.4. ANALISI DEL COVERAGE	65
3.4 ANALISI VARIANTI ETEROPLASMICHE	68
CAPITOLO 4	88
DISCUSSIONE	88
CAPITOLO 5	95
CONCLUSIONI	95
BIBLIOGRAFIA:	97

<i>SITOGRAFIA IMMAGINI</i>	100
<i>RINGRAZIAMENTI</i>	102

CAPITOLO 1

INTRODUZIONE

1.1 LA GENETICA FORENSE

La genetica forense è una disciplina che fa uso di moderne tecniche di biologia molecolare all'interno di un contesto forense, per l'identificazione di autori di reati, l'attribuzione individuale di resti umani sconosciuti, o l'attribuzione di parentela. I campioni prevalentemente utilizzati sono: capelli, sangue, tamponi buccali, o liquido seminale.⁽¹⁾ Ogni organismo vivente ha un proprio assetto genetico. Questo genoma, noto come DNA contiene al suo interno una serie di geni indispensabili per la produzione di proteine ed enzimi necessari per il corretto svolgimento delle funzioni vitali e fisiologiche degli organismi. Nel DNA vi sono specifiche sequenze che si distinguono per le funzioni: i tratti detti "codificanti" sono la sequenza che verrà trascritta in RNA e poi tradotta in proteine; i tratti detti "non codificanti" rappresentano la sequenza che ha funzione di regolazione dell'espressione genica. Grazie all'estrazione di DNA dal materiale biologico si riesce a delineare, il profilo di un individuo utile per poter risalire ad una probabile colpevolezza. Ecco che la genetica forense ha un

ruolo importante grazie alle sue indagini, e fa uso di varie tecniche e conoscenze come la biologia molecolare, genetica e medicina legale.

1.2 TIPIZZAZIONE DEL DNA NELL'IDENTIFICAZIONE FORENSE

Tipizzare il DNA ci permette di vedere la variabilità e di avere informazioni per poter identificare e distinguere un individuo dall'altro, la parte variabile che ci permette ciò è pari allo 0.3%.⁽²⁾ Tra i marcatori genetici più conosciuti e studiati in ambito forense abbiamo i microsatelliti o short tandem repeats (STRs). Molto spesso queste regioni vengono chiamate DNA satellite, e sono situati intorno al centromero dei cromosomi. Questi marcatori ovvero gli STRs sono i migliori candidati grazie alla loro ridotta sequenza di DNA che dà origine alla sequenza ripetuta in tandem la quale risulta essere formata da 2 a 7 paia di basi. L'insieme delle unità ripetute formano la 'regione ripetuta', formata da un numero variabile di ripetizioni nei vari individui, ciò li rende utili nell'identificazione umana, in quanto possiedono un alto potere discriminatorio. Gli STRs vengono inoltre definiti dinucleotidici, trinucleotidici, tetranucleotidici, pentanucleotidici ed esanucleotidici, costituiti rispettivamente da 2, 3, 4, 5 e 6 nucleotidi per unità ripetitiva. Quelli maggiormente utilizzati nelle indagini forensi sono gli STRs costituiti da ripetizioni trinucleotidiche, in quanto hanno la capacità di evitare la produzione di stutter durante la PCR, ovvero ampliconi, ossia regioni di DNA prodotte in

PCR e risultano essere più corti o formati da più unità ripetute rispetto all'allele. La loro creazione avviene in seguito ad uno scivolamento della polimerasi sullo stampo di DNA. La loro presenza ha un effetto negativo sull'indagine in quanto rende più difficile l'interpretazione di tracce miste di DNA, dove sono presenti più profili genetici appartenenti a individui diversi. Un vantaggio che comporta l'uso di STRs trinucleotidi, risiede nella loro grande facilità di essere distinguibili durante la separazione elettroforetica, dove il principio base della separazione è la lunghezza. Un'altra cosa che possiamo fare è estrarre e amplificare più STRs simultaneamente, usando un mix di primer diretti verso diversi STRs, tramite la 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. Andando a tipizzare i 13 microsatelliti del CODIS è possibile identificare in modo preciso e affidabile un individuo poiché la random match probability ovvero, 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 anche andare a tipizzare contemporaneamente questi microsatelliti in poco tempo e con poco DNA di partenza tramite l'uso vari kit che troviamo 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. Inoltre possono essere utilizzati per la ricostruzione di linee parentali, soprattutto per analisi di polimorfismi sul cromosoma Y, ciò ci permette di accertare una paternità, al contrario i polimorfismi del DNA mitocondriale possono essere utilizzati per confermare l'ereditarietà materna.

1.3 INTRODUZIONE MITOCONDRIO

Il mitocondrio⁽³⁾ è un organello a doppia membrana con forma allungata, presente in tutti gli eucarioti, ed è dotato di un DNA proprio, ovvero il DNA mitocondriale.

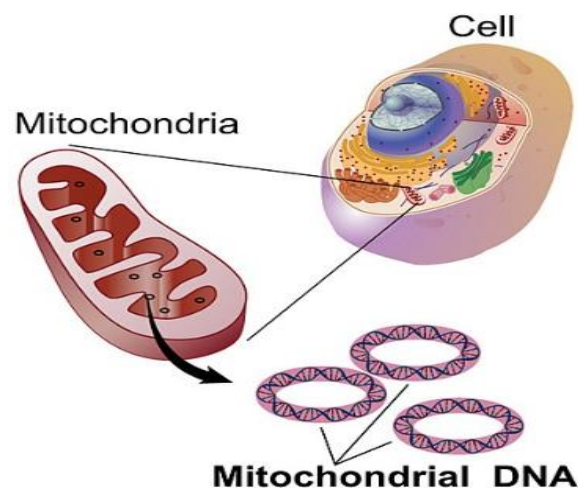


Fig.1. Rappresentazione mtDNA all'interno del mitocondrio

I mitocondri sono gli organelli addetti alla respirazione cellulare, costituiti da sacchette contenenti enzimi respiratori. Sono formati da due membrane: la

membrana interna e la membrana esterna; lo spazio compreso fra queste è detto spazio intermembrana. Lo spazio delimitato dalla membrana interna è detto matrice mitocondriale; la membrana interna si estende nella matrice formando delle pieghe dette creste mitocondriali essa delimita la camera interna definita matrice mitocondriale, costituita da un gel discretamente viscoso nel quale si trovano: gli enzimi per l'ossidazione degli acidi grassi e per il funzionamento del ciclo di Krebs, e i granuli contenenti calcio e magnesio per la produzione di ATP.

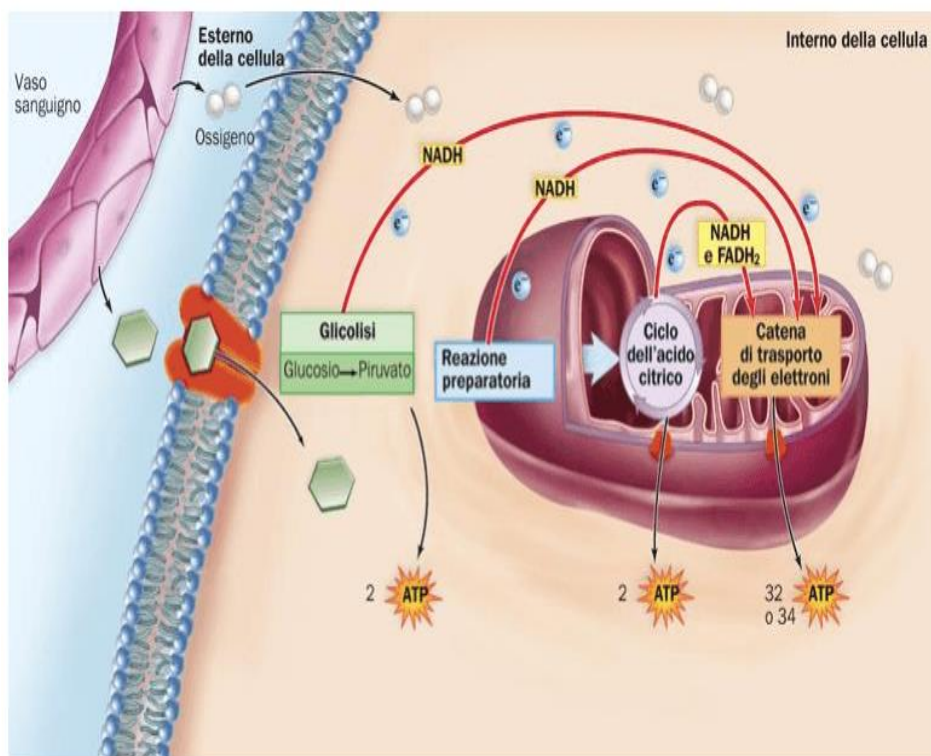


Fig.2 Schema respirazione cellulare

I mitocondri svolgono un ruolo essenziale nel ciclo dell'acido tricarbossilico (TCA), nella β -ossidazione degli acidi grassi e la manipolazione del calcio,

nella regolazione dell'apoptosi intrinseca e sono a loro volta coinvolti nel ciclo cellulare. Il DNA mitocondriale (mtDNA) è stato identificato e isolato per la prima volta da Margit Nass e Sylvan Nass nel 1963, i quali studiando alcune fibre mitocondriali ponendo la loro attenzione su alcune caratteristiche come: fissazione, stabilizzazione e colorazione e sembravano essere correlate al DNA⁽⁴⁾ L' mtDNA ha una struttura a doppio filamento chiuso, nelle cellule umane, il mtDNA consta di 16.569 paia di basi e codifica 37 geni, inclusi 13 polipeptidi, due RNA ribosomiali e 22 tRNA. Un polipeptide (ND6) e otto tRNA si trovano sul filamento L; gli altri 12 polipeptidi, due rRNA e 14 tRNA sono codificati dal filamento H. Si possono distinguere due filamenti: 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 (93%), ma troviamo una regione non codificante, chiamata 'regione di controllo' localizzata tra i geni per il tRNA della prolina e per il tRNA della fenilalanina, a questa regione è attribuito il tasso più alto di variabilità. Questa regione è lunga 1.122 bp ed è situata tra la posizione 16.024 e la posizione 576 e rappresenta il 5-7% del DNA mitocondriale. Nella regione di controllo è presente un OH, in seguito a ciò questa regione è definita 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). Nella regione di controllo inoltre troviamo la regione ipervariabile 1(HVR1) situata tra la posizione 16.024 e la 16.365 e la regione ipervariabile 2 (HVR2) situata tra la posizione 73 e 340, ma oltre a queste regione ne è presente una terza HVR3 situata tra la posizione 483 e 574, tutte dotate di elevato polimorfismo. La sequenza mtDNA ci permette di definire la sequenza dell'aplotipo, ovvero la combinazione di alleli a diversi siti polimorfici sullo stesso cromosoma. L'analisi della sequenza del mtDNA ci permette di definire 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. Gli aplogruppi vengono nominati con delle lettere in ordine dalla A alla Z, in base al momento della scoperta:

-la maggior parte 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);

-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;

-le popolazioni afro-americane, gli aplogruppi più comunemente osservati sono L2a, L1c, L1b e L3b;

-i principali aplogruppi trovati in individui provenienti da popolazioni asiatiche sono gli aplogruppi M e N3.⁽⁵⁾

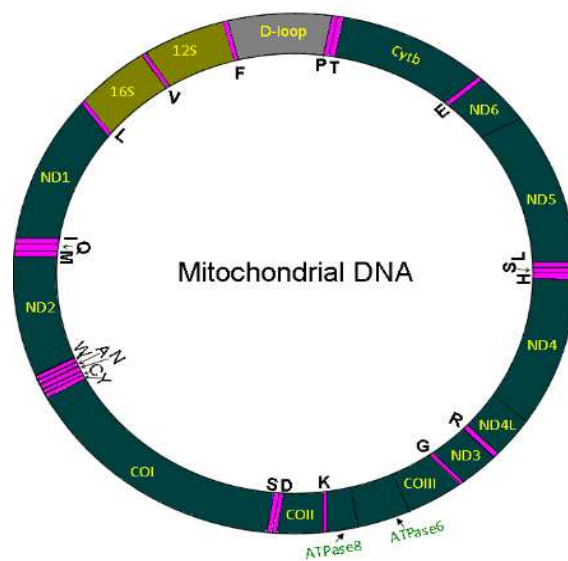


Fig.3 Struttura mtDNA

1.4 IMPORTANZA mtDNA IN GENETICA FORENSE

Il genoma mitocondriale⁽⁴⁾ viene 'ereditato uniparentalmente' dalla madre, in quanto i mitocondri vengono trasmessi dalla madre alla prole senza risentire di

alcuna influenza maschile, ciò fa sì che non sia soggetto alla ricombinazione genica. L'uso principale del mtDNA nell'attività giudiziaria è legato alla ricostruzione delle relazioni genetiche nel caso di identificazione di persone scomparse o indagini su disastri di massa. L'analisi del DNA mitocondriale (mtDNA) può fornire risultati attendibili in campioni forensi con DNA danneggiato. Il motivo principale che permette il recupero del mtDNA da campioni degradati è che in una cellula umana l'mtDNA è presente in un numero più alto di copie rispetto al DNA nucleare e le molecole di mtDNA hanno maggiore probabilità di sopravvivere rispetto al DNA autosomico, inoltre l'mtDNA risulta essere protetto dalla degradazione enzimatica rispetto al DNA genomico, grazie alla forma circolare che lo rende meno suscettibile all'attività esonucleasica. Pertanto possiamo concludere affermando che la probabilità di ottenere un profilo del mtDNA è maggiore. Quando non si può escludere che il profilo mtDNA di un campione di prova in dubbio e di un campione di riferimento provengano dalla stessa fonte o stessa linea materna, sarà valutata una stima statistica di corrispondenza, ottenendo un peso statistico che ci permetterà di poter a

rrivare ad una conclusione. Per fare ciò si procede facendo una stima della frequenza del tipo mtDNA in un database di popolazione appropriato. L'eredità materna senza ricombinazione ci sta ad indicare che le variazioni nucleotidiche

che caratterizzano gli aplotipi di un mtDNA devono essere considerata come unico locus. La tecnica comunemente utilizzata per determinare la rarità di un tipo di mtDNA tra gli individui, consiste nel contare il numero di volte in cui un particolare aplotipo viene riscontrato in un database, questo metodo viene denominato "metodo di conteggio". Nel caso in cui la dimensione del database è piccola e non riesce a rappresentare tutti i probabili contributori del tipo e della statistica del mtDNA per una corrispondenza casuale, la stima della frequenza potrebbe venire falsata. A tal proposito è importante prendere in considerazione l'intervallo di confidenza il quale tiene conto della dimensione del database e delle variazioni del campionamento. Gli intervalli di confidenza possono essere usati per stimare i limiti superiore e inferiore di un calcolo di frequenza e possono offrire la garanzia che tutti i dati ottenuti dovrebbero includere il valore reale del parametro nella proporzione di tempo stabilita dal livello di confidenza. Se si osserva un aplotipo del mtDNA, un determinato numero di volte (X), in un database di popolazione contenente N profili, la sua frequenza (p) può essere calcolata come segue: $p = X:N$. Alternativamente, la stima per gli approcci probabilistici viene calcolata come segue: nel caso in cui l'incertezza è dovuta a errori di campionamento, viene calcolata approssimativamente aggiungendo i profili dei casi al database:

$$p = \frac{(X+1)}{(N+1)} \quad \text{o in alternativa possiamo utilizzare } p = \frac{(X+2)}{(N+2)}$$

quando nel database sono presenti uno o più campioni, è possibile inserire un intervallo di confidenza del limite superiore del 95% sulla frequenza del profilo

utilizzando: $P + 1.96 \sqrt{\frac{p-(1-p)}{N}}$

Un approccio alternativo è il metodo Clopper-Pearson il quale può essere utilizzato anche per fornire una stima più conservativa per l'intervallo di confidenza superiore al 95%, dove si osservano conteggi molto bassi da un database di aplotipi. Nel caso in cui i profili non sono stati osservati in un database, il limite superiore del 95% sull'intervallo di confidenza è: $1-\alpha^{1/N}$

dove $\alpha=0,05$ e ci indica il coefficiente di confidenza, N è il numero di individui nel database. La rarità di un profilo mtDNA viene calcolata come segue:

$$p=1-(0.05)^{1/N}$$

Possiamo notare che l'intervallo di confidenza utilizzato risulta essere problematico in base alle dimensioni del campione. I rapporti di verosimiglianza (LR) per il mtDNA sono stimati da $1/(\text{probabilità di corrispondenza})$. In ambito forense, nei casi che coinvolgono sia la tipizzazione del mtDNA che gli STR autosomici, può essere utile combinare i risultati LR. Per poter combinare i risultati verrà una moltiplicazione, ma è necessario che i database per ciascun sistema tipizzato rappresentino adeguatamente la stessa

popolazione e che le ipotesi formulate siano le stesse, e si tenga conto che i parenti materni non possono essere distinti dal mtDNA. ⁽⁶⁾

Le linee guida per l'interpretazione delle prove del mtDNA sono state, fornite da ISFG e SWGDAM, e sono le seguenti:

-ESCLUSIONE: se i campioni risultano essere differenti per due o più posizioni nucleotidiche (escludendo le eteroplasmie di lunghezza), può essere escluso che essi derivino dallo stesso soggetto o appartengano alla stessa linea materna.

-INCLUSIONE: il confronto viene riportato come inconclusivo se i campioni differiscono solo per una singola posizione (indipendentemente, che condividano o meno una eteroplasmia di lunghezza).

NON ESCLUSIONE: se i campioni hanno la stessa sequenza, o sono concordanti (condividono una base comune ad ogni posizione nucleotidica), non si può escludere che essi derivino dallo stesso soggetto o appartengano alla stessa linea materna.

Il motivo di queste linee guida interpretative per l'analisi del mtDNA rispetto ad altri marcatori del DNA sono dovuti all'elevato tasso di mutazione della molecola e al collo di bottiglia rilevato soprattutto nei capelli.⁽⁷⁾

1.5 ETEROPLASMIA.

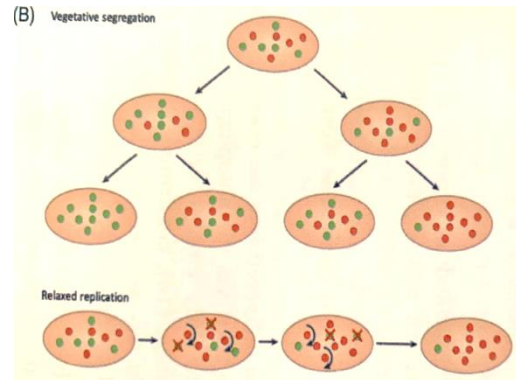
Con il termine eteroplasmia⁽⁸⁾ si intende la coesistenza di più di una variante del genoma mitocondriale (mtDNA) all'interno di un individuo. L'eteroplasmia si verifica attraverso mutazioni somatiche durante la vita di un individuo, e può avere conseguenze evolutive importanti sulla sopravvivenza del mtDNA dall'accumulo di mutazioni deleterie e per la sua coevoluzione con il genoma nucleare. I processi principali che portano all'eteroplasmia sono due e sono stati discussi da Stewart e altri, essi sono:

-Segregazione vegetativa: si verifica durante la divisione in cellule mitotiche, dove l'mtDNA presente nella cellula madre passa alle cellule figlie, se è presente una variazione eteroplasmica la cellula figlia riceverà una porzione della cellula mutata e questo può comportare dei cambiamenti di eteroplasmia molto significativi.

-Replicazione rilassata: in questo caso l'mtDNA viene distrutto e replicato in continuo senza seguire un meccanismo specifico, ciò non ci permette di assicurarci che venga replicato un numero corretto di copie.

La segregazione vegetativa e la replicazione rilassata non sono esclusive, probabilmente si combinano per influenzare le dinamiche della segregazione dell'eteroplasmia in vivo.

