

INDICE

Capitolo 1 - INTRODUZIONE	1
1 - MARCATORI GENETICI PER L'ANALISI DI CAMPIONI FORENSI	2
1.1 Analisi del DNA.....	2
1.2 I polimorfismi del DNA in genetica forense	3
1.3 STRs per l'identificazione personale.....	7
1.4 SNPs in genetica forense.....	10
1.5 Microplotipi del DNA	12
1.5.1 L'introduzione dei microplotipi del DNA	14
1.5.2 Nomenclatura dei microplotipi	15
1.5.3 Vantaggi e svantaggi dell'utilizzo dei microplotipi in genetica forense	17
1.5.4 Criteri per la selezione dei microplotipi	19
1.5.5 Applicazioni di microplotipi in genetica forense.....	21
<i>Microplotipi nell'identificazione personale</i>	<i>22</i>
<i>Microplotipi nell'identificazione di parentele</i>	<i>23</i>
<i>Microplotipi nella deconvoluzione di campioni misti.....</i>	<i>25</i>
<i>Microplotipi nell'ascendenza biogeografica</i>	<i>26</i>
<i>Microplotipi in genetica medica</i>	<i>27</i>
2 - LA GENETICA FORENSE: LIMITI E POTENZIALITÀ.....	29
2.1 Degradazione del DNA	34
2.2 Controlli di qualità e prevenzione della contaminazione	37
3 - SISTEMI DI RILEVAMENTO DEI MARCATORI GENETICI	40
3.1 Elettroforesi Capillare (CE).....	40
3.2 Sequenziamento del DNA: metodo di Sanger	41
3.3 Next Generation Sequencing	42
3.3.1 Tecnologia Ion Torrent: la piattaforma Ion PGM™ System	47
Capitolo 2 - SCOPO DELLO STUDIO.....	54
Capitolo 3 - MATERIALI E METODI.....	56
3.1 Selezione dei loci d'interesse	56

3.2 Disegno dei primer	56
3.3 Campioni utilizzati per la validazione del pannello	59
3.4 Estrazione e quantizzazione	61
3.4.1 Estrazione DNA con QIAamp® DNA Micro.....	62
3.4.2 Estrazione con fenolo/cloroformio	65
3.4.3 Estrazione con QIAamp® DNA Investigator	66
3.4.4 Protocollo di depurinazione	66
3.4.5 Quantizzazione DNA con Quantifiler® Trio DNA Quantification Kit	67
3.4.6 Quantizzazione con E-Gel™ Agarose Gels with SYBR Safe™	70
3.4.7 Quantizzazione con Agilent 2200 TapeStation System.....	71
3.4.8 Massive Parallel Sequencing	73
<i>Amplificazione delle regioni target</i>	73
<i>Parziale digestione degli ampliconi</i>	75
<i>Ligazione dei barcode</i>	76
<i>Purificazione delle librerie</i>	77
<i>Quantizzazione delle librerie</i>	79
<i>Amplificazione aggiuntiva delle librerie con Platinum™</i>	81
<i>Purificazione delle librerie riamplicate con Platinum™</i>	82
3.4.9 Emulsion PCR - Ion OneTouch 2™	84
3.4.10 Arricchimento/purificazione delle librerie - Ion OneTouch™ ES	87
3.4.11 Sequenziamento con Ion PGM™ System	89
Capitolo 4 - RISULTATI	92
4.1 Selezione dei markers e disegno dei primer	92
4.2 Risultati della quantizzazione delle librerie	94
4.3 Risultati della quantizzazione delle librerie riamplicate e purificate con Platinum™	96
4.4 Risultati del sequenziamento e analisi del coverage	97
4.5 Risultati della genotipizzazione dei campioni	104

Capitolo 5 - DISCUSSIONE.....	117
Capitolo 6 - CONCLUSIONI.....	128
Capitolo 7 - BIBLIOGRAFIA.....	130
RINGRAZIAMENTI.....	141

Capitolo 1

INTRODUZIONE**1. MARCATORI GENETICI PER L'ANALISI DEI
CAMPIONI FORENSI**

La genetica forense è una branca delle scienze forensi che tramite l'analisi del DNA è capace di risalire all'identità di individui a partire da campioni biologici di diverso genere. La traccia biologica di partenza può essere di diversa origine, quale sangue, saliva, sperma, tampone buccale, capelli, reperti ossei, frammenti di pelle o unghie.