Fig.6 Rappresentazione due processi che portano all'eteroplasmia



Nei tessuti ci possono essere vari livelli di eteroplasmia quindi per definire correttamente le mutazioni presenti nel tessuto è necessario raccogliere più campioni di uno stesso individuo. Gli studi si sono concentrati sull'analisi del sangue, delle cellule buccali e capelli, in quanto questi rappresentano un tipo di tessuto/fluido maggiormente riscontrato tra le tracce forensi. L'mtDNA mutante possono essere ereditati dalla madre o insorgere de novo, vari studi hanno dimostrato che i livelli di mtDNA mutante possono insorgere in maniera non casuale tra le generazioni e tra i tessuti di un individuo.⁽⁷⁾ In genetica forense, distinguiamo due tipi di eteroplasmia: *-eteroplasmia di lunghezza*, e *eteroplasmia puntiforme*. Il tipo di eteroplasmia maggiormente considerato in ambito forense è l'eteroplasmia puntiforme, e ci permette di avere una definizione dell'aplogruppo ovvero un insieme di aplotipi tra loro differenti, tutti però originati dallo stesso aplotipo ancestrale. Inizialmente si pensava che

la maggior parte degli individui era omoplasmica, ⁽⁹⁾ ovvero era presente un solo tipo di mtDNA, ciò venne smentito grazie a nuove tecnologie che hanno permesso di evidenziare svariati casi di eteroplasmia nell'uomo. Un primo caso in cui venne riscontrata l'eteroplasmia è stato analizzando mtDNA dei capelli strappati da un donatore anonimo riconducibile presumibilmente allo zar Nicola II di Russia e suo fratello, in posizione 16.169 della regione di controllo dell'mtDNA

1.6 ETEROPLASMIA E MALATTIA

L'eteroplasmia mitocondriale rappresenta una co-espressione dinamicamente determinata da polimorfismi ereditari e patologie somatiche in rapporti variabili all'interno dei singoli genomi del DNA mitocondriale (mtDNA) con modelli ripetitivi di specificità tissutale. I rapporti dei genomi mtDNA rappresentano un equilibrio tra esiti cellulari sani e patologici. Possiamo vedere nello specifico le cardiomiopatie le quali hanno profonde alterazioni della funzione mitocondriale. Alcuni squilibri allelici nel genoma mitocondriale nucleare sono associati a proteine mitocondriali. L'eteroplasmia mitocondriale può manifestarsi nei punti critici di espressione proteica, ad es. citocromo c ossidasi (COX). Anche le mutazioni patologiche del mtDNA sono associate allo sviluppo di insufficienza cardiaca congestizia. È interessante notare che nelle famiglie possono verificarsi rapporti mitocondriali "normali vs. anormali"

di varie popolazioni eteroplasmiche. Nel contesto traslazionale della salute e della malattia umana, è importante la necessità di determinare i focolai critici per sondare molteplici ruoli biologici dell'eteroplasma mitocondriale nella cardiomiopatia.⁽¹⁰⁾

1.7 ETEROPLASMA MITOCONDRIALE RILEVATA NEI CAPELLI

L'eteroplasma viene riscontrata maggiormente nei campioni di capelli, causata dai fenomeni di deriva genetica e dei colli di bottiglia che si verificano a causa della natura semiclonale del follicolo pilifero.⁽¹¹⁾ I capelli hanno origini ectodermiche⁽¹²⁾, con un certo contributo dello strato germinale mesodermico alla papilla dermica. A circa 9-12 settimane di gestazione, un gruppo di cellule dell'ectoderma inizia a crescere verso il basso formando un'invaginazione, qui ci sono un gruppo di cellule mesodermiche che successivamente formeranno la papilla dermica e la guaina del follicolo. La papilla dermica ha la funzione di influenzare le cellule della matrice nel bulbo pilifero dirigendo l'attività mitotica. Entro la 16-20 settimana di gestazione, i follicoli piliferi (costituiti dai gruppi cellulari sopra descritti) si stabilizzano e iniziano a produrre i capelli composti da una radice (parte nascosta) e da un fusto (parte esterna). Ogni radice dei capelli è formata da un piccolo gruppo di cellule staminali nel follicolo pilifero. Questo porta a un pool limitato di molecole di mtDNA, cioè

il collo di bottiglia somatico. Le cellule che formano il fusto del capello si differenziano e non si dividono mai più.

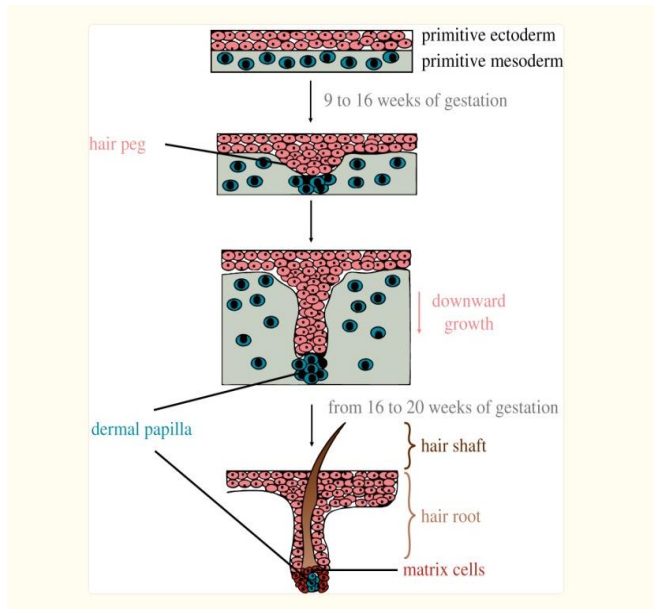


Fig. 7. Struttura capello

L'eteroplasma⁽¹³⁾ è una situazione che si riscontra spesso nella prima e nella seconda regione ipervariabile del mtDNA (HV1/HV2), in particolare nei campioni di capelli. Per indagare su possibili hotspot di eteroplasma, l'HV3 da campioni di capelli e sangue di 100 individui è stato sequenziato e confrontato. Non è stata osservata eteroplasma puntiforme, ma prevalentemente eteroplasma di lunghezza, presente sia in C-stretches che nelle regioni con ripetizione CA. Per osservare quali "alleli" CA erano presenti in ciascun tessuto, i prodotti PCR sono stati clonati e risequenziati. Tuttavia, non è stata osservata alcuna variazione tra gli alleli CA. Per quanto riguarda la pratica

forense, concludiamo che l'eteroplasmia puntiforme in HV3 non è così frequente come in HV1/HV2.

1.8 SEQUENZE MITOCONDRIALI NUCLEARI (NUMT)

Con il termine *NUMT*⁽¹⁴⁾ indichiamo segmenti di mtDNA che si inseriscono nel genoma nucleare, e ciò venne coniato dal genetista evoluzionista Jose V. Lopez et al, che descrisse la trasposizione di qualsiasi tipo di DNA mitocondriale nel genoma nucleare degli organismi eucarioti. Le *NUMT* vennero scoperte involontariamente da ricercatori i quali studiavano l' mtDNA, le *NUMT* differiscono per numero e dimensioni tra le diverse specie e ciò può essere spiegato dalla variazione interspecifica in tali fattori come: Stabilità della linea germinale e numero di mitocondri. Dopo il rilascio del mtDNA nel citoplasma, l' mtDNA viene trasferito nel nucleo provocando una rottura del doppio filamento. Non è stata trovata alcuna correlazione tra la frazione di DNA non codificante e l'abbondanza di *NUMT* nel genoma, ma è stato dimostrato che i *NUMT* non hanno una distribuzione casuale, hanno maggiori probabilità di trovarsi in una determinata zona del genoma rispetto ad altri. A seconda della posizione i *NUMT* possono perturbare la posizione dei geni. L'integrazione de novo degli pseudogeni *NUMT* nel genoma ha un effetto negativo provocando vari disturbi e l'invecchiamento. Il meccanismo con cui i *NUMT* entrano nel genoma nucleare non è del tutto chiaro, 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), ma studi sul lievito hanno dimostrato che i *NUMT* si integrano attraverso un meccanismo di giunzione delle estremità non omologa durante la riparazione della rottura del doppio filamento. Questo stesso meccanismo è stato indicato da Jensen-Seaman e colleghi sulla base di modelli di inserimento NUMT nei genomi umani e di scimpanzé. Con l'emergere di tecnologie basate su MPS, sono stati identificati più di 1077 NUMT umani, inoltre è stata sequenziata un'ampia varietà di genomi di specie, dimostrando che quasi in ogni genoma di qualsiasi specie sono presenti le *NUMT*.⁽¹⁵⁾ Balciuniene e Balciunas hanno ipotizzato che i *NUMT* possano anche essere presenti come inserti mitogenomici completi multicopia, un fenomeno che gli autori hanno chiamato mega-*NUMT*, costituiti da circa 39 bp,. Gli autori hanno confermato la presenza del mega-*NUMT* applicando gli approcci di sequenziamento del mtDNA di Sanger e MPS su DNA estratto da vari tessuti (ad es. sangue, saliva) e cloni sviluppati da questi tessuti. I mega-*NUMT* hanno avuto un particolare interesse per la genetica forense, si è visto infatti che questi inserti possono portare a interpretazioni sbagliate, come ad esempio la trasmissibilità mitocondriale paterna. Anche la trasposizione dei *NUMT* nel genoma è stata associata a malattie umane.

L'integrazione de novo degli pseudogeni *NUMT* nel genoma nucleare ha un effetto negativo in alcuni casi, promuovendo vari disturbi, o può provocare l'invecchiamento, o in alcuni casi quando l'integrazione avviene a livello dei geni codificanti nelle cellule germinali, questo può essere letale per lo sviluppo dell'embrione.

1.9 ETEROPLASMA RILEVATA CON MPS

Negli ultimi anni, la maggior parte dei laboratori di genetica forense ha iniziato ad utilizzare un'innovativa tecnica di sequenziamento ovvero il Massive Parallel Sequencing (MPS), il quale ci consente di rilevare marcatori genetici forensi, tra cui short tandem repeat (STR), single polymorphism nucleotide (SNP), la regione di controllo o l'intero genoma del DNA mitocondriale (mtDNA), in un solo giorno. L'utilizzo di MPS ci permette di avere informazioni altamente attendibili e accurate grazie alla sua capacità di lettura a un threshold più basso, rispetto a Sanger. Questa peculiarità rende l'MPS più sensibile nel rilevare l'eteroplasmia, consentendo quindi di identificare una maggiore quantità di varianti eteroplasmiche all'interno di un campione, rispetto al sequenziamento di Sanger. Questa caratteristica ha un risvolto importante nella pratica forense in quanto va ad aumentare il potere discriminatorio nei campioni che condividono lo stesso aplotipo.

E' stato dimostrato che la tecnologia di sequenziamento Massivo Parallel Sequencing ci permette di avere molte informazioni in un unico esperimento, grazie al suo elevato throughput, sia in termini di regioni genomiche analizzate che in termini di numero di campioni sequenziati. Mediante l'MPS ci consente di sequenziare l'intero genoma mitocondriale in molti campioni contemporaneamente, anche a partire campioni forensi danneggiati.

1.11 DATABASE MTDNA

Quando ci troviamo di fronte a due sequenze di mtDNA una proveniente da un campione di prova e l'altra da quello di riferimento, non bisogna escludere la probabilità che abbiano la stessa origine, per capirlo bisogna contare quante volte una sequenza specifica viene osservata all'interno di un database di popolazione. Il più importante database di aplotipi del mtDNA è l'EDNAP Mitochondrial DNA Population Database. EMPOP è stato progettato per fungere da database della popolazione di riferimento, da utilizzare nella valutazione delle prove del mtDNA in tutto il mondo, con l'obiettivo di fornire dati del mtDNA della massima qualità. EMPOP funge anche da strumento di controllo della qualità per gli scienziati di genetica forense e di altre discipline. EMPOP utilizza SAM, un algoritmo di ricerca basato su stringhe che converte sequenze di query e database in stringhe di nucleotidi senza allineamento e ci garantisce che un aplotipo venga trovato in una query di database

indipendentemente dal suo allineamento. Ad oggi viene utilizzato SAM-E, una versione aggiornata di SAM che considera il blocco InDels come eventi filogenetici. Il browser degli aplogruppi dello strumento rappresenta tutti gli aplogruppi Phylotree stabiliti in un comodo formato ricercabile e fornisce il numero di sequenze EMPOP assegnate ai rispettivi aplogruppi stimando gli aplogruppi del DNA mitocondriale utilizzando l'approccio di massima verosimiglianza EMMA. PhyloTree fornisce una filogenesi completa aggiornata della variazione globale del mtDNA umano, basata sia sulle mutazioni della regione di codifica che di controllo.⁽¹⁹⁾ Un altro importante database del mtDNA umano è Mitomap, contenente la variazione del mtDNA umano pubblicata insieme a varianti geografiche e specifiche per malattia. Mitomap memorizza l'elenco annotato delle varianti del mtDNA sia da individui sani che da pazienti. Le frequenze delle varianti sono calcolate dai mitogenomi umani recuperati dalla GenBank.⁽²⁰⁾

SCOPO DELLO STUDIO

Ad oggi l'importanza del mtDNA si conferma maggiormente, assumendo un ruolo fondamentale nel campo della genetica forense. Di pari passo a ciò sono state sviluppate nuove ed efficienti metodiche analisi, sequenziamento e interpretazione dei dati, le quali ci hanno permesso di rilevare più facilmente l'mtDNA da campioni come sangue o tamponi buccali. Si è riscontrato avere maggiore difficoltà nel rilevare mtDNA da campioni di capelli, in quanto ci sono molte varianti eteroplasmatiche. Per la rilevazione e interpretazione delle eteroplasmie nei capelli e nei campioni di sangue e tamponi buccali si è utilizzato il Massive Parallel Sequencing. Lo scopo del mio progetto di tesi è proprio di rilevare più varianti eteroplasmatiche possibili di mtDNA nei capelli. Successivamente i risultati, saranno analizzati accuratamente, per arrivare a risultati soddisfacenti per la genetica forense. Il mio progetto di tesi si concentra prevalentemente sulla rilevazione di eteroplasmia dell' mtDNA su campioni di capelli. Per fare si è proceduto con la raccolta di campioni di sangue, campioni buccali e capelli, estrazione del DNA, quantificazione e sequenziamento. Le varianti eteroplasmatiche riscontrate nei capelli sono state confrontate con quelle rilevate nei campioni di sangue e tamponi buccali, utilizzati come riferimenti. Uno degli obiettivi dello studio è stato quello di verificare se le varianti eteroplasmiche osservate fossero varianti vere oppure causate dalla

concomitante amplificazione e sequenziamento di specifiche regioni NUMT. Tale obiettivo è stato perseguito mediante un'analisi combinata di diversi strumenti bioinformatici che verranno ampiamente descritti nei capitoli successivi.

CAPITOLO 2

MATERIALI E METODI

2.1 RACCOLTA CAMPIONI

I campioni analizzati in questo studio sono stati raccolti da 5 diversi volontari. Per ciascun volontario è stato prelevato 1 campione di sangue e 1 tampone buccale, utilizzati come campioni di riferimento, e 3 fusti di capelli. I campioni sono stati codificati con una lettera specifica a seconda del tipo di tessuto analizzato ovvero “S” per i tamponi buccali, “B” per il sangue e infine “H” per i fusti di capelli.

La raccolta del tampone buccale è stata svolta utilizzando tamponi sterili. (CITOSWAB - Collection swabs, LABOINDUSTRIA), seguendo la procedura di seguito indicata. Dopo aver codificato ed irradiato le provette con luce UV per 5 minuti, si è proceduto al prelievo delle cellule della mucosa buccale mediante tamponatura interna della guancia evitando il contatto con superfici contaminate come lingua e denti. Al termine il campione prelevato è stato conservato a -20°C, per evitare la formazione di muffe che potrebbero avere un effetto deleterio sul materiale. La raccolta del campione di sangue è stata svolta, da personale addestrato specializzato adoperando tutte le opportune misure di

sicurezza. Anch'esso è stato conservato a 20°C in provette appositamente codificate. La raccolta dei campioni di capelli è stata svolta come indicato qui di seguito. Dopo aver preparato cinque piastre Petri, opportunamente codificate e sterilizzate con luce UV è stato chiesto a ciascun volontario di pettinarsi i capelli con le dita per raccogliere i capelli. I capelli raccolti sono stati trasferiti nelle piastre petri e conservati a temperatura ambiente prima dell'elaborazione. Successivamente, per ciascun individuo, si è proceduto alla selezione di 3 capelli in cui è visibile il bulbo pilifero, per consentire la distinzione tra la radice e la punta del fusto. Ogni capello è stato misurato per intero ed informazioni, la lunghezza, colore, spessore, e tutte le caratteristiche peculiari sono state registrate.

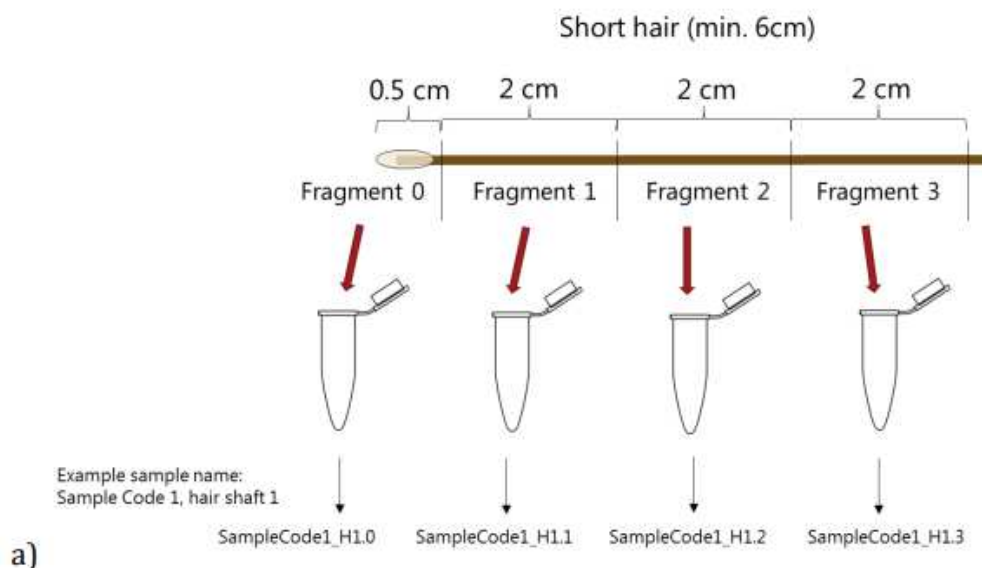


Fig.8. Step preparazione e identificazione del campione di capello corto

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.

Ogni fusto di capelli deve essere diviso in frammenti di 2 cm lungo la sua lunghezza. La sezione della radice va asportata a circa 0,5 cm dall'estremità prossimale dei capelli e trasferita in una provetta (etichettata SampleCode_HX.0) precedentemente codificata e sterilizzata. 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, mentre per quelli più lunghi (>6cm) i frammenti saranno selezionati dopo aver esaminato l'intera lunghezza del capello.

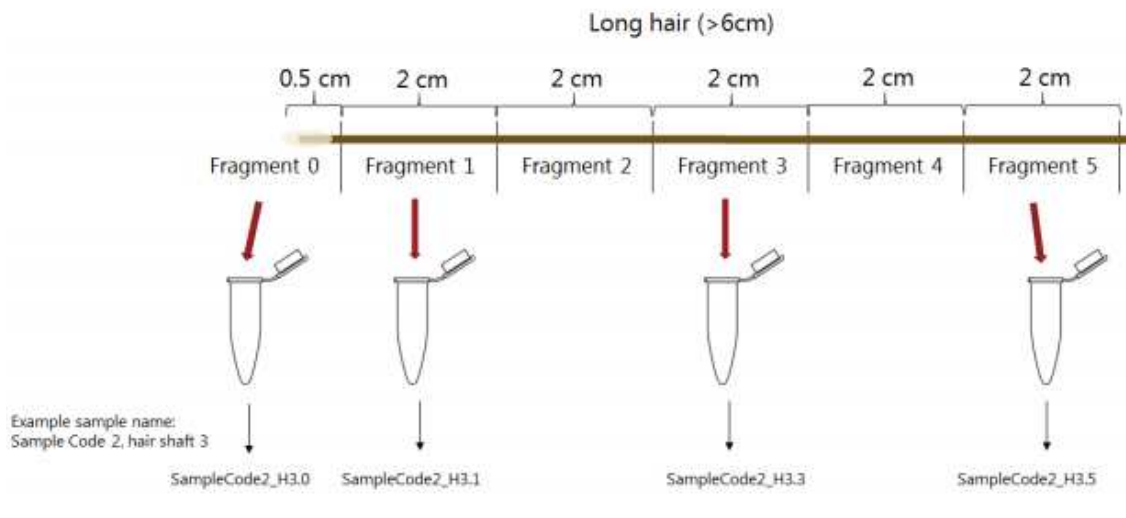


Fig. 9 Preparazione e identificazione del campione di capello lungo

2.2 ESTRAZIONE

L'estrazione è stata svolta usando il Maxwell FSC DNA IQ Casework Kit (Promega) mediante lo strumento Maxwell 16 (Promega).