La genetica forense assume un ruolo fondamentale nelle attività investigative, come l'identificazione di soggetti indagati per omicidi, violenze sessuali o aggressioni, l'attribuzione individuale di resti umani sconosciuti, il riconoscimento cadaverico nei disastri di massa, l'accertamento delle relazioni parentali.

Ad oggi la biologia molecolare è la base della genetica forense, ma inizialmente l'analisi forense si basava su caratterizzazione dei polimorfismi proteici, quali sistema ABO, Rh e HLA, con metodiche immunologiche ed

elettroforetiche (1). Queste tecniche però sono state presto sostituite dalla biologia molecolare in quanto numerose tracce biologiche tendono ad andare in contro a degradazione proteica in tempi brevi. Il DNA invece è una molecola molto stabile e presente in qualsiasi cellula del corpo, per cui ogni tipo di cellula nucleata dell'individuo può permettere l'identificazione.

1.1 Analisi del DNA

Nel 1990 parte in America il progetto nazionale *Human Genome Project* (HGP) come collaborazione tra gli enti *American Department of Energy e National Institute of Health* (NIH). Il progetto si propone di sequenziare 3.3 miliardi di basi del genoma umano entro 15 anni. Il costo previsto era di 15 miliardi di dollari. Gli obiettivi alla base di questa impresa erano:

- Identificare tutti i geni del genoma umano (circa 30.000);
- Memorizzare le informazioni in banche dati apposite;
- Approntare programmi per l'analisi dei dati;
- Affrontare i problemi etici, legali e sociali legati al progetto.

Nel 1998 entra in gioco anche una compagnia privata, la *Celera Genomics*, che introduce una nuova strategia di sequenziamento chiamata *Whole Genome Shotgun* (WGS), molto efficiente, rapida e facilmente automatizzabile.

Nel 2001 si conclude lo *Human Genome Project*.

Dal sequenziamento del genoma umano è emerso che solo l'1,5% del DNA è composto da regioni codificanti, mentre il restante 98,5% è rappresentato da sequenze non codificanti, ripetute o in singola copia (2).

È importante sottolineare però che la maggior parte del materiale genetico, circa il 99,5% del DNA non cambia da individuo a individuo. Infatti solo una piccolissima porzione del nostro genoma è soggetto a variabilità interindividuale. Questa minima variabilità è la causa dell'unicità di ogni individuo ed è proprio su queste regioni che puntano le analisi di genetica forense.

1.2 I polimorfismi del DNA in genetica forense

Le informazioni genetiche di ogni individuo sono depositate nella sequenza del DNA. In particolare l'uomo possiede il DNA nucleare, composto da 23 coppie di cromosomi (22 coppie di autosomi e 1 coppia di cromosomi sessuali) e il DNA mitocondriale (composto da una molecola di DNA circolare) (*Fig.1*).

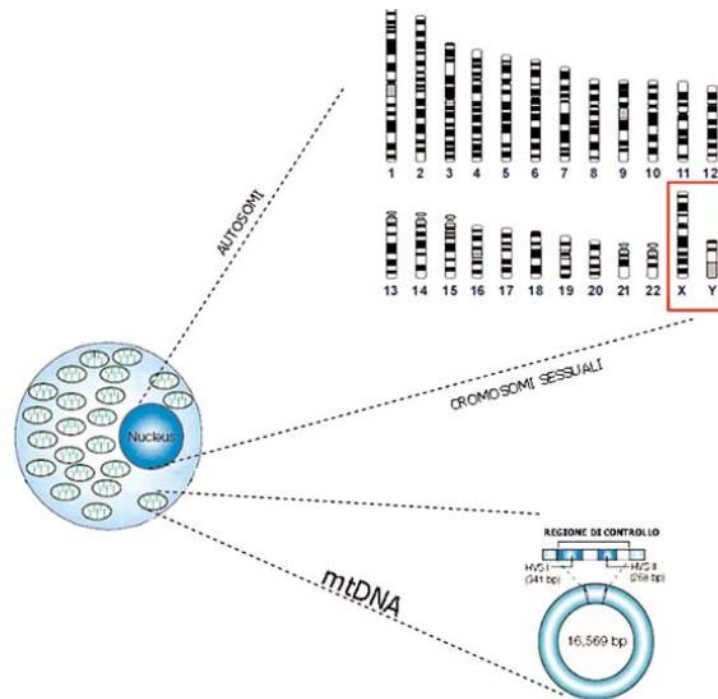


Fig.1 Patrimonio genetico della cellula eucariote (maschile): DNA nucleare, composto da 22 autosomi e da 2 cromosomi sessuali (X e Y), e DNA mitocondriale (3).