2.2.3 ESTRAZIONE DA CAMPIONI DI SANGUE

Per l'estrazione di campioni di sangue è stato utilizzato Per l'estrazione dai campioni di sangue è stato utilizzato il kit QIAamp® DNA Mini kit (Qiagen). I passaggi della procedura sono riportati di seguito:

- 1 Pipettare 20 µl QIAGEN Protease (or proteinase K) nella provetta da 1.5 ml.
- 2 Aggiungere 200 µl di campione nella provetta per la microcentrifuga

3. Aggiungere 200 μ l di tampone al campione. Vortexare per 15 s, per garantire una lisi efficiente, è essenziale che il campione e il tampone vengano miscelati accuratamente per ottenere una soluzione omogenea
- 4 . Incubare a 56 °C per 10 min.
5. Centrifugare brevemente la provetta da 1.5 ml per rimuovere le gocce dall'interno del coperchio
6. Aggiungere 200 ml di etanolo (96-100%) al campione e mescolare nuovamente mediante vortexare per 15 s. Dopo la miscelazione, centrifugare brevemente la provetta da 1.5 ml per rimuovere le gocce dal coperchio.
7. Applicare con attenzione la miscela del passaggio 6 alla colonnina spin QIAamp Mini in una provetta da 2 ml senza bagnare il bordo. Chiudere il tappo e centrifugare a 8000 giri/min per 1 minuto. Posizionare la colonnina QIAamp Mini in una provetta pulita da 2 ml ed eliminare la provetta contenente il filtrato.
8. Aprire con attenzione la colonnina QIAamp Mini e aggiungere 500 μ l di tampone AW1 senza bagnare il bordo. Chiudere il tappo e centrifugare a 8000 giri per 1 minuto. Posizionare la colonnina QIAamp Mini in una provetta pulita da 2 ml ed eliminare la provetta contenente il filtrato.

9. Aprire con attenzione la colonnina QIAamp Mini e aggiungere 500 µl di tampone AW2 senza bagnare il bordo. Chiudere il tappo e centrifugare alla massima velocità (14.000 giri/min) per 3 minuti

10. Posizionare la colonnina QIAamp Mini in una nuova provetta da 2 ml ed eliminare la precedente provetta con il filtrato. Centrifugare alla massima velocità per 1 minuto.

11. Posizionare la colonnina QIAamp Mini in una provetta per microcentrifuga pulita da 1.5 ml ed eliminare la provetta contenente il filtrato. Aprire con attenzione la colonnina QIAamp Mini e aggiungere 100 µl di tampone AE o distillato acqua. Incubare a temperatura ambiente (15–25°C) per 5 minuti, quindi centrifugare a 8.000 giri/min per 1 minuto.

2.2.3 ESTRAZIONE DA TAMPONI BUCCALI

Per l'estrazione dai tamponi buccali è stato utilizzato il protocollo di estrazione Solid Support con DNA IQ™ Spin Basket (Promega), riportato qui di seguito:

Tab. 1. Miscela di estrazione

Reagenti	Volume (μl)
Buffer di estrazione casework	386 μ l
Proteinasi K	10 μ l
1-tioglicerolo	4 μ l
Volume totale	400 μ l

- 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 (Tab.1):

Tab. 1. Miscela di estrazione

- Posizionare il tampone sul fondo di una provetta da 2 ml opportunamente codificata e aggiungere la miscela di estrazione.

- Chiudere il coperchio della provetta, vortexare il campione ad alta velocità per 5 secondi. I campioni sono stati suddivisi in due gruppi in base al tempo di incubazione, incubando il primo gruppo per un'ora, il secondo overnight.

- Centrifugare il campione alla massima velocità per rimuoverne eventuali gocce presenti sul coperchio.

- Trasferire tampone buccale in un DNA IQ™ Spin Basket che verrà posizionato nella provetta contenete il lisato.

- Centrifugare alla massima velocità per 2 minuti.
- Rimuovere il DNA IQ™ Spin Basket e aggiungere 200 µl di Lysis Buffer
- Chiudere il coperchio della provetta e vortexare il campione per 5-10 secondi.
- Posizionare le cartucce Maxwell 16 LEV unitamente alle elution tube negli appositi alloggiamenti del supporto.
- Aggiungere 50 µl di tampone di elution buffer in ciascuna elution tube.
- Aprire le 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).

2.2.5 ESTRAZIONE DA CAPELLI

Prima dell'estrazione del DNA, i fusti dei capelli sono decontaminati utilizzando una soluzione di etanolo al 70% per 5 min, e successivamente sottoposti a tre risciacqui con acqua ultrapura. Tra i vari passaggi vortexiamo il campione a velocità moderata, così che avremo il distacco di cellule adese. Infine, i campioni sono tagliati in porzioni da 0.5 cm ciascuno, in modo che, durante l'estrazione, tutti i frammenti siano coperti dai reagenti. Procedere alla conservazione dei campioni a -20°C nel caso in cui non si proceda direttamente all'estrazione. Il processo di estrazione è semiautomatico ovvero ad una prima fase di lisi manuale segue il processamento dei campioni tramite lo strumento Maxwell FSC DNA IQ™ secondo il protocollo DNA IQ™ Casework Pro Kit for Maxwell® 16.

Per ogni campione aggiungere 386 µl di Casework extraction buffer, 10 µl di Proteinasi K e 4 µl di 1-Tioglicerolo, vortexare alla massima velocità per mescolare bene;

1. Incubare il campione sul thermomixer a 56°C in agitazione per 1 ora;
2. Aggiungere 200 µl di Lysis buffer +100 µl di DTT;

3. Posizionare le Maxwell16 LEV cartridges sul supporto, e aggiungere il campione nel pozzetto 1 e posizionare un plunger nel pozzetto 8;

4. Posizionare un elution tube in cui sono stati aliquotati 50 µl di Elution Buffer;

5. Porre sullo strumento e seguire le istruzioni.

Al termine del processo verificare la presenza di capelli nel pozzetto n.1, un'ottima estrazione si ottiene quando nessun capello è presente nel pozzetto.

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

2.3 QUANTIFICAZIONE

2.3.1 FUSTI DI CAPELLI

Per quantizzare l'mtDNA estratto dai fusti di capelli è stato utilizzato un saggio multiplex Real Time PCR (qPCR TaqMan), in grado di valutare simultaneamente sia il DNA mitocondriale che il DNA nucleare umano²³. Il saggio combina due target mtDNA con prodotti di amplificazione di dimensioni diverse (mtND1=69bp e mt143=143bp) utili anche alla valutazione dello stato di degradazione del mtDNA, un target per il nDNA (nuRNU2_181 bp) e un controllo positivo interno (IPC) per la rilevazione di potenziali inibitori della PCR.

Tutte le reazioni sono state eseguite in un volume finale di 20 μ l, contenente 5 μ l di 4x QuantiNova Multiplex PCR Master Mix (Qiagen, QuantiNova Multiplex PCR kit Handbook 06/2016), 2 μ l di campione, la rispettiva concentrazione di primer e sonda (Tabella 2) e acqua nuclease-free a raggiungimento del volume finale. L'amplificazione e il rilevamento della fluorescenza sono stati eseguiti utilizzando Rotor Gene Corbett 6000 (Qiagen) con le seguenti condizioni PCR: 2 minuti a 95 ° C per la denaturazione e l'attivazione enzimatica e 40 cicli a 95 ° C per 3 sec e 60 ° C per 30 sec. Per la costruzione della curva standard è stato utilizzato lo standard DNA - G1521 (Promega). La preparazione della curva e il numero delle copie di mtGE presenti all'interno delle diluizioni seriali è stato eseguito come descritto in precedenza²². Brevemente, per il target nucleare, è stata costruita una curva standard di otto punti a partire da una concentrazione di 62.5ng/ μ l fino a 0.0008ng/ μ l del DNA standard G1521 (Promega). Per il genoma mitocondriale, una breve sequenza a doppio filamento di mtDNA è stata usata in cinque concentrazioni di diluizione seriale per stimare il numero di equivalenti del genoma mitocondriale (mtGE) presenti nello standard DNA-G1521.

	Conc. Iniziale (μM)	Conc. finale (nM)	Volume (μl)
4x Quantinova Multiplex PCR MM			5
IPC- primer Fw	4	100	1

IPC- primer Rv	4	100	1
IPC-probe	4	100	1
IPC-oligo	600.000	600.000	1
nuRNU2_181bp- primer Fw	20	900	0.9
nuRNU2_181bp- primer Rv	20	900	0.9
nuRNU2_181bp- probe	10	500	1
mtND1-primer Fw	6.25	312.5	1
mtND1-primer Rv	6.25	312.5	1
mtND1-probe	2	100	1
mt143bp- primer Fw	12.5	312.5	0.5
mt143bp- primer Rv	12.5	312.5	0.5
mt143bp- primer probe	4	100	0.5
Standard o campione			2
H ₂ O			3.2

Tab.2 Quantivona Multiplex PCR

2.3.2 QUANTIFICAZIONE QUBIT™ per campioni di riferimento

I campioni di sangue e tamponi buccali sono stati quantizzati mediante l'utilizzo del fluorimetro Qubit ® (ThermoFisher Scientific), che utilizza coloranti fluorescenti per determinare la concentrazione di acidi nucleici o proteine in un campione. Nello studio è stato utilizzato il dsDNA HS Assay Kits sviluppato e distribuito da Thermo Fisher Scientific. Questo kit è specifico e selettivo per DNA a doppia elica ed è indicato per campioni concentrati tra 10 pg/µl e 100 ng/µl. Il protocollo di "Qubit dsDNA HS Assay Kits" (MAN0002326 Rev. C.0), prevede la preparazione degli standard per la

creazione della curva di calibrazione e la preparazione dei campioni come descritto qui di seguito:

- Per gli standard aggiungere 1908 μL di working solution Qubit™ dsDNA a ciascuna provetta e aggiungere 10 μL di ciascun campione da analizzare all'interno di una provetta

-Per i campioni aggiungere 198 μL di working solution Qubit™ dsDNA a ciascuna provetta e

-Aggiungere 2 μL di ciascun campione da analizzare all'interno di una provetta

-Mescolare vigorosamente ogni campione vortexando per 3-5 secondi.

-Lasciare incubare tutte le provette a temperatura ambiente per 2 minuti

-Inserire la provetta all'interno del Qubit™, chiedere il coperchio e procedere con la lettura dei campioni

2.4. PREPARAZIONE MANUALE DELLA LIBRERIE

Per valutare la presenza di varianti eteroplasmiche nei campioni presi in esame è stato usato 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 utilizzato solo per le applicazioni forensi. Il pannello per il genoma del mtDNA è formato da

162 ampliconi composto da due pool multiplex (con 81 coppie di primer ciascuno) mirato all'intero genoma mitocondriale umano (16.569 bp).

Il templating delle librerie costruite è stato eseguito mediante Ion Chef™ System (Thermo Fisher Scientific) e il sequenziamento con lo strumento Ion Gene Studio S5™ System (Thermo Fisher Scientific).

2.4.1 REAZIONE DI AMPLIFICAZIONE DEL mtDNA TARGET

Il Precision ID library kit (Thermo Fisher Scientific) comprende i reagenti per la preparazione delle librerie utili al Precision ID mtDNA Whole Genome Panel e necessita di un minimo di 0.1 ng di DNA genomico, corrispondente a 2.900 copie del mtDNA per la reazione di amplificazione del target. Dopo la 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). Il metodo 2-in-1, utilizzato per i frammenti di capelli, prevede la preparazione di due mix per ogni pool di primer, come riportato qui di seguito:

-Preparare la master mix per il pannello Precision ID mtDNA Pool 1, per poi aggiungerla ad ogni provetta di reazione (Tab.3):

Reagente	Volume
HIFI 5X Ion AmpliSeq™	4µl
Precision ID WGP mtDNA panel, Pool1	10µl
gDNA 0.1ng	XµL
Nulease free water	6-X µL
Totale	20µL

Tab 3. mix Pool 1 metodo 2 in 1. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

-Preparare la master mix per il pannello Precision ID mtDNA Pool 2 ,per poi aggiungerla ad ogni provetta di reazione (Tab.4):

Reagente	Volume
Precision ID WGP mtDNA panel, Pool2	10µl
gDNA 0.1ng	XµL
Nulease free water	6-X µL
Totale	20µL

Tab 4. mix per pool 2 metodo 2 in 1. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

-Vortexare le provette di reazione, poi centrifugare per raccogliere le goccioline.

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

Il metodo che viene utilizzato per i campioni di sangue e tamponi buccali è il metodo conservativo. Per l'applicazione bisogna preparare due mix, uno per ogni pool di primer, come riportato qui di seguito (Tab. 5 e Tab. 6):

Reagente	Volume
HIFI 5X Ion AmpliSeq™	2µl
Precision ID WGP mtDNA panel, Pool1	5µl
gDNA 0.1ng	XµL
Nulease free water	3-X µL
Totale	10µL

Tab 5. Mix Pool 1 metodo semiconservativo. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

Reagente	Volume
HIFI 5X Ion AmpliSeq™	2µl
Precision ID WGP mtDNA panel, Pool2	5µl
gDNA 0.1ng	XµL
Nulease free water	3-X µL
Totale	10µL

Tab 6.mix Pool 2 metodo semiconservativo. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

Per amplificare bisogna eseguire il profilo termico che segue (Tab.7) sul termociclatore Veriti™ 96-Well Fast Thermal Cyclor (Thermo Fisher Scientific):

Stage	Step	Temperatute	Time
Hold	Activate the enzyme	99° C	2 min
Cycle number	Denature	99° C	15 sec
	Anneal and exted	60° C	4 min
Hold	-	10° C	Hold

Tab. 7. Rappresentazione del profilo termico per reazione di amplificazione

A seconda della differente matrice sottoposta a PCR diversi numeri di cicli sono stati utilizzati come da indicazioni del manuale; ovvero per i frammenti di capelli sono stati utilizzati 26 cicli di amplificazione (low copy number samples) viceversa i nostri riferimenti sono stati sottoposti a 21 cicli di amplificazione (non-degraded samples).

2.4.2 DIGESTIONE PARZIALE DEGLI AMPLICONI

Per la parziale digestione dell'amplicone viene aggiunto 2 µL di FuPa reagent a ciascun campione amplificato ottenendo un volume totale ~ 22 µl. Questo passaggio permette la creazione di estremità in grado di legarsi al barcode e all'adattatore successivamente aggiunti, così da costruire correttamente la libreria finale. Questo passaggio è preferibile svolgerlo su ghiaccio o blocco freddo, e procedere rapidamente all'incubazione. Successivamente le provette

verranno vortexate e centrifugate per raccogliere eventuali goccioline presenti sul tappo. Il profilo termico da utilizzare è il seguente (Tab.8):

Temperatura	Tempo
50° C	10 minuti
55°C	10 minuti
60°C	20 minuti
10°C	Hold (fino a 1 ora)

Tab 8. Profilo termico per step FuPa. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System.

2.4.3 LEGAME DEI BARCODE E PURIFICAZIONE DELLA LIBRERIA

Le librerie verranno sequenziate sullo stesso chip, quindi è necessario aggiungere dei barcode diversi (Ion Xpress™ Barcode Adapters, Thermo fisher Scientific) per ciascuna libreria. Bisogna fare molta attenzione quando si maneggiano i barcode poiché molto sensibili alle contaminazioni. Viene preparata la reazioni di ligazione come di seguito riportato (Tab.9):

Reagente	Volume
Soluzione Switch	4 µl
Ion Xpress Barcode	2 µl
DNA ligasi	2 µl
Volume totale	-30 µl

Tab 9. Preparazione reazione di legasi. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

Al termine della preparazione vortexare e centrifugare i campioni e caricarli sul termociclatore Veriti utilizzando il profilo termico riportato di seguito (Tab.10):

Temperatura	Tempo
22° C	30 minuti
72°C	10 minuti
10°C	Hold

Tab.10 Profilo termico per svolgere la reazione di ligazione. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

Successivamente, procediamo preparando i reagenti per le fasi di purificazione, mediante l'Agencourt AMPure XP Reagent (Beckman Coulter) riequilibrato a temperatura ambiente, e vortexare per risospendere le sfere magnetica prima di utilizzarlo. Per i successivi passaggi di lavaggio si procederà a preparare etanolo al 70% (230 µl di etanolo assoluto combinato con 100 µl di acqua priva di nucleasi per ogni campione).

Ad ogni campione viene aggiunto 45 µl di Agencourt AMPure XP Reagent spipettando la miscela su è giù per circa 5 volte così da uniformare la resina al campione. Incubare il campioni per 5 min a temperatura ambiente e successivamente trasferire le provette su uno stendino magnetico e incubare per

2 minuti, o fin quando non avremo una soluzione trasparente, e successivamente verrà eliminato il surnatante. Seguono due fasi di lavaggi con 150 µl di etanolo al 70%. Al termine il pellet, costituito dalle sfere magnetiche a cui è adesa la nostra libreria NGS, viene fatto asciugare a temperatura ambiente per circa 2-5 minuti così da far evaporare l'etanolo residuo dal precedente lavaggio, in quanto la presenza di esse potrebbe inibire la reazione di PCR. Successivamente, le provette verranno rimosse dallo stendino e verranno aggiunti 50 µl di low TE al pellet per eluire la libreria finale procedendo ad un vortex ed ad un incubazione per 5 minuti a temperatura ambiente. Le provette verranno poi riposizionate sullo stendino magnetico e lasciate per 2 minuti ad incubare a temperatura ambiente, infine recuperando il surnatante.

2.4.4 QUANTIFICAZIONE LIBRERIA TRAMITE qPCR

Le librerie così costruite sono quantizzate mediante qPCR con il kit di quantificazione Ion Library TaqMan™ (Cat. N. 4468802). Di seguito viene riportato il procedimento utilizzato:

- 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 del campione E. coli DH10B Control Library (concentrato ~68pM) utile alla costruzione della curva standard alle concentrazioni elencate nella tabella seguente (Tab.11):

Standard	Volume	Nucleasi free-water	Concentrazione
1	5µl Control Library	45µl	6.8 pM
2	5µl Std1	45µl	0.68 pM
3	5µl Std 2	45µl	0.068 pM

Tab. 11. Diluizioni seriali E. coli DH10B Control Library per curva standard qPCR. Da manuale Precision ID mtDNA Panels with the HID ION S5™ S5 System

-Preparare la mix di reazione per ciascun campione, per il controllo negativo (TE), e gli standard. 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) o standard o TE, per un volume totale di 20 µl.

-Impostare il profilo termico per la Real-time PCR sul Rotor-Gene 6000 Corbett (Qiagen) corretto.

Al termine della qPCR, per ciascuna libreria (non diluita) quantizzata viene calcolata la concentrazione media usando la formula:

Concentrazione media della libreria non diluita=(qPCR quantità media) x (diluizione della libreria).

Successivamente tutte le librerie sono diluite alla concentrazione di 40pM come richiesto dal manuale per il templating sullo strumento Ion Chef™ System. Le librerie sia diluite che non diluite vanno conservate a una temperatura tra 2° e 8° per un massimo di un mese. Differentemente per conservazione per tempi prolungati vanno messe a temperature comprese tra -30° a -10°.

2.4.5 LIBRERIE PREPARATE MEDIANTE Ion Chef™ SYSTEM

Parte delle librerie dei fusti di capello sono state preparate automaticamente sullo strumento Ion Chef™ System (Thermo Fisher Scientific) mediante Precision ID DL8 Kit (Thermo Fisher Scientific, MAN0017770, C.0).

2.4.6 PREPARAZIONE TEMPLATE SU HID Ion Chef™ SYSTEM

La preparazione del template sullo strumento Chef™ System consiste nell'amplificazione clonale (emulsion PCR, emPCR) delle sequenze target mediante Ion Sphere Particles (ISP) per procedere poi al caricamento automatico sul chip di sequenziamento utilizzato. Per ogni chip Ion 520, un totale di 25 campioni possono essere sequenziati contemporaneamente ed è necessario creare una corsa seguendo le istruzioni ((Precision ID mtDNA Panel

with the HID Ion S5/HID Ion GeneStudio S5 System, MAN0017770, rev. C.0) utilizzando il software Torrent Suite (TSS, v.5.10.1). In totale sono state eseguite 5 corse indipendenti sullo strumento HID Ion GeneStudio S5™ System, tre utilizzando chip 520 e due con chip 530.