Le regioni studiate in genetica forense vengono chiamate polimorfismi.

Un polimorfismo del DNA viene definito come la presenza, nella popolazione di una specie, di più alleli di uno stesso gene; fisicamente un polimorfismo è rappresentato da una mutazione puntiforme per sostituzione, per inserzione o per delezione. Per convenzione, si considerano polimorfismi genetici solo le forme che si presentano con una frequenza maggiore dell'1%. Nel caso in cui la frequenza sia inferiore si parla di variante rara.

I polimorfismi sono dispersi su tutto il genoma (autosomi, cromosomi sessuali e DNA mitocondriale). Esistono due tipologie di polimorfismi del DNA: i polimorfismi di sequenza e i polimorfismi di lunghezza.

I polimorfismi di sequenza vengono chiamati anche SNPs (*Single Nucleotide Polymorphism*) e sono formati da una sostituzione nucleotidica che determina la differenza tra due sequenze di DNA omologhe. Solitamente gli SNPs producono polimorfismi biallelici e forniscono un'informazione limitata per l'identificazione personale. È possibile quindi andare a studiare una serie di SNPs per avere una conoscenza più completa del profilo personale.

Nella specie umana esistono anche dei polimorfismi triallelici o tetrallelici ma hanno una frequenza molto bassa nella popolazione.

La seconda tipologia di polimorfismo, quello di lunghezza, è costituito da un numero variabile di sequenze di DNA ripetute in *tandem*. Questi polimorfismi portano quindi a diverse forme multialleliche, ciascuna definita dalla lunghezza variabile della sequenza che è determinata dal numero di ripetizioni.

In base alla lunghezza della sequenza ripetuta è possibile suddividerli in minisatelliti, o VNTRs (*Variable Number of Tandem Repeats*), e microsatelliti, o STRs (*Short Tandem Repeats*).

I minisatelliti sono costituiti da circa 20-50 ripetizioni di una sequenza lunga dalle 10 alle 100 bp (*Fig.2A*). Mentre i microsattelliti sono brevi sequenze di 2-6 bp ripetute da 5 a oltre 40 volte (*Fig.2B*).

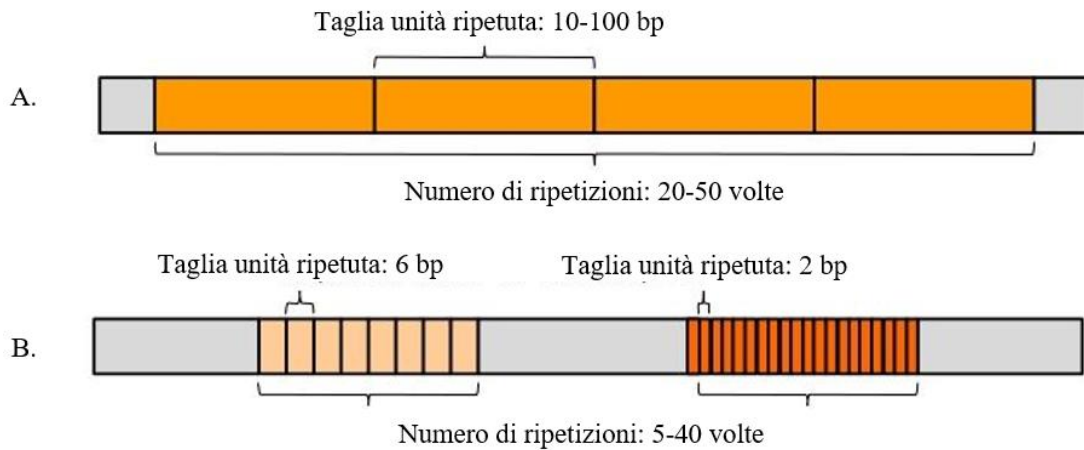


Fig.2 Esempio di minisatellite (A) ed esempio di microsattellite (B).