2.5 ANALISI DEI RISULTATI

Dopo il termine della reazione di MPS, le sequenze ottenute sono state analizzate con il software Torrent Suite Software (TSS, v.5.10.1) e le reads sono state allineate sulla reference Precision_ID_mtDNA_rCRS. I plugin Coverage Analysis (v. 5.12.0.0) e VariantCaller (v5.12.0.4, TVC) sono stati utilizzati per l'analisi delle reads. Il primo permette di ottenere informazioni relativamente alle mapped reads, on target e i valori di coverage medio per ogni campione; il secondo permette di identificare le varianti presenti nei campioni rispetto alla reference di utilizzo. La soglia della frequenza per valutare la presenza di eteroplasmia era del 10%. La profondità di lettura (coverage) è al 20x, ciò indica che un amplicone deve essere letto almeno 20 volte per poter identificare una variante mitocondriale. L'analisi secondaria è stata svolta mediante il software Integrative Genomics Viewer ⁽²³⁾, ci ha permesso di valutare e confermare le varianti ottenute dal TVC. Ventiquattro librerie sono state analizzate anche con il software CONVERGE seguita da un'ispezione

manuale mediante software Integrative Genomics Viewer. Le reads ottenute vengono prima mappate su PhyloTree e quindi riallineate utilizzando un algoritmo di allineamento Smith-Waterman personalizzato che integra le informazioni PhyloTree ed EMPOP nella funzione di punteggio. Le varianti vengono chiamate con riferimento al rCRS. Inoltre, viene calcolato l'aplogruppo più vicino e le varianti vengono valutate in base alla loro presenza nell'aplogruppo e ad altri parametri generali tra cui frequenza, strand bias e coverage. Per le posizioni eteroplasmiche riscontrate nei nostri campioni è stata eseguita un'analisi approfondita per determinare se tali posizioni fossero PHP vere o 'PHP apparenti' (artefatti di sequenziamento o NUMTs) mediante la valutazione della presenza di strand bias, l'utilizzo di EMPOP¹⁹ e il tool BLAT (<https://genome.ucsc.edu/cgi-bin/hgBlat>) come descritto nel lavoro di Cihlar[Jennifer Churchill Cihlara,b,*, Christina Stroblc, Robert Lagacéd, Melissa Muenzlera,Walther Parsonc,e, Bruce Budowlea,b]. EMPOP è stato utilizzato per fornire informazioni filogenetiche e valutare quanto spesso la variante è stata osservata in precedenza. Il tool BLAT è stato utilizzato per identificare le regioni mitocondriali con alta omologia con il DNA nucleare ovvero per le identificazioni delle NUMTs.

CAPITOLO 3

RISULTATI

3.1 QUANTIFICAZIONE ESTRATTI

Per la quantificazione degli estratti dei riferimenti quali tamponi buccali e sangue è stato utilizzato il kit Qubit dsDNA HS Assay (Thermo Fisher Scientific). I valori sono riportati nella tabella qui di seguito (Tab.12):

Qubit assay	
2_S	1.51
2_B	16.9
4_S	3.09
4_B	22
5_S	2.1
5_B	24.2
7_S	3.74
7_B	14
8_S	6.9
8_B	44.8

Tab.12. Risultati quantificazione dei campioni di riferimento

Per la quantificazione del mtDNA nei 59 fusti di capelli è stato utilizzato un saggio qPCR che permetteva di valutare simultaneamente sia il DNA mitocondriale che il DNA nucleare. Per la valutazione della reazione di qPCR sono stati presi in considerazione i seguenti parametri: linearità (R^2), l'efficienza di amplificazione e la slope. Per l'analisi dei dati è stato utilizzato il software Rotor Gene Q v. 2.1.0 (Qiagen), il Ct (il ciclo soglia) che viene

calcolato automaticamente dal software e applicato a tutti i campioni. Grazie all'uso dello standard Human Genomic DNA - G1521 (Promega), ci ha permesso di costruire la curva di calibrazione a partire da una concentrazione iniziale di 62.5 ng/ μ l fino a 0.004 ng/ μ l. Osservando e valutando la linearità e l'efficienza della qPCR possiamo notare che ci sono dei risultati ottimali, confermando la robustezza del saggio. Nel test è presente IPC, ovvero il controllo positivo interno, che permette la rilevazione di possibili inibitori della PCR. Per tutti i campioni analizzati non sono stati riscontrati inibitori. I valori di quantizzazione del mtDNA per i fusti di capelli sono riportati nella tabella 13. Poiché nel saggio qPCR utilizzato sono presenti due target mitocondriali con differente lunghezza è stato possibile fare una valutazione qualitativa del mtDNA calcolando l'indice di degradazione (sonda corta/sonda lunga) e l'indice di frammentazione (1-sonda lunga/sonda corta).

Campioni	mtND1 (mtGE/ul)	mt143 (mtGE/ul)	DI	% Frammentazione
2_H1.1	168.9	116.2	1.5	31.2%
2_H1.6	109.3	2.4	45.5	97.8%
2_H1.12	73.1	nd	nc	
2_H2.1	617.2	497	1.2	19.5%
2_H2.7	144.1	6.4	22.5	95.6%
2_H2.14	145.9	0.4	364.8	99.7%
2_H3.1	194.5	86.3	2.3	55.6%
2_H3.6	146.2	14.8	9.9	89.9%

2_H3.12	55.9	4.4	12.7	92.1%
2_H4.1	1101.64	558.65	1.97	49.3%
2_H4.7	633.90	265.67	2.39	58.1%
2_H4.14	582.18	187.88	3.10	67.7%
2_H5.1	830.71	445.57	1.86	46.4%
2_H5.5	755.14	266.83	2.83	64.7%
2_H5.9	581.49	14.91	39.00	97.4%
2_H6.1	Non quantizzati			
2_H6.8	Non quantizzati			
2_H6.15	Non quantizzati			
4_H1.1	179.8	65.9	2.7	63.3%
4_H1.6	279.5	20.42	13.7	92.7%
4_H1.12	343.1	16.2	21.2	95.3%
4_H1.19	219.1	4.2	52.2	98.1%
4_H2.1	422.9	113.1	3.7	73.3%
4_H2.6	544.5	190.4	2.9	65.0%
4_H2.11	347.6	43	8.1	87.6%
4_H2.17	248.6	15.1	16.5	93.9%
4_H3.1	566.2	313.5	1.8	44.6%
4_H3.5	97.3	86.4	1.1	11.2%
4_H3.9	405.8	170.9	2.4	57.9%
4_H4.1	Non quantizzati			
4_H4.6	Non quantizzati			
4_H4.11	Non quantizzati			
5_H1.1	749.8	355.5	2.1	52.6%
5_H1.6	252.8	12.5	20.2	95.1%
5_H1.10	252.2	49.4	5.1	80.4%
5_H2.1	547.7	176.8	3.1	67.7%

5_H2.5	653.2	304.6	2.1	53.4%
5_H2.9	323.7	59.8	5.4	81.5%
5_H3.1	1004.9	573.5	1.8	42.9%
5_H3.7	246.6	95.7	2.6	61.2%
5_H3.14	311.7	28.8	10.8	90.8%
7_H1.1	117	32.1	3.6	72.6%
7_H1.2	213.3	132.5	1.6	37.9%
7_H1.3	165.7	7.1	23.3	95.7%
7_H2.1	288.8	98	2.9	66.1%
7_H2.2	190.7	27.5	6.9	85.6%
7_H2.3	227.9	33.3	6.8	85.4%
7_H3.1	158	55.9	2.8	64.6%
7_H3.2	211.5	42.4	5.0	80.0%
7_H3.3	177.3	11	16.1	93.8%
8_H1.1	197.9	153.3	1.3	22.5%
8_H1.10	200.3	1.5	133.5	99.3%
8_H1.19	57.5	2.2	26.1	96.2%
8_H2.1	493.3	298.5	1.7	39.5%
8_H2.6	135.2	6.1	22.2	95.5%
8_H2.12	113.4	2.6	43.6	97.7%
8_H3.1	129.9	31.9	4.1	75.4%
8_H3.7	228.9	212.3	1.1	7.3%
8_H3.13	107.7	4.7	22.9	95.6%
8_H4.1	1587.11	561.86	2.82	64.6%
8_H4.8	815.43	347.22	2.35	57.4%
8_H4.16	711.09	88.23	8.06	87.6%
8_H5.1	949.15	611.99	1.55	35.5%
8_H5.10	641.97	226.35	2.84	64.7%

8_H5.19	485.17	2.36	205.81	99.5%
----------------	--------	------	--------	-------

Tab13. Valori di quantizzazione fusti capelli con qPCR. nd=non determinato, nc=non calcolabile

Come è possibile vedere in tabella 13, per il capello 6 del donatore 2 e il capello 4 del donatore 4 non si è proceduto alla quantizzazione dell'estratto. Tutti i precedenti fusti sono stati amplificati utilizzando il volume massimo di DNA (6ul) per reazione PCR poiché per nessuno è stato possibile ottenere 2900mt copie come richiesto dal manuale. Come possiamo vedere in Tab. 13 alcuni frammenti presentavano concentrazioni di mtDNA maggiori indipendentemente dal donatore. I valori con concentrazione più alta sono corrispondenti al maggior tempo di incubazione (overnight). Infatti un tempo di incubazione più lungo (overnight) ha comportato una migliore digestione dei capelli, permettendo un aumento del mtDNA recuperato; i frammenti incubati per tutta la notte risultavano completamente digeriti e appariva un pellet scuro sul fondo della cartuccia di estrazione. Viceversa, i fusti incubati per 1 h presentavano frammenti ben visibili nella provetta della cartuccia di estrazione che ha comportato il recupero di concentrazioni minori di mtDNA.

Nella figura 10, vengono riportati i valori di quantizzazione dei fusti divisi per la vicinanza al bulbo, misurato sia con la sonda del target mitocondriale corta che con quella lunga.⁽²⁴⁾ Si può notare come la concentrazione di mtDNA diminuisce all'aumentare della distanza del bulbo.

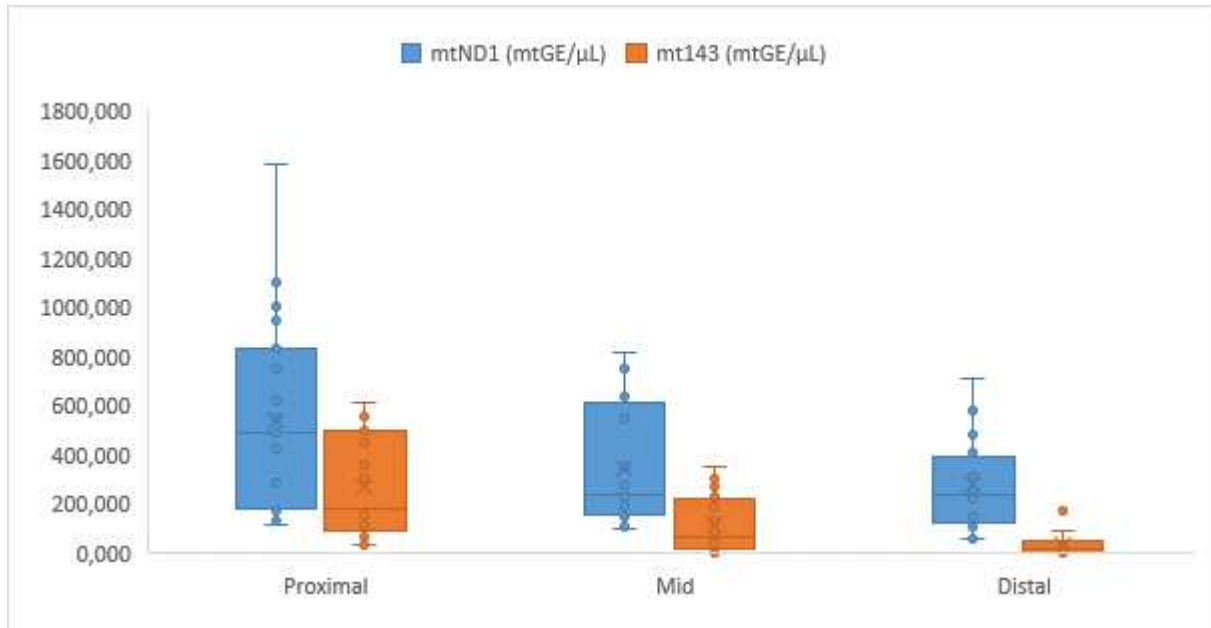


Fig.10. Recupero quantità di mtDNA nei frammenti di capelli incubati durante la notte rispetto a quelli incubati per 1 ora (corto: $p = 7,13E-06$; lungo: $p = 0,009$) e % di frammentazione ($p = 0,717$)

Il saggio utilizzato permetteva anche di valutare la qualità del DNA mitocondriale ovvero la degradazione. Per questa analisi è stato utilizzato come parametro di riferimento l'indice di degradazione (DI) dato dal rapporto tra il target mtDNA corto (69bp) e il target mtDNA lungo (143bp). Quando il valore del DI è 1 o inferiore mtDNA è definito non degradato, mentre un $DI > 1$ indica che negli estratti sono presenti meno frammenti di 143 bp o più grandi. Se ci troviamo di fronte a un DI indeterminato questo ci indica che mtDNA è molto degradato poiché non viene rilevata la sonda lunga (mt143bp). Come è possibile vedere in tabella 13 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 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. I valori di % frammentazione riportati nella tabella 13 sono stati rappresentati nel grafico (fig.11) di seguito riportata.

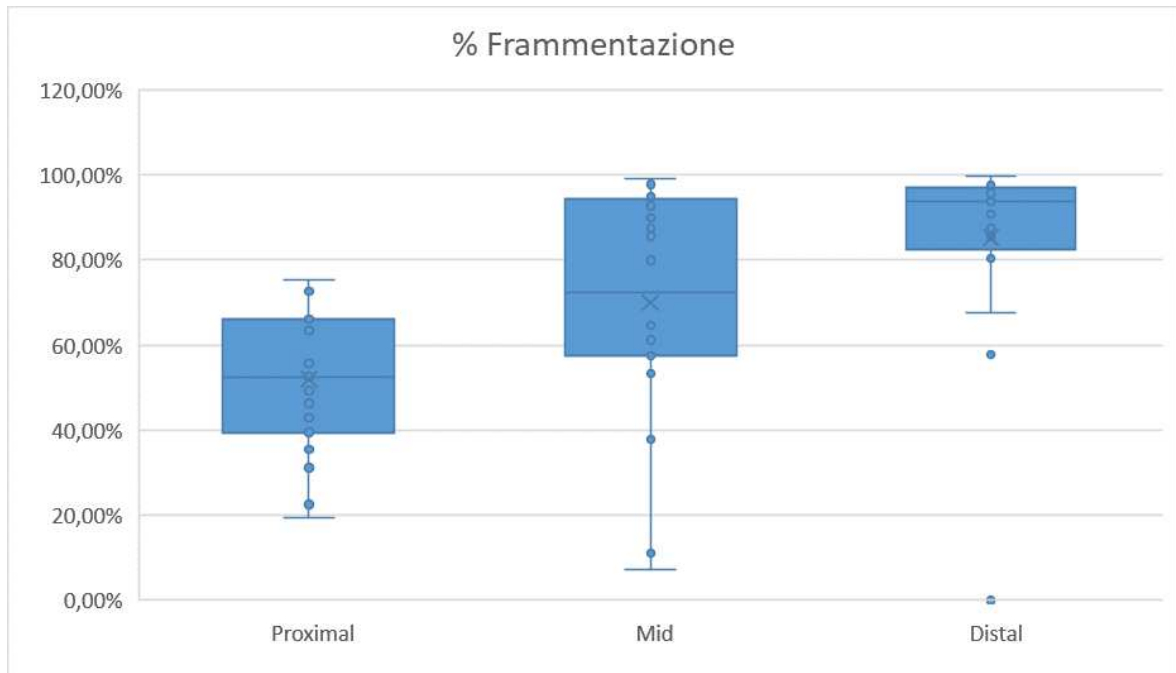


Fig.11. Quantità di mtDNA osservata nei frammenti di capelli a diverse distanze dalla radice (prossimale vs medio e prossimale vs distale: $p < 0,05$; medio vs distale $p > 0,05$) e % di frammentazione ($p < 0,05$ in tutti i confronti).

Come è possibile vedere nel box plot (fig.11) la frammentazione del mtDNA aumenta mano a mano che ci allontaniamo dal bulbo.

Da notare che per tutti i fusti si è ottenuto un valore negativo per il target nDNA presente nel saggio qPCR utilizzato.

3.2 RISULTATI QUANTIFICAZIONE LIBRERIE

Le 55 librerie finali sono state quantificate tramite lo strumento Rotor Gene 6000 Corbett (Qiagen) utilizzando il kit Ion Library TaqMan Quantitation, ciascuna libreria è stata analizzata in duplicato in reazioni da 20 μ l. Terminata la reazione è stata calcolata la media dei valori della quantificazione dei duplicati e successivamente la media ottenuta è stata moltiplicata per 100 data la diluizione effettuata come da manuale 1:100. L'analisi dei dati è riportata nella tabella 14.

Campione	Qpcr1	Qpcr2	pM medio	pM medio*100
2_H4.1	Quantizzazione Ion Chef			
2_H4.7				
2_H4.14				
2_H5.1				
2_H5.5				
2_H5.9				
2_H6.1	14.184	13.598	13.891	1389.133
2_H6.8	8.920	10.239	9.579	957.920
2_H6.15	3.090	3.195	3.143	314.286
2_S	Quantizzazione Ion Chef			
2_B				
4_H2.1	Quantizzazione Ion Chef			
4_H2.6				
4_H2.11				
4_H3.1	23.426	25.304	24.365	2436.50
4_H3.5	6.281	7.072	6.677	667.68
4_H3.9	0.749	0.748	0.749	74.87
4_H4.1	10.686	10.126	10.406	1040.56
4_H4.6	12.327	11.872	12.100	1209.98
4_H4.11	5.119	4.862	4.991	499.06
4_S	17.120	17.183	17.151	1715.11

4_B	11.901	13.213	12.557	1255.70
5_H1.1	29.62	31.00	30.3	3030.9
5_H1.6	1.48	1.37	1.4	142.6
5_H1.10	3.23	3.67	3.4	344.8
5_H2.1	0.55	0.56	0.6	55.8
5_H2.5	15.36	16.13	15.7	1574.6
5_H2.9	2.85	2.80	2.8	282.3
5_H3.1	28.13	28.73	28.4	2843.2
5_H3.7	3.69	5.12	4.4	440.6
5_H3.14	1.80	1.99	1.9	189.7
5_S	27.60	29.56	28.6	2857.9
5_B	13.342	14.172	13.757	1375.68
7_H1.1	4.827	4.928	4.878	487.75
7_H1.2	2.332	2.355	2.343	234.33
7_H1.3	0.433	0.445	0.439	43.90
7_H2.1	6.842	7.835	7.338	733.85
7_H2.2	0.711	0.738	0.725	72.45
7_H2.3	0.339	0.329	0.334	33.37
7_H3.1	2.049	3.018	2.534	253.35
7_H3.2	2.010	2.049	2.030	202.96
7_H3.3	0.856	0.763	0.810	80.95
7_S	14.831	14.518	14.675	1467.45
7_B	7.194	8.142	7.668	766.83
8_H3.1	Quantizzazione Ion Chef			
8_H3.7				
8_H3.13				
8_H4.1				
8_H4.8				
8_H4.16				
8_H5.1				
8_H5.10				
8_H5.19				
8_S				
8_B	2.302	2.532	2.417	241.71

Tab 14. Risultati quantificazione librerie

Seguendo le istruzioni del manuale, ogni libreria è stata diluita a 40 pM aggiungendo low TE per procedere allo step del templating.

3.3 RISULTATI SEQUENZIAMENTO

Le metriche recuperate alla fine della corsa ci permettono di valutare il sequenziamento e la qualità dei dati del mitogenoma. Per le 55 librerie sono state effettuate 5 corse di cui 2 con il Chip Ion 530™ e 3 con il Chip Ion 520™. Le metriche di sequenziamento ottenute dal TSS sono illustrate in Fig. 12. Per praticità vengono mostrati e valutati i parametri solo di una delle cinque corse indipendenti effettuate per il sequenziamento dei campioni, essendo l'interfaccia uguale per tutte.