Attualmente gli STRs sono i marcatori genetici comunemente utilizzati nella routine dei laboratori di genetica forense. La scelta ricade sugli STRs poiché essi possiedono diverse caratteristiche distintive. La caratteristica più importante è che si tratta di marcatori multiallelici. A differenza di quanto accade per gli SNPs, si hanno più forme alleliche, determinate dal numero di ripetizioni dell'unità ripetuta. Questo determina la loro elevata capacità discriminativa.

Inoltre gli STRs sono amplificabili in *PCR multiplex*, con altri *markers*. Questi marcatori possiedono bassa probabilità di produrre picchi *stutter* e basso tasso mutazionale per generazioni. Per picchi *stutter* si intendono dei

picchi elettroforetici che si localizzano appena prima (o dopo) del reale picco allelico e che rappresentano degli ampliconi che sono formati da una ripetizione in meno (o in più) dell'allele realmente presente. Questo picco è determinato da uno slittamento enzimatico durante la PCR (*Fig.3*).

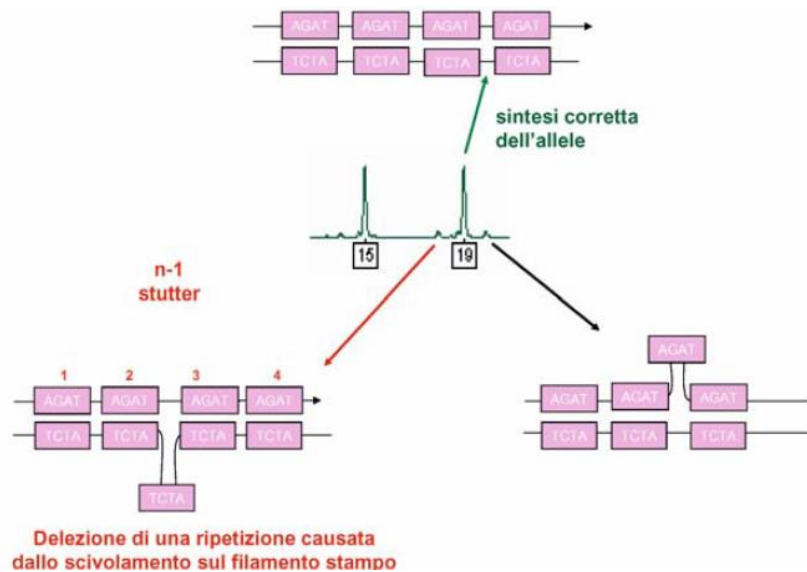


Fig.3 Meccanismo di formazione delle stutter (3).

1.3 STRs per l'identificazione personale

Nel 1990 l'FBI (*Federal Bureau Investigation*) sponsorizza il primo progetto per l'identificazione di un set di STRs utili per la creazione di un database di DNA personale noto come CODIS (*Combined DNA Index System*) (4).

Vengono selezionati 13 *loci* che presentano un elevato potere di identificazione personale, cioè CSF1PO, D3S1358, D5S818, D7S820,

D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, vWA e locus Amelogenina, per l'identificazione personale (5). Dal gennaio 2017 sono stati aggiunti altri loci: D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045 (3, 4 e 6) (*Fig.4 e Tab.1*).

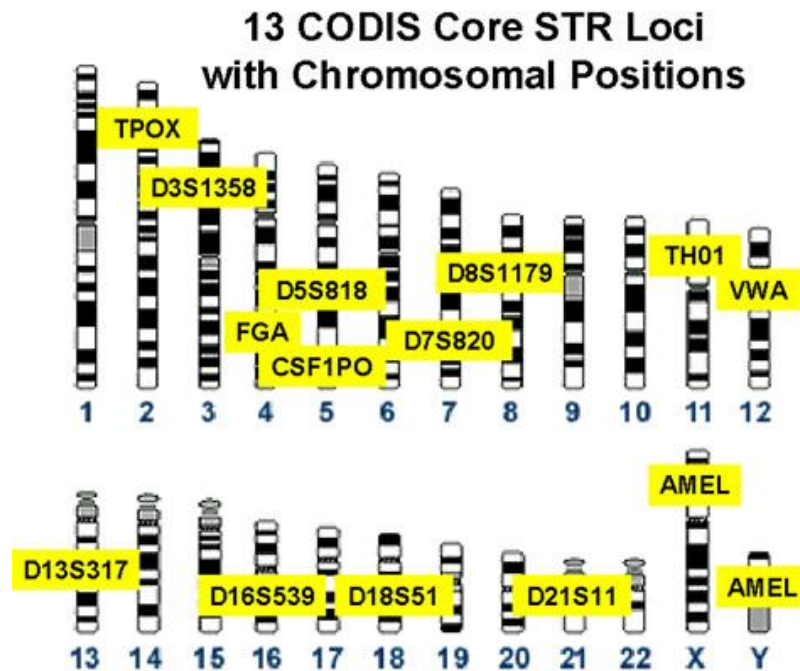


Fig.4 Loci del CODIS.