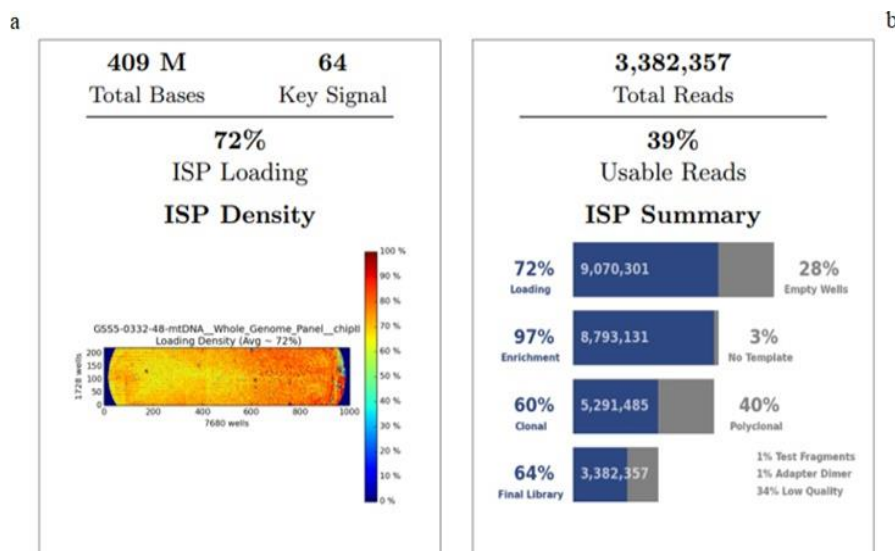


Fig 12. Metriche risultanti per ISP loading (a) e usable reads (b) ottenute.

Nella figura 12 (a) è riportata la percentuale delle ISPs loading sul chip (72% in questo caso). Oltre al valore percentuale riportato, la qualità del caricamento è osservabile anche dalla scala colorimetrica riportata in figura. Le aree in rosso sono quelle con un caricamento ottimale, le aree con densità di caricamento

inferiori hanno una colorazione che oscilla dal giallo al verde, mentre le aree blu indicano un caricamento non riuscito. Il numero di reads totali è riportato in Fig. 12 (b) ed è pari a 3.382.357, di queste il 39% sono indicate come reads utilizzabili ai fini analitici.

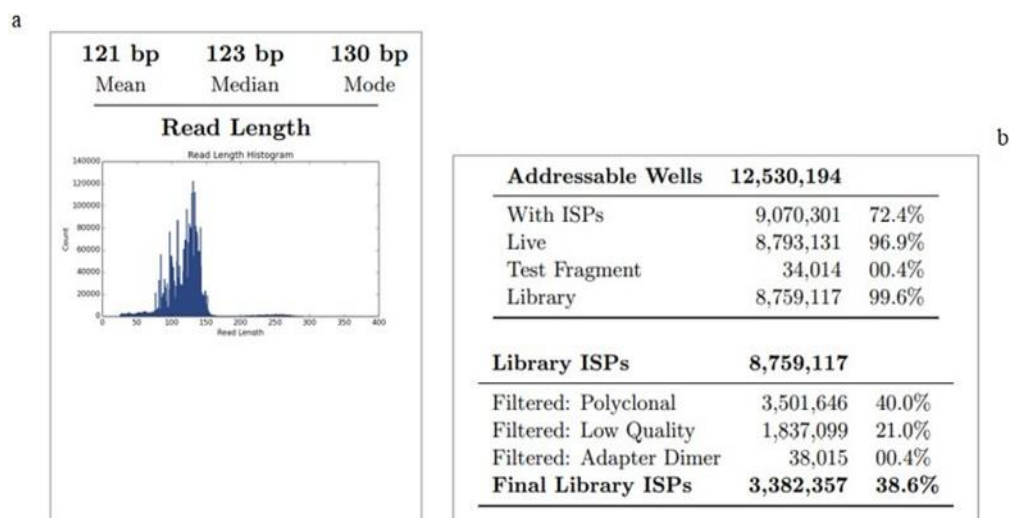


Fig. 13. Metriche risultanti per read lenght (a) e final library ISPs (b) ottenute

L'istogramma in Fig. 13 (a) mostra la lunghezza di tutte le reads presenti nei file di output, inoltre sono riportati i valori di media, mediana e moda relativi alle reads trovate. La Fig. 13 (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. 14).

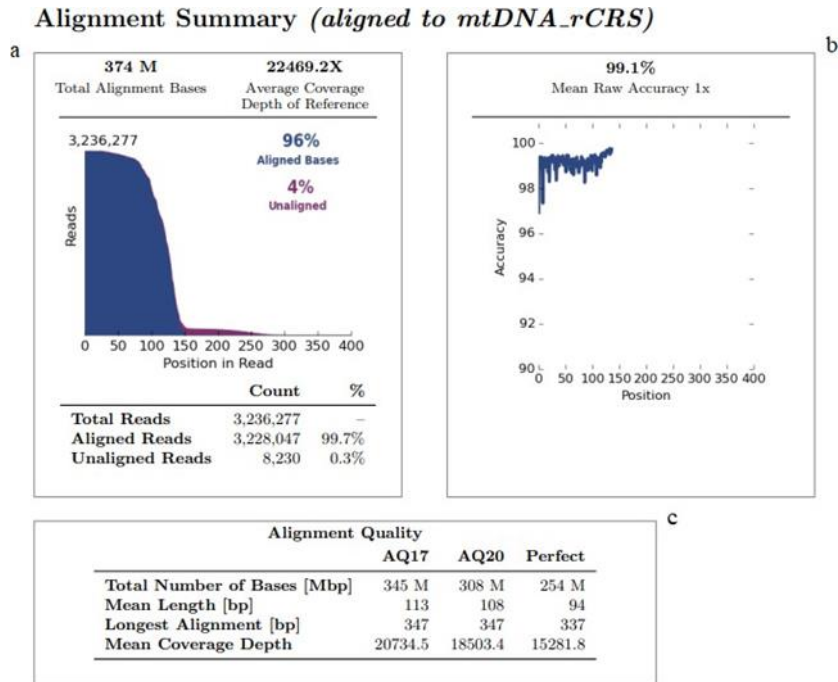


Fig.14 Rappresentazione metriche post allineamento

Le basi totali incluse nell'allineamento sono visualizzate in Fig. 14 (a) dove vengono riportate sia quelle allineate (96%) che quelle non allineate (4%) al riferimento. Nella Fig. 14 (b) viene rappresentata l'accuratezza media calcolata per ciascuna posizione in ogni read. Infine, i parametri di qualità dell'allineamento vengono visualizzati nella Fig. 14 (c). Tutti i valori presenti nelle metriche di corsa superano le soglie imposte. quindi possiamo dedurre che il sequenziamento è riuscito e si è proceduto con l'analisi del coverage

3.4. ANALISI DEL COVERAGE

Il plug-in del coverage analysis genera un Coverage Analysis Report, il report riassume i codici a barre utilizzati, i campioni, il numero di mapped reads, percentuale reads on target, la profondità media del coverage (mean depth) e

l'uniformità di amplificazione del pannello MPS utilizzato. I valori sono riportati in tabella 15.

Campioni	Mapped Reads	On Target	Mean Depth	Uniformità
2_H4.1	536232	100.00%	3300	87.24%
2_H4.7	606487	100.00%	4231	86.40%
2_H4.14	496808	100.00%	2388	86.00%
2_H5.1	1289189	100.00%	9282	90.15%
2_H5.5	953616	100.00%	6654	85.93%
2_H5.9	532976	100.00%	3229	82.38%
2_H6.1	103511	100.00%	755.9	86.93%
2_H6.8	109009	100.00%	777.9	85.55%
2_H6.15	121196	100.00%	857.4	84.22%
2_S	404357	100.00%	2936	87.03%
2_B	1975289	100.00%	14248	88.80%
4_H2.1	970489	100.00%	6947	85.88%
4_H2.6	856241	100.00%	5908	84.56%
4_H2.11	942513	100.00%	6549	86.69%
4_H3.1	101748	100.00%	689.3	73.17%
4_H3.5	128413	100.00%	901.3	98.34%
4_H3.9	126777	100.00%	916.8	88.54%
4_H4.1	112634	100.00%	816.5	86.41%
4_H4.6	64925	100.00%	461.7	73.15%
4_H4.11	101997	100.00%	725.0	85.58%
4_S	446106	100.00%	3247	89.70%
4_B	156301	100.00%	1126	98.97%
5_H1.1	305493	100.00%	2149	88.44%
5_H1.6	208055	100.00%	1449	88.04%
5_H1.10	244738	100.00%	1705	89.09%
5_H2.1-1	118330	100.00%	657.2	54.40%
5_H2.1-2	140861	100.00%	990.3	61.01%
5_H2.5	141100	100.00%	1008	96.06%
5_H2.9	150532	100.00%	1048	97.66%
5_H3.1	333413	100.00%	2377	90.38%
5_H3.7	146061	100.00%	1035	96.48%
5_H3.14	328867	100.00%	2297	88.92%
5_S	373121	100.00%	2691	88.11%
5_B	141233	100.00%	1019	98.24%
7_H1.1	142355	100.00%	1002	96.64%
7_H1.2	113909	100.00%	797.9	78.21%

7_H1.3	192473	100.00%	1382	88.71%
7_H2.1	149420	100.00%	1064	98.19%
7_H2.2	106195	100.00%	764.7	87.27%
7_H2.3	1110832	100.00%	7876	90.59%
7_H3.1	172000	100.00%	1188	75.86%
7_H3.2	114003	100.00%	824.8	89.88%
7_H3.3	131545	100.00%	926.6	96.57%
7_S	332161	100.00%	2383	89.11%
7_B	143073	100.00%	1037	99.98%
8_H3.1	1136815	100.00%	7898	89.91%
8_H3.7	732808	100.00%	5021	84.52%
8_H3.13	397179	100.00%	1799	69.10%
8_H4.1	198808	100.00%	1301	83.87%
8_H4.8	1068574	100.00%	7620	86.26%
8_H4.16	1234437	100.00%	8684	90.13%
8_H5.1	163384	100.00%	1123	85.24%
8_H5.10	957204	100.00%	6713	87.46%
8_H5.19	1029087	100.00%	5165	88.22%
8_S	1756854	100.00%	12796	86.99%
8_B	132617	100.00%	978.7	88.37%

Tab. 15. Risultati coverage

Come si nota dalla tab.15 per le librerie si è ottenuto un'uniformità non ottimale per alcuni campioni ciò è dovuto a un problema di amplificazione dei due pool di primer. I valori di uniformità, come possiamo notare, vanno da un minimo di 54.40% fino a 99%, dopo l'ispezione manuale mediante il software IGV è stato possibile determinare che fino al 73% riuscivamo ad ottenere un profilo completo. Possiamo notare che il campion 5_H2.1 è stato sequenziato due volte perché nel primo sequenziamento si è ottenuto un profilo parziale (uniformità 54.40%) e un secondo (61.01%), ma il consenso dei due campioni ci ha permesso di ottenere un aplotipo completo analizzabile.

3.4 ANALISI VARIANTI ETEROPLASMICHE

Al fine di determinare le varianti presenti nei 55 campioni selezionati, è stata eseguita un'analisi primaria delle reads mediante il plug-in Variant Caller (v5.12.0.4, TVC). Questo plugin non è ottimizzato per l'analisi del DNA mitocondriale, per tale motivo tutte le varianti riportate dal TVC sono state ispezionate manualmente mediante un'analisi secondaria utilizzando il software IGV. L'analisi secondaria è stata eseguita in modo indipendente da due analisti diversi. Le soglie impostate per l'analisi corrispondono al 10% per la rilevazione delle eteroplasmie e 20x per il coverage minimo di lettura. Per 24 dei 55 campioni analizzati è stata eseguita anche l'analisi con il software Converge, i cui risultati sono stati confrontati con gli aplotipi ottenuti dall'analisi primaria con il TVC e l'ispezione manuale con IGV. Nonostante il Converge sia ottimizzato per l'analisi del DNA mitocondriale richiede comunque un'ispezione manuale mediante IGV per confermare o meno le posizioni indicate nel report come "possible", "likely", "unlikely".

Nella tabella 16,17,18,19,20 sono riportati le varianti confermate, ovvero gli aplotipi mitocondriali, per i 55 campioni selezionati per lo studio.

2_B	2_H4.1	2_H4.7	2_H4.14	2_H5.1	2_H5.5	2_H5.9	2_H6.1	2_H6.8	2_H6.15	
T195C	T195C	T195C	Contamination, haplotype not readable	T195C	T195C	Contamination, haplotype not readable	T195C	T195C	T195C	
A263G	A263G	A263G		A263G	A263G		A263G	A263G	A263G	A263G
309.1C	309.1C	rCRS		309.1C	309.1C		309.1C	309.1C	309.1C	309.1C
315.1C	315.1C	315.1C		315.1C	315.1C		315.1C	315.1C	315.1C	315.1C
A750G	A750G	A750G		A750G	A750G		A750G	A750G	A750G	A750G
A1438G	A1438G	A1438G		A1438G	A1438G		A1438G	A1438G	A1438G	A1438G
A4769G	A4769G	A4769G		A4769G	A4769G		A4769G	A4769G	A4769G	A4769G
A8860G	A8860G	A8860G		A8860G	A8860G		A8860G	A8860G	A8860G	A8860G
rCRS	rCRS	rCRS		9942R (G 81%, A 19%)	rCRS		rCRS	rCRS	rCRS	rCRS
rCRS	rCRS	rCRS		11312R (G 87%, A 13%)	rCRS		rCRS	rCRS	rCRS	rCRS
G13759A	13759R (G 12%, A 88%,)	G13759A		G13759A	G13759A		G13759A	G13759A	G13759A	G13759A
A15326G	A15326G	A15326G		A15326G	A15326G		A15326G	A15326G	A15326G	A15326G
15431R (G 84%, A 16%)	15431R (G 40%, A 60%)	rCRS		15431R (G 69%, A 31%)	rCRS		rCRS	rCRS	rCRS	rCRS
rCRS	rCRS	rCRS		rCRS	rCRS		rCRS	rCRS	rCRS	16189Y-1219 (C23%, T77%)
C16192T	C16192T	C16192T	C16192T	C16192T	C16192T	16192Y-443 (C 32%, T 68%)	16192Y-1532 (C19%-T 81%)	16192Y-1353 (C 27%, T 73%)		
A16219G	A16219G	A16219G	A16219G	A16219G	A16219G	A16219G	A16219G	A16219G		
C16261T	C16261T	C16261T	C16261T	C16261T	C16261T	C16261T	C16261T	C16261T		
T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C	T16311C		
H+195	H+195	H+195	H+195	H+195	H+195	H+195	H+195	H+195		

Tab.16. Aplotipi identificati per il tampone buccale, sangue e fusti di capelli per donatore 2, in grassetto sono evidenziate le eteroplasmie, in giallo l'aplogruppo di appartenenza.

Degli 11 campioni appartenenti al donatore 2 solo per 9 di questi è stato possibile ottenere un aplotipo mitocondriale, mentre per 2 campioni (2_H4.14

e 2_H5.9) non è stato possibile ottenere un aplotipo "leggibile" per via di una probabile contaminazione presente nei campioni. Relativamente al numero di varianti identificate queste vanno da un minimo di 13 ad un massimo di 17 per campione. Per i 9 campioni analizzati 11 PHP (Point Heteroplasmy Position) sono state identificate in 6 diversi siti. Dopo l'ispezione manuale tramite il tool IGV è stato possibile identificare le PHP in posizione 9942R e 11312R nel campione 2_H5.1 e la PHP nella posizione eteroplasmica 15431R, presente nei riferimenti 2_S e 2_B. Da notare che per questi 3 campioni è stata eseguita anche l'analisi con il software Converge, quest'ultima ha rilevato senza problemi la presenza delle PHP 9942R, 11312R e 15431R a differenza del TVC.

Per il donatore 4:

4_S	4_B	4_H2.1	4_H2.6	4_H2.11	4_H3.1	4_H3.5	4_H3.9	4_H4.1	4_H4.6	4_H4.11
A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G	A73G
C150T	C150T	C150T	C150T	C150T	C150T	C150T	C150T	C150T	C150T	C150T
A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G	A263G	cov<20x	A263G
C295T	C295T	C295T	C295T	C295T	C295T	C295T	C295T	C295T	cov<20x	C295T
315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	315.1C	cov<20x	315.1C
T489C	T489C	T489C	T489C	T489C	T489C	T489C	T489C	T489C	T489C	T489C
A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G	A750G
A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G	A1438G
A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G	A2706G
rCRS	rCRS	3042Y (T 86%, C 14%)	rCRS	3042Y (T 70%, C 30%)	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS
T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C	T4216C
A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G	A4769G
5460R (G 76%, A 24%)	5460R (G 70%, A 30%)	rCRS	rCRS	rCRS	G5460A	rCRS	rCRS	rCRS	rCRS	5460R (G 87%, A 13%)
C5633T	C5633T	C5633T	C5633T	C5633T	C5633T	C5633T	C5633T	C5633T	C5633T	C5633T
C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T	C7028T
C7476T	C7476T	C7476T	C7476T	C7476T	cov<20X	C7476T	C7476T	C7476T	cov<20x	C7476T
A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G
G10172A	G10172A	G10172A	G10172A	G10172A	G10172A	G10172A	G10172A	G10172A	G10172A	G10172A
A10398G	A10398G	A10398G	A10398G	A10398G	A10398G	A10398G	A10398G	A10398G	A10398G	A10398G
A11251G	A11251G	A11251G	A11251G	A11251G	A11251G	A11251G	A11251G	A11251G	A11251G	A11251G
G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A	G11719A
A12612G	A12612G	A12612G	A12612G	A12612G	A12612G	A12612G	A12612G	A12612G	A12612G	A12612G
G13708A	G13708A	G13708A	G13708A	G13708A	G13708A	G13708A	G13708A	G13708A	G13708A	G13708A
14302Y (T 85%, C 15%)	14302Y (T 72%, C 28%)	rCRS	rCRS	rCRS	T14302C	rCRS	rCRS	rCRS	rCRS	14302Y (T 89%, C 11%)
C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T	C14766T

G15257A	G15257A	G15257A	G15257A	G15257A	G15257A	G15257A	G15257A	G15257A	G15257A	G15257A
A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G	A15326G
C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A	C15452A
T15453C	T15453C	T15453C	T15453C	T15453C	T15453C	T15453C	T15453C	T15453C	T15453C	T15453C
G15812A	G15812A	G15812A	G15812A	G15812A	G15812A	G15812A	G15812A	G15812A	G15812A	G15812A
C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T	C16069T
T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C	T16126C
C16168T	C16168T	C16168T	C16168T	C16168T	C16168T	C16168T	C16168T	C16168T	C16168T	C16168T
J2b1f	J2b1f	J2b1f	J2b1f	J2b1f	J2b1f	J2b1f	J2b1f	J2b1f	J2b1f	J2b1f

Tab.

17. Aplotipi identificati per il tampone buccale, sange e fusti di capelli per donatore 4, in nero sono evidenziate le eteroplasmie, in giallo l'aplogruppo di appartenenza.

Degli 11 campioni appartenenti al donatore 4 solo per 10 di questi è stato possibile ottenere un aplotipo mitocondriale mentre per un campione (4_H4.6) non è stato possibile ottenere un aplotipo completo, ma solo parziale. Relativamente al numero di varianti identificate queste vanno da un minimo di 30 ad un massimo di 32 per campione. Per i 10 campioni analizzati 8 PHP sono state identificate in 3 diversi siti. Dopo l'ispezione manuale tramite il tool IGV è stato possibile identificare la posizione eteroplasmica 14302Y nel riferimento 4_S e nel frammento di fusto 4_H4.11, la PHP in posizione 3042Y nel campione 4_H2.11 e la PHP in posizione 5460R nel campione 4_H4.11. Inoltre è stata possibile identificare la variante omoplasmica in posizione C295T del campione 4_H3.1 filtrata dal TVC. Per il campione 4_S è stato possibile fare un'analisi omologa con il software Converge, quest'ultimo ha rilevato la presenza della posizione eteroplasmica 14302Y ma riportandola come

“Unclear”. Nonostante questo, dopo ispezione manuale, è stata confermata e riportata nell'aplotipo finale. Stesso risultato è stato ottenuto per la PHP 3042Y presente nel campione 4_H2.1.

Per il donatore 5:

5_S	5_B	5_H1.1	5_H1.6	5_H1.10	5_H 2.1 Con sens o	5_H 2.5 (chi p II)	5_H 2.9 (chi p II)	5_H 3.1	5_H 3.7 (chi p II)	5_H 3.14
A73G	A73G	A73G	A73G	A73G	A73 G	A73 G	A73 G	A73 G	A73 G	A73 G
T146C	T146C	T146C	T146C	T146C	T14 6C	T14 6C	T14 6C	T14 6C	T14 6C	T14 6C
T152C	T152C	T152C	T152C	T152C	T15 2C	T15 2C	T15 2C	T15 2C	T15 2C	T15 2C
T195C	T195C	T195C	T195C	T195C	T19 5C	T19 5C	T19 5C	T19 5C	T19 5C	T19 5C
A263G	A263G	A263G	A263G	A263G	A26 3G	A26 3G	A26 3G	A26 3G	A26 3G	A26 3G
C285T	C285T	C285T	C285T	C285T	C28 5T	C28 5T	C28 5T	C28 5T	C28 5T	C28 5T
315.1C	315.1C	315.1C	315.1C	315.1C	315. 1C	315. 1C	315. 1C	315. 1C	315. 1C	315. 1C
A750G	A750G	A750G	A750G	A750G	A75 0G	A75 0G	A75 0G	A75 0G	A75 0G	A75 0G
rCRS	rCRS	rCRS	rCRS	rCRS	G10 97A	G10 97A	G10 97A	rCR S	rCR S	rCR S
A1438G	A1438G	A1438G	A1438G	A1438G	A14 38G	A14 38G	A14 38G	A14 38G	A14 38G	A14 38G
T2387C	T2387C	T2387C	T2387C	T2387C	T23 87C	T23 87C	T23 87C	T23 87C	T23 87C	T23 87C
A2706G	A2706G	A2706G	A2706G	A2706G	A27 06G	A27 06G	A27 06G	A27 06G	A27 06G	A27 06G
A4769G	A4769G	A4769G	A4769G	A4769G	A47 69G	A47 69G	A47 69G	A47 69G	A47 69G	A47 69G
C7028T	C7028T	C7028T	C7028T	C7028T	C70 28T	C70 28T	C70 28T	C70 28T	C70 28T	C70 28T
C8395T	C8395T	C8395T	C8395T	C8395T	C83 95T	C83 95T	C83 95T	C83 95T	C83 95T	C83 95T

A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G	A8860G
rCRS	rCRS	9925R (G 86%, A 14%)	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS
T10885C	T10885C	T10885C	T10885C	T10885C	T10885C	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 (A 11%, G 89%)	A12308G	A12308G	A12308G	12308R (A 12%, G 88%)	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 (G 10%, A 90%)	G15148A	G15148A	G15148A	G15148A	G15148A	G15148A	G15148A
G15172A	G15172A	G15172A	15172R (G 10%, A 90%)	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	A15 954 C	A15 954 C	A15 954 C	A15 954 C	A15 954 C	A15 954 C
T16172C	T16172C	T16172C	T16172C	T16172C	T16 172 C	T16 172 C	T16 172 C	T16 172 C	T16 172 C	T16 172 C
T16249C	T16249C	T16249C	T16249C	T16249C	T16 249 C	T16 249 C	T16 249 C	T16 249 C	T16 249 C	T16 249 C
16261Y (C 86%, T 14%)	16261Y (C 85%, T 15%)	16261Y (C 35%, T 64%)	16261Y (C 36%, T 63%)	16261Y (C 42%, T 58%)	C16 261 T	C16 261 T	C16 261 T	C16 261 T	rCR S	rCR S
T16311C	T16311C	T16311C	T16311C	T16311C	T16 311 C	T16 311 C	T16 311 C	T16 311 C	T16 311 C	T16 311 C
C16327T	C16327T	C16327T	C16327T	C16327T	C16 327 T	C16 327 T	C16 327 T	C16 327 T	C16 327 T	C16 327 T
U1b3	U1b3	U1b3	U1b3	U1b3	U1b 3	U1b 3	U1b 3	U1b 3	U1b 3	U1b 3

Tab.
18. Aplotipi
identificati
per il
tampone
buccale,
sange e
fusti di
capelli per

donatore 5 in nero sono evidenziate le eteroplasmie, in giallo l'aplogruppo di appartenenza.

Per il donatore 5, degli 11 campioni sottoposti ad analisi è stato possibile ottenere un profilo completo per tutti i campioni, da notare che per il campione 5_2.1 l'aplotipo finale è stato ottenuto dal consenso del campione sequenziato due volte per problemi di amplificazione dei due pool primer del WGP. Relativamente al numero di varianti identificate queste vanno da un minimo di 34 ad un massimo di 36 per campione. Per i 11 campioni analizzati 10 PHP sono state identificate in 5 diversi siti. Dopo l'ispezione manuale tramite il tool IGV è stato possibile identificare la posizione eteroplasmica 16261Y nei riferimenti 5_S e 5_B e nel frammento di fusto 5_H1.1, la PHP in posizione 9925R nel campione 5_H1.1.

	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	Contami nation, haplotyp e not readable	A73G	A73G	A73G	A73G	A73G	A73G
rCRS	rCRS	215R (A 88%, G 12%)-1656	215R-1316 (A50%, G50%)		rCRS	rCRS	rCRS	rCRS	rCRS	rCRS
A263 G	A263 G	A263G	A263G		A263 G	A263 G	A263 G	A263 G	A263 G	A263 G
309.1 C	309.1 C	309.1C	309.1C		309.1 C	309.1 C	309.1 C	309.1 C	309.1 C	309.1 C
315.1 C	315.1 C	315.1C	315.1C		315.1 C	315.1 C	315.1 C	315.1 C	315.1 C	315.1 C
573.1 C	573.1 C	573.1C	573.1C		573.1 C	573.1 C	573.1 C	573.1 C	573.1 C	573.1 C
G709 A	G709 A	G709A	G709A		G709 A	G709 A	G709 A	G709 A	G709 A	709R (G 11%, A 89%) - 1931
A750 G	A750 G	A750G	A750G		A750 G	A750 G	A750 G	A750 G	A750 G	A750 G
G930 A	G930 A	G930A	G930A		G930 A	G930 A	G930 A	G930 A	G930 A	G930 A
A143 8G	A143 8G	A1438G	A1438G		A143 8G	A143 8G	A143 8G	A143 8G	A143 8G	A143 8G
G188 8A	G188 8A	G1888A	G1888A		G188 8A	G188 8A	G188 8A	G188 8A	G188 8A	1888R (G 13%, A 87%) -624
A270 6G	A270 6G	A2706G	A2706G		A270 6G	A270 6G	A270 6G	A270 6G	A270 6G	A270 6G
T421 6C	T421 6C	T4216C	T4216C		T421 6C	T421 6C	T421 6C	T421 6C	T421 6C	T421 6C
A476 9G	A476 9G	A4769G	A4769G		A476 9G	A476 9G	A476 9G	A476 9G	A476 9G	A476 9G
A491 7G	A491 7G	A4917G	A4917G		A491 7G	A491 7G	A491 7G	A491 7G	A491 7G	A491 7G
G514 7A	G514 7A	G5147A	5147R- 345 (G 20%, A 80%)		G514 7A	G514 7A	G514 7A	G514 7A	G514 7A	G514 7A
rCRS	rCRS	rCRS	6261R- 2039 (G 80%, A 20%)	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	
C702 8T	C702 8T	C7028T	C7028T	C702 8T	C702 8T	C702 8T	C702 8T	C702 8T	C702 8T	
rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	7206R (G 85%, A 15%)-588	
rCRS	rCRS	rCRS	8455Y- 203 (C 88%, T 12%)	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	
G869 7A	G869 7A	G8697A	G8697A	G869 7A	< 20X (A86 97G)	G869 7A	G869 7A	G869 7A	G869 7A	
rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	8765Y (C 86%,	

										T 14%) -37
A886 0G	A886 0G	A8860G	A8860G	A886 0G	A886 0G	A886 0G	A886 0G	A886 0G	A886 0G	A886 0G
T104 63C	T104 63C	T10463C	T10463C	T104 63C	1046 3Y- 305 (T 13%, C 87%)	T104 63C	T104 63C	T104 63C	T104 63C	T104 63C
rCRS	rCRS	rCRS	10646R- 175 (G 88%, A12%)	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS
A112 51G	A112 51G	A11251G	A11251G	A112 51G	A112 51G	A112 51G	A112 51G	A112 51G	A112 51G	A112 51G
G117 19A	G117 19A	G11719A	G11719A	G117 19A	G117 19A	G117 19A	G117 19A	G117 19A	G117 19A	G117 19A
A118 12G	A118 12G	A11812G	A11812G	A118 12G	1181 2R- 752 (A 23%, G 77%)	A118 12G	A118 12G	A118 12G	A118 12G	A118 12G
rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	1241 3Y (C 16%, T 84%) - 1173	1241 3Y-53 (C 55%, T 45%)	rCRS	rCRS
G133 68A	G133 68A	G13368A	G13368A	G133 68A	G133 68A	G133 68A	G133 68A	G133 68A	G133 68A	G133 68A
A142 33G	A142 33G	A14233G	A14233G	A142 33G	A142 33G	A142 33G	A142 33G	A142 33G	A142 33G	A142 33G
C147 66T	C147 66T	C14766T	C14766T	C147 66T	C147 66T	C147 66T	C147 66T	C147 66T	C147 66T	C147 66T
G149 05A	G149 05A	G14905A	G14905A	G149 05A	G149 05A	G149 05A	G149 05A	G149 05A	G149 05A	G149 05A
A153 26G	A153 26G	A15326G	A15326G	A153 26G	A153 26G	A153 26G	A153 26G	A153 26G	A153 26G	A153 26G
C154 52A	C154 52A	C15452A	C15452A	C154 52A	C154 52A	C154 52A	C154 52A	C154 52A	C154 52A	C154 52A
A156 07G	A156 07G	A15607G	A15607G	A156 07G	A156 07G	A156 07G	A156 07G	A156 07G	A156 07G	A156 07G
G159 28A	G159 28A	G15928A	G15928A	G159 28A	G159 28A	G159 28A	G159 28A	G159 28A	G159 28A	G159 28A
T161 26C	T161 26C	T16126C	T16126C	T161 26C	T161 26C	T161 26C	T161 26C	T161 26C	T161 26C	T161 26C
rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	rCRS	1616 9Y- 963 (C 63%, T 36%)	1616 9Y (C 39%, T 61%) - 2232	rCRS

C162 94T	C162 94T	C16294T	C16294T		C162 94T	C162 94T	C162 94T	C162 94T	C162 94T	C162 94T
T163 04C	T163 04C	T16304C	T16304C		T163 04C	T163 04C	T163 04C	T163 04C	T163 04C	T163 04C
T165 19C	T165 19C	T16519C	T16519C		T165 19C	T165 19C	T165 19C	T165 19C	T165 19C	T165 19C
T2b	T2b	T2b	T2b		T2b	T2b	T2b	T2b	T2b	T2b

Tab. 19. Aplotipi identificati per il tampone buccale, sange e fusti di capelli per donatore 7, in nero sono evidenziate le eteroplasmie, in giallo l'aplogruppo di appartenenza.

Per il donatore 7, sopra riportato:

Degli 11 campioni appartenenti al donatore 7 solo per 10 di questi è stato possibile ottenere un aplotipo mitocondriale mentre per un campione (7_H1.3) non è stato possibile ottenere un aplotipo "leggibile" per via di una contaminazione. Relativamente al numero di varianti identificate queste vanno da un minimo di 34 ad un massimo di 39 per campione. Per i 10 campioni analizzati 16 PHP sono state identificate in 13 diversi siti. Dopo l'ispezione manuale tramite il tool IGV è stato possibile identificare la posizione eteroplasmica 215R (7_H1.1), 8455Y e 10646R (7_H1.2), 7206R e 8765Y (7_H3.3). Solo per due campioni 7_S e 7_H2.3 è stata fatta un'analisi di confronto con il software Coverge, entrambi le analisi hanno rilevato lo stesso aplotipo mitocondriale per questi due campioni.

8_S	8_B	8_H3.1	8_H3.7	8_H3.13	8_H4.1	8_H4.8	8_H4.16	8_H5.1	8_H5.10
A73G	A73G	A73G	A73G	Contamination , haplotype not readable	A73G	A73G	A73G	A73G	A73G
249del	249del	249del	249del		249del	249del	249del	249del	249del

A263G	A263G	A263G	A263G
315.1C	315.1C	315.1C	315.1C
T489C	T489C	T489C	T489C
595.1C	595.1C	595.1C	595.1C
A750G	A750G	A750G	A750G
A1438G	A1438G	A1438G	A1438G
A1670T	A1670T	1670W (A 10%, T 90%)- 1810	A1670T
A2706G	A2706G	A2706G	A2706G
rCRS	rCRS	rCRS	rCRS
T3552A	rCRS	T3552A	T3552A
rCRS	rCRS	rCRS	rCRS
A4715G	A4715G	A4715G	A4715G
A4769G	A4769G	A4769G	A4769G
C7028T	C7028T	C7028T	C7028T
C7196A	C7196A	C7196A	C7196A
C7694T	C7694T	C7694T	C7694T
G8251A	G8251A	G8251A	G8251A
8271del	8271del	8271del	8271del
8272del	8272del	8272del	8272del
8273del	8273del	8273del	8273del
8274del	8274del	8274del	8274del
8275del	8275del	8275del	8275del
8276del	8276del	8276del	8276del
8277del	8277del	8277del	8277del
8278del	8278del	8278del	8278del
8279del	8279del	8279del	8279del
G8584A	G8584A	G8584A	G8584A
A8701G	A8701G	A8701G	A8701G

A263G	A263G	A263G	A263G	A263G
315.1C	315.1C	315.1C	315.1C	315.1C
T489C	T489C	T489C	T489C	T489C
595.1C	595.1C	595.1C	595.1C	595.1C
A750G	A750G	A750G	A750G	A750G
A1438G	A1438G	A1438G	A1438G	A1438G
A1670T	A1670T	A1670T	A1670T	A1670T
A2706G	A2706G	A2706G	A2706G	A2706G
rCRS	rCRS	rCRS	rCRS	3167Y (T 84%, C 16%)- 14157
T3552A	T3552A	T3552A	T3552A	T3552A
rCRS	rCRS	rCRS	4219R (G 81%, A 19%)- 355	4219R (G 64%, A 36%)- 4420
A4715G	A4715G	A4715G	A4715G	A4715G
A4769G	A4769G	A4769G	A4769G	A4769G
C7028T	C7028T	C7028T	C7028T	C7028T
C7196A	C7196A	C7196A	C7196A	C7196A
C7694T	C7694T	C7694T	C7694T	C7694T
G8251A	G8251A	G8251A	G8251A	G8251A
8271del	8271del	8271del	8271del	8271del
8272del	8272del	8272del	8272del	8272del
8273del	8273del	8273del	8273del	8273del
8274del	8274del	8274del	8274del	8274del
8275del	8275del	8275del	8275del	8275del
8276del	8276del	8276del	8276del	8276del
8277del	8277del	8277del	8277del	8277del
8278del	8278del	8278del	8278del	8278del
8279del	8279del	8279del	8279del	8279del
G8584A	G8584A	G8584A	G8584A	G8584A
A8701G	A8701G	A8701G	A8701G	A8701G

A8860G	A8860G	A8860G	A8860G
T9540C	T9540C	T9540C	T9540C
A9545G	A9545G	9545R (A 10%, G 90%)-2129	A9545G
A10398 G	A10398 G	A10398 G	A10398G
C10400T	C10400T	C10400T	C10400T
T10454C	T10454C	T10454C	T10454C
T10873C	T10873C	T10873C	T10873C
rCRS	rCRS	rCRS	rCRS
G11719 A	G11719 A	G11719 A	G11719A
G11914 A	G11914 A	G11914 A	G11914A
C12705T	C12705T	C12705T	C12705T
A13263 G	A13263 G	A13263 G	A13263G
T14318C	T14318C	T14318C	T14318C
C14766T	C14766T	C14766T	C14766T
T14783C	T14783C	T14783C	T14783C
G15043 A	G15043 A	G15043 A	G15043A
G15301 A	G15301 A	G15301 A	G15301A
A15326 G	A15326 G	A15326 G	A15326G
A15487T	A15487T	A15487T	A15487T
rCRS	rCRS	15639Y (T88%, C12%)-4979	rCRS
T16093C	T16093C	16093Y (T 22%, C 78%)-21843	16093Y (T 87%, C 13%)-17726
16111Y (C 48%, T 52%)-19843	16111Y - 1199 (C 33%, T 67%)	16111Y (C44%, T 56%)-3564	C16111T
C16223T	C16223T	C16223T	C16223T
C16234T	C16234T	C16234T	C16234T

A8860G	A8860G	A8860G	A8860G	A8860G
T9540C	T9540C	T9540C	T9540C	T9540C
A9545G	A9545G	A9545G	A9545G	A9545G
A10398 G	A10398 G	A10398G	A10398G	A10398 G
C10400T	C10400T	C10400T	C10400T	C10400T
T10454C	T10454C	T10454C	T10454C	T10454C
T10873C	T10873C	T10873C	T10873C	T10873C
rCRS	11505Y (T 84%, C 16%)-4447	rCRS	rCRS	rCRS
G11719 A	G11719 A	G11719A	G11719A	G11719 A
G11914 A	G11914 A	G11914A	G11914A	G11914 A
C12705T	C12705T	C12705T	C12705T	C12705T
A13263 G	A13263 G	A13263G	A13263G	A13263 G
T14318C	T14318C	T14318C	T14318C	T14318C
C14766T	C14766T	C14766T	C14766T	C14766T
T14783C	T14783C	T14783C	T14783C	T14783C
G15043 A	G15043 A	G15043A	G15043A	G15043 A
G15301 A	G15301 A	G15301A	G15301A	G15301 A
A15326 G	A15326 G	A15326G	A15326G	A15326 G
A15487T	A15487T	A15487T	A15487T	A15487T
rCRS	rCRS	rCRS	rCRS	rCRS
16093Y (T88%, C12%)-2333	rCRS	16093Y (T 77%, C 23%)-18778	16093Y (T 13%, C 87%)-3173	T16093C
16111Y (C64%, T 36%)-3113	16111Y (C 74%, T 26%)-20050	16111Y (C 75%, T 25%)-23491	16111Y (C69%, T 31%)-4658	16111Y (C 35%, T 65%)-20306
C16223T	C16223T	C16223T	C16223T	C16223T
16234Y (C10%, T 90%)-217	C16234T	C16234T	C16234T	C16234T

T16288C	T16288C	T16288C	T16288C	T16288C	T16288C	T16288C	T16288C	T16288C	T16288C
T16298C	T16298C	T16298C	T16298C	T16298C	T16298C	T16298C	T16298C	T16298C	T16298C
C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T	C16327T
G16518 T	G16518T	G16518T	G16518T	G16518T	G16518T	G16518T	G16518T	G16518T	G16518T
T16519C	T16519C	T16519C	T16519C	T16519C	T16519C	T16519C	T16519C	T16519C	T16519C
C16527T	C16527T	C16527T	C16527T	C16527T	C16527T	C16527T	C16527T	C16527T	C16527T
C5c1a	C5c1a	C5c1a	C5c1a	C5c1a	C5c1a	C5c1a	C5c1a	C5c1a	C5c1a

Tab. 20. Aplotipi identificati per il tampone buccale, sange e fusti di capelli per donatore 8, in nero sono evidenziate le eteroplasmie, in giallo l'aplogruppo di appartenenza.

Per il donatore 8, riportato su, degli 11 campioni solo per 9 di questi è stato possibile ottenere un aplotipo mitocondriale mentre due campioni (8_H3.13 e 8_H5.19) presentavano contaminazione. Relativamente al numero di varianti identificate queste vanno da un minimo di 55 ad un massimo di 58 per campione. Per i 9 campioni analizzati 21 PHP sono state identificate in 9 diversi siti. Dopo l'ispezione manuale tramite il tool IGV è stato possibile identificare la posizione eteroplasmica 15639Y (8_H3.1), 16093Y (8_H4.1), 4129R (8_H5.1) e 3167Y (8_H5.10). Per tutti questi campioni, eccetto il riferimento 8_B, è stata fatta un'analisi di confronto con il software Converge, entrambi le analisi hanno rilevato lo stesso aplotipo mitocondriale per questi campioni, rilevando senza problemi le posizioni eteroplasmiche riscontrate solo su IGV.

Gli aplotipi risultanti sono stati sottoposti a una ricerca nel database EMPOP mtDNA v4/R13. Il tool haplogrouping ha permesso di determinare l'aplogruppo di appartenenza per ogni donatore (tab.21).

Donatore	Aplogruppo
2	H+195
4	J2b1f
5	U1b3
7	T2b

Tab.21 Aplogruppo di appartenenza rilevato tramite EMPOM

L'analisi secondaria effettuata mediante il tool IGV ha permesso di individuare le "PHP apparenti" ovvero varianti che soddisfacevano la soglia impostata per le PHP, i.e. 10%. Un totale di 28 PHP apparenti (tab. 18) sono state identificate nei campioni di cui 15 classificate come "Possibili NUMT" e 13 classificate come "Possibile artefatto di sequenza".

Campione	PHP Apparente	Strand bias	In fase con altre varianti?	Variante attesa nell'aplogruppo o subaplogruppi?	Predizione	NUMT nota?	Risultato BLAT
4_H3.1	6645	si	si	no	NUMT	si	chr5: 5396435- 5397074
4_H3.1	14554	si	si	no	NUMT	no	chr1: 568766- 568888
2_H4.1	16093	no	no	no	NUMT	si	Poly_NumtS_1259 Dayama 2014
2_H6.1,2_H6.15,2_H6.8,4_B,4_H3.1,4_H3.5,4_H4.1,5_B,5_H2.1,5_H2.5,5_H2.9,5_H3.7,7_B,7_H1.1,7_H1.2,7_H2.1,7_H2.2,7_H3.1,7_H3.2,7_H3.3,8_B	12060M-12061M	si	si	no	Possibile artefatto di sequenziamento	-	-
5_H2.1	12927	si	si	no	NUMT	si	chr5: 5396435- 5397074
5_H2.1 , 7_H3.1	13129	si	si	no	NUMT	si	chrX: 95807916-95807935
7_H3.1	13419	si	si	no	Possibile artefatto di sequenziamento	-	-
4_H3.1	13750-13754	si	si	no	Possibile artefatto di sequenziamento	-	-
2_H6.1,2_H6.15,2_H6.8,4_B,4_H3.1,4_H3.5,4_H3.9,4_H4.1,5_H2.5,5_H2.9,5_H3.7,7_B,7_H1.1,7_H1.2,7_H2.1,7_H2.2,7_H3.1,7_H3.2,7_H3.3,	13883M	si	no	no	Possibile artefatto di sequenziamento	-	-
4_H2.1, 4_H2.11, 8_H4.16	14774Y-14775W-14777M	si	si	no	NUMT	si	chr5: 134259648- 134259780
5_H2.1	14811	si	no	no	Possibile artefatto di sequenziamento	-	-
2_H4.1	15303B	si	no	no	Possibile artefatto di sequenziamento	-	-
5_H2.1	15746	si	no	no	Possibile artefatto di sequenziamento	-	-
4_H3.1	15880-15884	si	si	no	NUMT	no	chr5: 16890582-16890601
2_H4.1,2_H4.7,4_H2.1,4_H2.11,4_H2.6,7_H2.3,8_H3.1,8_H3.7,8_H5.10	2230M-2232R-2233W-2245R	si	si	no	NUMT	no	chr5: 79946281-79946302
8_H4.1	2245R	si	si	no	NUMT	no	chr5: 79946281-79946302
4_H3.1	402-438-451-452	si	si	no	NUMT	no	chr5: 16890582-16890601

2_H6.1, 2_H6.15, 2_H6.8	4184Y-4186Y-4187Y	si	si	no	NUMT	si	chr1:238105788-238106144
2_H6.1,2_H6.15,2_H6.8,4_H3.9,4_H4.1,4_H4.11,7_H1.2,7_H2.2,7_H3.2,8_B,4_H3.1	4883M-4884M	si	si	no	Possibile artefatto di sequenziamento	-	-
5_H2.1	5205	si	no	no	Possibile artefatto di sequenziamento	-	-
5_H2.1	5434	si	no	no	Possibile artefatto di sequenziamento	-	-
2_H6.1,2_H6.15,2_H6.8,5_B,5_H2.1,5_H2.5,5_H2.9,5_H3.7,7_B,7_H1.1,7_H1.2,7_H2.1,7_H3.1,7_H3.3,8_B	7471M-7474R-7475S	si	si	no	NUMT	si	chr5: 5396435- 5397074
2_H6.1,2_H6.15,2_H6.8,4_B,4_H3.5,4_H3.9,4_H4.1,4_H4.11,5_H2.5,5_H2.9,5_H3.7,7_B,7_H1.1,7_H1.2,7_H2.1,7_H2.2,7_H3.1,7_H3.2,7_H3.3,4_H3.1	8248R-8249V-8250S-8252S-8256K	si	si	no	NUMT	si	chr1: 568766- 568888
2_B,2_H4.7,2_H5.1,4_H2.1,4_H2.11,4_S,7_H2.3,7_S	8249S	si	no	no	NUMT	no	chr1: 568803-568888
4_H3.1	8405-8406	si	si	no	NUMT	no	chr5: 16890582-16890601
2_H4.1	8649M-8652H-8653B	si	si	no	Possibile artefatto di sequenziamento	-	-
4_H2.11, 8_H3.1	8651W-8652M-8653M-8654Y	si	si	no	Possibile artefatto di sequenziamento	-	-
4_H3.1	9525-9526-9531-9532	si	no	no	Possibile artefatto di sequenziamento	-	-

Tab.22 PHP apparenti riscontrate nei 55 campioni selezionati per lo studio.

La predizione finale è stata ottenuta seguendo la flow chart dei checkpoints di analisi sviluppata da Cihlar e colleghi (Cihlar, 2020). Per i dati qui descritti, i risultati di strand bias e la presenza di ciascuna variante su EMPOP⁽²⁷⁾ ha permesso di determinare se queste PHP apparenti fossero il risultato di artefatti di sequenziamento o errori bioinformatici. L'omologia di sequenza determinata mediante il tool BLAT unitamente ai database contenenti NUMT osservate (Li 2012, Dayama 2014) e le informazioni filogenetiche sono state impiegate come passaggi informativi per determinare se queste PHP fossero di origine mitocondriale o nucleare. Se la variante della sequenza non era stata osservata in precedenza nell'aplogruppo, si è ritenuto meno probabile che derivasse dal mtGenome del donatore di quel campione. L'omologia con il genoma nucleare e/o la potenziale amplificazione di una regione nucleare tendeva a sostenere che PHP apparente fosse il risultato della co-amplificazione dei NUMT. Uno dei maggiori vantaggi delle tecnologie MPS è l'ottenimento di informazioni sulla fase (phasing) poiché ciascuna read è stata generata da un'unica molecola di DNA. Quindi, phasing o il riconoscimento di più di un aplotipo all'interno di un amplicone aiuta a identificare i NUMT. In figura 8 è riportato un esempio di una possibile NUMT identificata in pos. 8248-8256, poiché presenta più di un aplotipo all'interno di un amplicone.

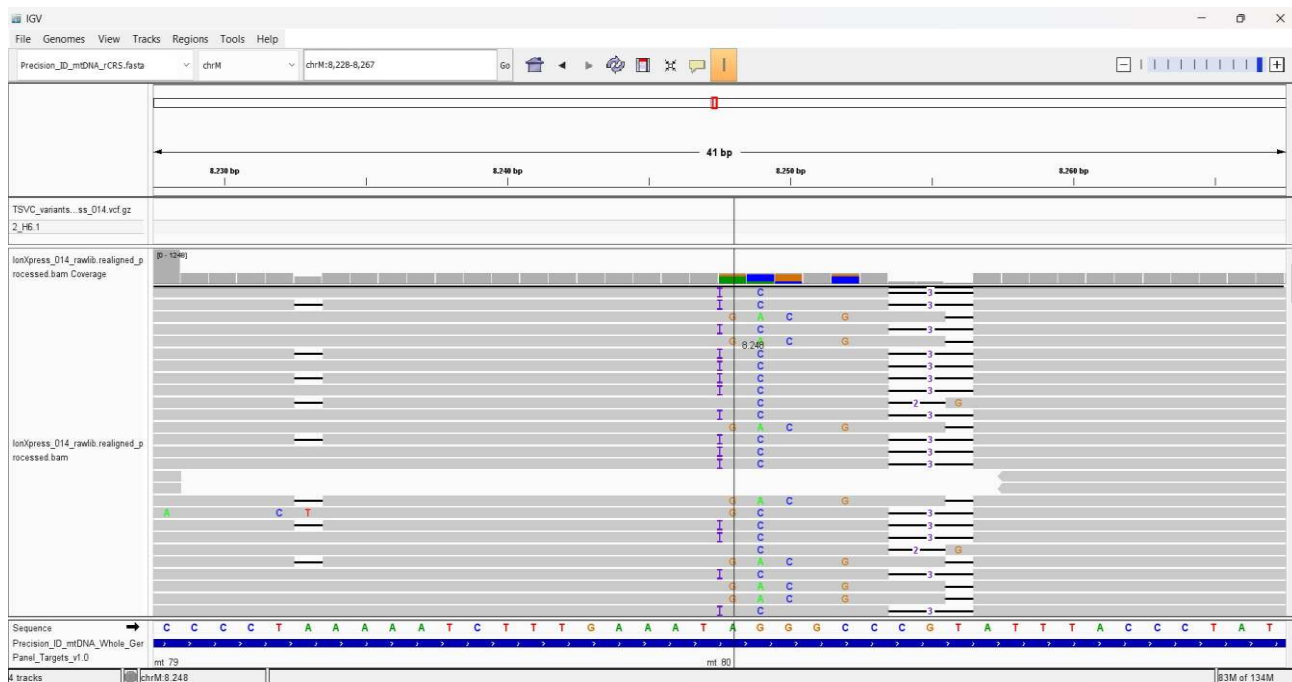


Fig.8 visualizzazione di NUMT multialleliche

Questa NUMT è stata già precedentemente osservata in (Strobl 2018, Faccinnetto 2019), per verificare l'omologia con il DNA nucleare è stata eseguita l'analisi con il tool BLAST identificando una regione omologa al chr1: 568766- 568888. Tuttavia, non sempre sono presenti varianti in fase, quindi l'informazione sulla fase non è informativa in questi casi poiché solo una variante è presente in un amplicone. In questi casi è necessaria un'analisi di omologia di sequenza per determinare se queste varianti con soglia del 10% fossero PHP apparenti o varianti di origine mitocondriale. Nei nostri risultati questa condizione si è verificata per le PHP apparenti alle posizioni 6645, 8249, 12927, 13129, 14554, 16093.

Le PHP apparenti classificate come “Possibile artefatto di sequenziamento” si riferiscono a varianti identificate solo in uno dei due strand di sequenziamento e/o nelle regioni omopolimeriche e ripetute, già segnalate in precedenza come regioni critiche per l’interpretazione dei dati per i sistemi ionici. L’estremo strand bias del filamento suggeriva che non erano il risultato di una contaminazione poiché il DNA esogeno sarebbe stato sequenziato in entrambe le direzioni. Pertanto, questi artefatti erano probabilmente il risultato di errori di sequenziamento e/o bioinformatici. Da notare che la maggior parte delle possibili NUMT/possibili artefatti di sequenziamento identificate sono state rilevate nel campione 5_H2.1; probabilmente questo è dovuto dai problemi di amplificazione inefficiente dei due pool di primer. La riduzione del numero di reads potrebbe aver comportato l’aumento, in termini di percentuale, della PHP apparente su di esse. L’analisi effettuata con il Converge non ha rilevato la presenza di nessun possibile artefatto e/o NUMT, ad eccezione di quelle note e catalogate nei database di NUMT pubblicati. Questo perché il Converge contiene statistiche NUMT e può rilevare e filtrare questo tipo di contaminazione.

CAPITOLO 4

DISCUSSIONE

Nel campo della genetica forense è noto che le applicazioni del DNA mitocondriale, rispetto al DNA nucleare, sono ridotte sia come tipologia che come potere discriminativo ma sono in relazione alle due più importanti caratteristiche biologiche del genoma mitocondriale: l'elevato numero di copie e l'assenza di ricombinazione. Proprio l'elevato numero di copie lo rendono indispensabile nell'analisi di reperti in cui il materiale cellulare è ridotto oppure è stato sottoposto a fenomeni di degradazione. Un reperto che viene tipizzato quasi esclusivamente attraverso l'analisi del DNA mitocondriale è rappresentato dalle formazioni pilifere, in particolare i capelli che spesso vengono rinvenuti sulla scena del crimine.

All'interno dei fusti di capelli possono esserci delle variazioni dell'mtDNA dovuto all'effetto collo di bottiglia che di conseguenza comporta una fissazione delle varianti eteroplasmatiche. Molti studi si sono concentrati sullo studio del collo di bottiglia e sui cambiamenti nelle frequenze alleliche dell'eteroplasmia che si verificano nei capelli umani poiché la presenza di tali varianti può complicare l'analisi del mtDNA specialmente se i risultati devono essere confrontati con il materiale di riferimento. D'altra parte una corretta

interpretazione dell'eteroplasmia può rafforzare il potere di identificazione del mtDNA.⁽²⁵⁾

Per la corretta interpretazione dei dati del mtDNA esperti nel campo della genetica forense hanno progettato questo studio, basato sulla rilevazione e interpretazione dell'eteroplasmia nei capelli e nei campioni di riferimento attraverso l'applicazione dell'innovativa tecnologia MPS. Per questo progetto sono stati raccolti per 5 diversi donatori due matrici da utilizzare come riferimento , quali sangue e tampone buccale, e fusti di capelli; ciò per confrontare gli aplotipi rilevati nei diversi tessuti. Il primo risultato ottenuto dai nostri dati riguarda la qualità/quantità del DNA mitocondriale dei fusti di capelli. Lo step della quantizzazione, oltre ad ottimizzare le successive fasi di amplificazione e creazione della libreria, ha sottolineato come la quantità di DNA mitocondriale diminuisce mano a mano che ci allontaniamo dal bulbo mentre aumenta la frammentazione del mtDNA quando ci avviciniamo alla punta del fusto. Questi risultati concordano con quanto riportato in letteratura negli studi condotti da Brandhagen ⁽²⁶⁾ , Turchi ⁽²⁴⁾ e Canale 2020 . Inoltre come riportato nello studio Turchi ⁽²⁴⁾ aumentare il tempo di incubazione dei fusti permette di ottenere una resa maggiore di DNA mitocondriale, non comportando un aumento della frammentazione del mitogenoma. Per il sequenziamento dell'intero genoma mitocondriale è stato utilizzato il Precision

ID mtDNA Whole Genome Panel 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 degradati. Le librerie create sono state sequenziate mediante lo strumento IonGene Studio 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). Successivamente un'analisi secondaria con il tool IGV ed un confronto con i report del Converge software ha permesso di determinare l'aplotipo mitocondriale per 50 dei 55 campioni analizzati. Le 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; Faccinetto. 202121). Grazie al software EMPOP mtDNA Database ci ha permesso di risalire all'aplotipo di appartenenza di ogni donatore: Donatore 2, Donatore 4, Donatore 5, Donatore 7, Donatore 8, appartenenti rispettivamente a questi aplogruppi: H+195, J2b1, 5U1B3, T2b, C5c1a. La conferma delle varianti per mezzo di IGV ha permesso di rilevare 18PHP vere (215R, 3042Y, 3167Y, 4219R, 5460R, 7206R, 8455Y, 8765Y, 9925R, 9942R, 10646R, 11312R, 14302Y, 15431R, 15639Y, 16093Y, 16261Y) e un variante omoplasmica (C295T); senza l'analisi secondaria queste posizioni eteroplasmiche sarebbero andate perse. Questo dimostra quanto sia

importante controllare i risultati di sequenziamento utilizzando diverse piattaforme, soprattutto in questo caso poiché il TVC è stato originariamente progettato per le mutazioni nucleari e non è stato ottimizzato per il genoma mitocondriale o il rilevamento dell'eteroplasmia, specialmente a bassi livelli. Dal confronto degli aplotipo ottenuti per i riferimenti con quelli ottenuti dai fusti di capelli sono state osservate delle modifiche, infatti i livelli di eteroplasmia possono variare da un tessuto all'altro di uno stesso individuo o anche tra individui imparentati per via materna.. Per il donatore 2 la 15431R è presente nei riferimenti 5_S e 5_B, questa posizione è presente anche nel primo frammento 2_H4.1 e in 2_H5.1, nel primo frammento è presente anche 13759R, non rilevata però nei riferimenti. I frammenti 2_H4.14 e 2_H5.9 presentano un numero molto alto di eteroplasmie, ciò ci fa ipotizzare ad una contaminazione. Nel 2_H5.1 è presente la 9942R, 11312R, non rilevate nei riferimenti.

Per il donatore 4, nei riferimenti 4_S, 4_B, è presente 5460R e 14302Y, quest'ultima presente anche nel frammento 4_H4.11. Nel secondo frammento 4_H2.1, 4_H2.11 è presente 3042Y, non rilevata nei riferimenti. Nel frammento 4_H4.6, era un frammento parziale, di conseguenza non si è riuscito a risalire all'aplotipo. Nel frammento 4_H4.11 è presente 5460R, ritrovata nei riferimenti. Per il donatore 5, la 16261Y presente nel 5_S,5_B, nel primo primo

frammento 5_H1.1, 5_H1.6, 5_H1.10, la 12308R nel 5_S e nel primo frammento 5_H1.10. 9925R presente nel frammento 5_H1.1. Le posizioni 15148 e 15172 nel frammento 5_H1.6, sono omoplasmiche. Per il donatore 7, i riferimenti 7_S, 7_B, presentano aplotipi completi, nel primo frammento è presente 215R 7_H1.1, 7_H1.2. Nel 7_H1.2 è presente 5147R, 6162R,8455Y,106464R, non rilevata anche nei riferimenti. Il 7_H1.3 risulta essere pieno di contaminazioni e non è stato possibile rilevare l'aplotipo. Nel secondo frammento 7_H2.2 è presente 10463Y , 11812R- 752 , non rilevati nei riferimenti. Nel terzo frammento 7_H3.1, 7_H3.2 è presente 12413Y e 16169Y-963. Nell'ultima parte del terzo frammento 7_H3.3 è presente 709R, 1888R, 7206R , 8765Y. Per il donatore 8, è presente 16111Y, nei riferimenti 5_S, 5_B, rilevata anche nel, 8_H3.1, 8_H4.1, 8_H4.8, 8_H4.16, 8_H5.1, 8_H5.10. Nella prima parte del terzo frammento 8_H3.1 è presente 1670W, 9545R, 15639Y, e 16093Y rilevata anche nei frammenti 8_H3.7, 8_H4.1, 8_H4.16, 8_H5.1. Nella frammento 8_H4.1 è presente inoltre 16234Y, non rilevata in nessun'altra posizione. Nel frammento 8_H4.8 11505Y, anch'essa non rilevata né nei riferimenti, né in altri frammenti. Nel frammento 8_H5.1, 8_H5.10 è presente 4219R, non rilevate in altre posizioni. Nel frammento 8_H5.10 è presente 3167Y, anche questa non rilevata in altre posizioni. Queste due posizioni dei frammenti, presentavano un numero molto alto di

eteroplasmie riconducibili a una probabile contaminazione, e di conseguenza non è stato possibile identificare l'aplotipo 8_H3.13,8_H5.19.

Il pannello utilizzato per l'amplificazione del DNA mitocondriale in regioni così piccole può esacerbare la co-amplificazione dei NUMT (Nuclear mitochondrial DNA segment). A causa della breve distanza tra le coppie di primer in questi multiplex, le regioni omologhe nel genoma nucleare hanno una maggiore possibilità di essere amplificate. Di conseguenza un'analisi accurata del mtDNA richiede l'identificazione del maggior numero possibile di NUMT e artefatti. Un totale di 28 PHP apparenti (tab. 18) sono state identificate nei campioni di cui 15 classificate come "Possibili NUMT" e 13 classificate come "Possibile artefatto di sequenza". La predizione finale è stata ottenuta seguendo la flow chart dei checkpoints di analisi sviluppata da Cihlar e colleghi. Per i nostri dati, l'omologia di sequenza è stata determinata mediante il tool BLAT unitamente ai database contenenti NUMT note e informazioni filogenetiche sono state impiegati per determinare se queste PHP fossero di origine mitocondriale o nucleare. Le PHP apparenti classificate come "Possibile artefatto di sequenziamento" si riferiscono a varianti non presenti in EMPOP, identificate solo in uno dei due strand di sequenziamento e/o nelle regioni omopolimeriche e ripetute, già segnalate in precedenza come regioni critiche per l'interpretazione dei dati per i sistemi ionici. Ecco che l'interpretazione dei

dati può risultare complicata, proprio a causa della presenza delle NUMT, le quali vengono considerate erroneamente eteroplasmie. Oggi grazie all'avanzata tecnologia delle piattaforme MPS è possibile rilevare varianti a bassa frequenza (<10%), non rilevabili dal Sanger con lettura 15-20%, questo permette un'analisi molto più accurata e precisa.

CAPITOLO 5

CONCLUSIONI

Ad oggi possiamo vedere molteplici miglioramenti nell'ambito della genetica forense, questo è stato possibile anche attraverso l'applicazione del Massive Parallel Sequencing il quale analizza simultaneamente milioni di frammenti, consentendo l'analisi dei genomi mitocondriali completi in poco tempo, con un'alta sensibilità, riuscendo a rilevare posizioni che il sequenziamento di Sanger non riuscirebbe a leggere. La maggiore sensibilità e la natura quantitativa dei dati forniti dalle piattaforme MPS ci permettono il rilevamento e la caratterizzazione dell'eteroplasmia di livello inferiore rispetto alle consolidate tecnologie STS-CE. L'MPS è stato testato su molti campioni forensi, ma pochi studi hanno fatto uso di questo per valutare i livelli di varianti eteroplasmatiche nel mitogenoma. Questo progetto ha valutato la presenza mutazioni somatiche nei capelli umani mediante l'utilizzo dell' MPS.

Lo scopo della mia tesi si è concentrata sullo studio della variazione della sequenza del mtDNA nei fusti dei capelli, dove il fenomeno del collo di bottiglia porta alla segregazione delle molecole del mtDNA durante lo sviluppo dei capelli, causando eteroplasmia o fissazione di una variante che differisce dall'aplotipo osservato in altri tessuti, in questo caso come riferimenti abbiamo raccolto un

tampone buccale e un campione di sangue, comunemente raccolti nel contesto forense. Le NUMT, generate dal trasferimento del mtDNA nel genoma nucleare, possono co-amplificare con i piccoli ampliconi amplificati con il Precision ID mtDNA Whole Genome Panel. Questo fenomeno a sua volta può generare una miscela di sequenza che possono essere interpretate erroneamente come eteroplasmie. Lo studio qui illustrato ha evidenziato molteplici PHP apparenti che, a causa della loro frequenza e della mancanza di altre varianti, non erano immediatamente identificabili come NUMT. È stato quindi utilizzata una valutazione cumulativa di diversi parametri quali strand bias, phasing, presenza variante in EMPOP, omologia di sequenza per predire se l'allele alternativo in ciascuna PHP fosse una NUMT, un artefatto o una variante mitocondriale. Un totale di 28 PHP apparenti sono state identificate nel nostro studio; la maggior parte di esse sono state trovate nei fusti di capelli rispetto agli altri tessuti ciò ci conferma proprio come queste possano compromettere l'analisi dei risultati. Ad oggi grazie alle moderne tecnologie si cerca sempre più di superare questi ostacoli al fine di poter avere dei risultati ottimali. Da questo nostro studio e successivamente analisi possiamo confermare l'alta sensibilità del Massive Parallel Sequencing, in quanto ci ha permesso la rilevazione delle varianti anche di basso livello, grazie alla sua capacità di lettura utilizzando un threshold molto più basso rispetto a Sanger.

BIBLIOGRAFIA:

1. Introduzione alla genetica forense - Indagini di identificazione personale e di paternità di Adriano Tagliabracci; Springer Editore
2. Fundamentals of Forensic DNA Typing, Jonh M. Butler
3. Adriano Tagliabracci and Chiara Turchi Section of Legal Medicine, Department of Excellence of Biomedical Sciences and Public Health, Polytechnic University of Marche, Ancona Italy
4. Biologia- cellula e tessuti; Edi- ermes
5. Mitomap, Sequenza Cambridge
6. The human mitochondrial genome, mtDNA exploitation in forensic . Adriano Tagliabracci e Chiara Turchi
7. (Bethany Kathleen Forsythe), A thesis submitted that meets the requirements for the title of Doctor of Philosophy in Forensic Science, the University of Auckland, 2023.
8. mtDNA Heteroplasmy: Origin, Detection, Significance, and Evolutionary Consequences, Maria-Eleni Parakatselaki 1, Emmanuel D Ladoukakis 1
9. Stephen P. Burr and Patrick F. Chinnery Department of Clinical Neurosciences and MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, United Kingdom
10. Adriano Tagliabracci and Chiara Turchi Section of Legal Medicine, Department of Excellence of Biomedical Sciences and Public Health, Polytechnic University of Marche, Ancona Italy

11. 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
12. Greiciane G Paneto 1, Larissa V.G. Longo , Joyce A. Martins , Maria Angelica de Camargo , Jeane C Costa , Aline CO de Mello , Bety Chen , Rogério N Oliveira , Mario H Hirata , Regina MB Cicarelli.
13. 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
14. Lopez, J.V., Yuhki, N., Modi, W., Masuda, R. and O'Brien, S.J. (1994). " Numt, a recent transfer and tandem amplification of mitochondrial DNA in the nuclear genome of the domestic cat ". *J Mol Evol.* 39 (2): 174–190. Bibcode:1994JMolE..39..174L. doi:10.1007/BF00163806. PMID 7932781. S2CID 22165470
15. Interpreting NUMTs in forensic genetics: Seeing the forest for the trees Charla Marshall, Walther Parson. PMID: 33740708 DOI: 10.1016/j.fsigen.2021.102497
16. Greiciane G Paneto, Larissa V.G. Longo , Joyce A. Martins , Maria Angelica de Camargo , Jeane C Costa , Aline CO de Mello , Bety Chen , Rogério N Oliveira , Mario H Hirata , Regina MB Cicarelli.
17. Parson et al. (2013) Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Lagace R. *Forensic science international: valutazione genetica del*

sequenziamento mtGenome di nuova generazione utilizzando la Ion Torrent Personal Genome Machine (PGM) 2013. 7, 543–549

18. Amorim & Pinto (2018) Amorim A, Pinto N. Big data in genetica forense. *Forensic Science International: Genetica*. 2018; 37 (agosto): 102–105. doi: 10.1016/j.fsigen.2018.08.001

19. Parson & Dür (2007) Parson W, Dür A. EMPOP—un database mitocondriale forense. *Forensic Science International: Genetica*. 2007; 1 (2): 88–92. doi: 10.1016/j.fsigen.2007.01.018.

20. Ruiz-Pesini et al. (2007) Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, Mishmar D, Wallace DC. Una MITOMAP potenziata con una filogenesi mutazionale del mtDNA globale. *Ricerca sugli acidi nucleici*. 2007; 35 (SUPPL. 1):823–828. doi: 10.1093/nar/gkl927

21. Stoneking et al. (1991) Stoneking M, Hedgecock D, Higuchi RG, Vigilant L, Erlich HA. Variazione della popolazione delle sequenze della regione di controllo del mtDNA umano rilevate dall'amplificazione enzimatica e dalle sonde oligonucleotidiche specifiche della sequenza. *Giornale americano di genetica umana*. 1991; 48 (2): 370–382.

22. Xavier C, Eduardoff M, Strobl C, Parson W. SD quant- Sistema tetraplex di rilevamento sensibile per la quantificazione e l'inferenza della degradazione del DNA nucleare e mitocondriale. *Forensic Sci Int Genet*. 2019 Sep;42:39-44. doi: 10.1016/j.fsigen.2019.06.004. Epub 2019 Jun 10. PMID: 31216503.

23. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): Visualizzazione ed esplorazione dei dati genomici ad alte prestazioni. *Brief. Bioinform*. 2013, 14, 178–192)

24. Chiara Turchi, Filomena Melchionda, Federica Alessandrini, Valerio Onofri, Mauro Pesaresi, Loredana Buscemi, Adriano Tagliabracci, Sezione di Medicina Legale, Dipartimento di Scienze Biomediche e Sanità Pubblica, Università Politecnica delle Marche, Ancona, Italia. *Forensic Science International: Supplemento di genetica serie 8* (2022) 219–221

25. Just RS, Irwin JA, Parson W. Eteroplasmia del DNA mitocondriale nel campo emergente del sequenziamento massivo parallelo. *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

26. Brandhagen MD, Loreille O, Irwin JA, Fragmented nuclear DNA is the predominant genetic material in human hair stems, *Genes (Basel)*, 2018 Dec 18;9(12):640, DOI: 10.3390/genes9120640, PMID: 30567392; PMCID: PMC6316335.

27. Jennifer Churchill Cihlara,b,*, Christina Stroblc, Robert Lagacéd, Melissa Muenzlera, Walther Parsonc,e, Bruce Budowlea,b. Distinguishing mitochondrial DNA and NUMT sequences amplified with the precision ID mtDNA whole genome panel.

SITOGRAFIA IMMAGINI:

Fig. 1. <https://testdimedicina.altervista.org/blog/genoma-mitocondriale/>

Fig. 2. microbiologiaitalia.it/didattica/mitocondri/

Fig. 3. [Structure-of-mitochondrial-DNA-Mammalian-mitochondrial-DNA-mtDNA-is-a-double-stranded_fig2_51089175](#)

Fig. 4.5 capitolo 7 the human mitochondrial.

Fig. 5. 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

Fig 6. The human mitochondrial

Fig. 7. Stephen P. Burr and Patrick F. Chinnery Department of Clinical Neurosciences and MRC Mitochondrial Biology Unit. University of Cambridge. Cambridge. United Kingdom

Fig.10,11 Rotor Gene Q v, 2, 1, 0, di Qiagen

Tab. 1.2.3.4.5.6.7.8.9.10,11 Precision ID mtDNAPanels with the HID Ion S5™ /HID Ion GeneStudio™ S5 System

Tab 12,13,14,15 Torrent Suite Software (v5.12.3)

RINGRAZIAMENTI

Con il raggiungimento di questo traguardo, si chiude per me un capitolo molto importante, e non posso che ringraziare tutti coloro che sono sempre stati al mio fianco, ma un ringraziamento particolare lo devo a me stessa per non aver mai mollato, alla mia tenacia e alla mia forza che non sapevo di avere.

Un ringraziamento speciale, va alla mia famiglia, i miei più grandi sostenitori, il mio punto di riferimento, coloro che hanno sempre creduto in me ancor prima che tutto avesse inizio, senza voi sicuramente oggi non sarei qui e a voi, che con grande orgoglio dedico questa mia laurea.

A te Papà, grazie per non avermi fatto mai mancare nulla, in ogni momento ti ho sempre sentito vicino, pronto sempre a darmi supporto. La mia gioia più grande era sentir gioire te ogni volta in cui ti chiamavo comunicandoti di aver fatto l'esame, era lì che mi sentivo felicissima e soddisfatta. Non mi sono mai sbilanciata nel dire cose, ma oggi mi sento di dirti che tu per me sei l'esempio di vita più bello, perché mi hai insegnato, che quando la vita ci pone davanti a degli ostacoli, in realtà non è così, ma rappresentano solo dei punti di partenza, per arrivare ad un finale del tutto inaspettato.

A te Mamma, tu che per me sei la mia roccia, il mio più grande punto di riferimento, grazie per tutto il tempo che mi dedichi, grazie per aver condiviso con me ogni singolo momento, ogni singola mia paura, ogni esame, grazie per le lunghe telefonate, in cui ti assillavo sulle mie mille paranoie, ma in quei momenti di grande sconforto le tue parole erano più importanti di un abbraccio. Grazie per avermi incoraggiata a dare sempre il meglio di me, e per aver visto in me capacità a me ignare, grazie per avermi dato forza sempre e per essermi stata sempre vicina. Ma soprattutto grazie per avermi fatto capire che nella vita non bisogna mai arrendersi, ma lottare per ciò che si vuole, ed io sono così tanto orgogliosa e fiera di te.

A te Arturo, l'altra parte del mio cuore, grazie per essere così come sei, sei la cosa più preziosa che ho. Non basterebbero parole per descrivere il nostro legame, non c'è stato giorno in cui io non ti ho sentito vicino, non c'è stato giorno in cui non ricevevo un messaggio da parte tua, anche un semplice 'Lu', per me era più importante di tutto. Grazie per aver alleviato le mie ansie, con la tua solarità contagiosa, ogni volta in cui ti scoccavo con le mie paranoie, posso dirti che sei il mio supporto più grande, e il mio più grande rifugio. Senza te sarei persa. Sappi che io sarò sempre la tua spalla destra, e in ogni momento io ti sarò sempre vicina. Ti voglio bene.

Un ringraziamento particolare lo devo alla mia relatrice Chiara Turchi, la quale ha permesso la stesura di questo mio elaborato. Un ringraziamento va anche alla Dottoressa Filomena Melchionda, per avermi guidata passo dopo passo, in questo mio elaborato ricercando sempre la perfezione. Vi ringrazio calorosamente per il vostro supporto.

Un pensiero particolare va a mia nonna Lucia, anche se non sei qui fisicamente, so che da lassù, stai gioendo insieme a noi, e che sarai così tanto orgogliosa. Grazie per i tuoi consigli che custodisco preziosamente nel mio cuore.

Grazie a mia nonna Nina per la tua vicinanza e affetto.

A mia cugina Lucia, mia sorella, la mia complice in tutto, grazie per essermi stata sempre vicina nonostante gli svariati chilometri che ci separavano, grazie per aver condiviso con me ogni step della mia vita.

Un grazie particolare e speciale va a coloro che sono al mio fianco da 12 anni, Rossella, Laura, Titti, Silvana, Federica, nonostante la lontananza, siete sempre state ad un passo da me, confermando giorno dopo giorno il nostro legame, grazie per il vostro affetto costante, per il vostro supporto ai miei sfoghi paranoici, grazie per le nostre serate folli e indimenticabili, siete per me il dono più prezioso, e sono così orgogliosa di condividere con voi questi miei traguardi. Vi voglio bene.

A te Rossella, grazie per la persona che sei, grazie per la tua vicinanza sempre, per aver condiviso con me le mie ansie, paure e per il tuo conforto costante. Grazie per essere sempre riuscita a farmi sentir meno l'ansia angosciante con quelle piccole battute che ormai sono routine, e quasi quasi mi mancano. Sembra assurdo che circa sei anni fa iniziammo insieme questo percorso, che durò solo per poco, ma il destino ha voluto che entrambe ci laureassimo nello stesso periodo, e non sai quanto io sia orgogliosa e contenta di poter condividere questa mia gioia con te.

Grazie a Titti, la mia sostenitrice più grande, tu che hai sempre creduto in me, sin dal liceo. Grazie per essermi stata accanto, per avermi rassicurata, e soprattutto per aver fatto emergere in me la parte più razionale. Non dimenticherò mai la tua contentezza e orgoglio, quando comunicai che ero prossima alla laurea, riempiendomi il cuore di gioia. Grazie per ciò che sei, sono così tanto orgogliosa di te, e della donna che sei.

Grazie a Laura, la piccola grande donna, tu che mi sei vicina dai tempi delle medie, grazie per la tua vicinanza, per i tuoi incoraggiamenti così diretti e precisi. Grazie per avermi sempre dato forza, anche quando eri tu ad averne bisogno.

Grazie a Silvana, anche se nell'ultimo anno non ci siamo viste molto, ti ho sempre sentita vicina, e pronta a darmi sempre supporto.

Grazie a Federica, grazie per essermi sempre stata accanto con la tua estrema dolcezza, grazie per aver placato le mie ansie con le tue sedute psicologiche, grazie per il tuo supporto costante.

Grazie a voi Giorgia e Gigia, le mie las chicas de cable, così opposte tra noi, ma così perfettamente complici e unite. Grazie a voi che mi avete insegnato a vivere la vita con più leggerezza, e a studiare e affrontare gli esami con più spensieratezza, siete state un punto di riferimento importante in questo mio percorso, siete state le mie ombre, per me siete state casa. Grazie per aver

condiviso con me le mie paure, ansie, per avermi ascoltata ripetere fino alla nausea, grazie per tutto il supporto che mi avete dato, ma soprattutto grazie per tutto il tempo trascorso insieme. Siete delle persone speciali e vi porterò sempre con me.

A te Gigia, la scoperta più bella che la vita mi abbia riservato. Ti ho conosciuta in un momento in cui ero diffidente da tutti, ma tu con la tua solarità, con la tua vitalità, dolcezza, e tra una ricetta fit e l'altra sei riuscita a farti spazio nel mio cuore. Grazie per le nostre lunghe chiacchierate, per avermi sempre ascoltata, per aver placato le mie ansie, per avermi dato forza in ogni istante, grazie per avermi sempre teso la mano, molte volte facendoti male. Non potrò mai dimenticare tutto ciò che abbiamo vissuto e condiviso a partire dalle ansie, risate e a concludere con un semplice bagnoschiuma, sei stata la mia complice in tutto, bastava uno sguardo per capirci.. Non sono mancati però momenti in cui la contraddizione prendeva il sopravvento, definirei dire che è routine, ma se inizialmente mi sembrava strano, ad oggi posso dire che è ciò che contraddistingue il nostro rapporto. Sono così tanto fiera di te, della persona che sei, continua a credere sempre in te e nelle tue grandi capacità e sono sicura che riuscirai a realizzare il tuo sogno più grande.

A te Giorgia, come ti ho sempre detto sei stata la prima persona che ho conosciuto al mio primo esame in triennale, e l'ultima persona che è stata con me al mio ultimo esame in magistrale. La vita avvolte può sembrare strana, ma al tempo stesso ci riserva sorprese inaspettate. Ricordo ancora quando circa due anni fa iniziammo a parlare un po di più, ma ti vedevo sempre così fredda, poi col passare del tempo ho conosciuto la vera Giorgia, solare, con un cuore d'oro, e sempre pronta ad aiutarmi in ogni situazione. Grazie per essermi stata vicina, per aver condiviso con me tutto, casa, gioie, momenti di sconforto, esami che per osmosi hai studiato anche tu, ma soprattutto le nostre amate camomille alle 3 di notte in preda al panico. Sei stata una presenza fondamentale in questo mio percorso, più di una semplice amica, e io sono così tanto felice, orgogliosa di averti conosciuto, e di averti qui con me oggi.

Grazie alla mia amiga Zoi, la mia shipe preferita, la mia compagna di vida loca, impossibile spiegare la velocità con cui ci siamo legate così tanto, ma tu

con la tua estrema dolcezza, solarità, e premura verso ogni singola cosa, sei riuscita a farti voler bene sin da subito. Spero che questo rapporto così bello perduri nel tempo. Grazie per il tuo incoraggiamento costante, grazie per le nostre serate 'folli', ma soprattutto grazie per essermi stata vicina.

A Teresa, o meglio l'amica di Arturo, grazie per avermi rallegtrato le giornate con la tua spensieratezza e la tua risata così contagiosa, per essermi stata vicina, e avermi incoraggiata sempre, ti porterò sempre con me.

A loukia e katerina, grazie per il vostro affetto, per i vostri consigli, e la vostra vicinanza.

A Tina, la nostra prof, grazie per essermi stata accanto in ogni momento, grazie per i tuoi consigli preziosi. Non dimenticherò mai le nostre cene improvvisate, che avevano sempre profumo di casa, conserverò ogni singolo ricordo.

Grazie a te, Giulia, per aver condiviso con me tutto il mio percorso universitario, partendo dalla triennale, sino ad oggi. Nonostante tutto, mi sei sempre stata vicina, dandomi supporto e continuando a condividere tutto ciò che insieme avevamo iniziato.

Un grazie particolare va al mio compagno di vita, Pierdante. Grazie per avermi seguita passo dopo passo, in ogni punto della mia vita. Grazie per avermi dato forza, per avermi stimolata a ricercare sempre il meglio e perfezionare ogni singola cosa. Grazie per avermi supportato e sopportato e non a caso meriteresti un'altra laurea ad honorem in psicologia. Grazie per aver sempre cercato di placare le mie ansie e le mie mille insicurezze. Sono così tanto fiera e grata di avere nella mia vita te, sei sempre stato a un passo da me, non solo nelle gioie, ma anche in momenti di forte sconforto, ma insieme abbiamo sempre vinto tutto, nonostante sappiamo entrambi quanto la distanza sia stata dura. Come dissi tu qualche mese fa, questa è la nostra vittoria, e ora finalmente con tanto orgoglio non posso che assecondare questa frase. Senza te nulla sarebbe stato così. Ti amo.