NOME LOCUS	CODICE UniSTS	LOCALIZZAZIONE CROMOSOMICA	POSIZIONE FISICA GRCh37	SEQUENZA RIPETUTA	GENBANK ACCESSION (NUMERO ALLELE)	RANGE ALLELICO
D1S1656	58809	1q42	Chr 1 230.905 Mb	TAGA	G07820 (15.3)	8-20.3
TPOX	240638	2p25.3 (perossidasi tiroidea, 10° introne)	Chr 2 1.493 Mb	AATG	M68651 (11)	4-16
D2S441	71306	2p14	Chr 2 68.239 Mb	TCTA/TCAA	AC079112 (12)	8-17
D2S1338	30509	2q35	Chr 2 218.879 Mb	TGCC/TTCC	AC010136 (23)	10-31
D3S1358	148226	3p21.31	Chr 3 45.582 Mb	TCTG/TCTA	AC099539 (16)	6-26
FGA	240635	4q31.3 (alfa-fibrinogeno, 3° introne)	Chr 4 155.509 Mb	CTTT/TTCC	M64982 (21)	12.2-51.2
D5S818	54700	5q23.2	Chr 5 123.111 Mb	AGAT	AC008512 (11)	4-29
CSF1PO	156169	5q33.1 c-fms proto-oncogene, 6° introne	Chr 5 149.455 Mb	AGAT	X14720 (12)	5-17
SE33	ACTBP2	6q14 Pseudogene correlto β - actina	Chr 6 88.987 Mb	AAAG	V00481 (26.2)	3-42
D7S820	74895	7q21.11	Chr 7 83.789 Mb	GATA	AC004848 (13)	5-16
D8S1179	83408	8q24.13	Chr 8 125.907 Mb	TCTA/TCTG	AF216671 (13)	6-20
D10S1248	51457	10q26.3	Chr 10 131.093 Mb	GGAA	AL391869 (13)	7-19
TH01	240639	11p15.5 tirosina idrossilasi, 1° introne	Chr 11 2.192 Mb	TCAT	D00269 (9)	3-14
vWA	240640	12p13.31 fattore di von Willebrand, 40° introne	Chr 12 6.093 Mb	TCTG/TCTA	M25858 (18)	10-25
D12S391	2703	12p13.2	Chr 12 12.450 Mb	AGAT/AGAC	G08921 (20)	13-27.2
D13S317	7734	13q31.1	Chr 13 82.692 Mb	TATC	AL353628 (11)	5-17
Penta E	/	15q26.2	Chr 15 97.374 Mb	AAAGA	AC027004 (5)	5-32
D16S539	45590	16q24.1	Chr 16 86.386 Mb	GATA	AC024591 (11)	4-17
D18S51	44409	18q21.33	Chr 18 60.949 Mb	AGAA	AP001534 (18)	5.3-40
D19S433	33588	19q12	Chr 19 30.416 Mb	AAGG/TAGG	AC008587 (14)	5.2-20
D21S11	240642	21q21.1	Chr 21 20.554 Mb	TCTA/TCTG	AP000433 (29.1)	12-43.2
Penta D	/	21q22.3	Chr 21 45.056 Mb	AAAGA	AP001752 (13)	1.1-19
D22S1045	49680	22q12.3	Chr 22 37.536 Mb	ATT	AL022314 (17)	7-20

Tab.1 Tabella con i loci per l'identificazione personale del CODIS.

1.4 SNPs in genetica forense

Per SNPs si intendono delle variazioni di un singolo nucleotide presenti in determinati tratti di DNA. I segmenti di DNA da amplificare per studiare gli SNPs sono molto brevi, quindi l'uso di questi *markers* permette l'analisi anche su DNA molto degradato.

È importante notare che esistono variazioni notevoli tra le varie popolazioni umane. Uno SNP molto comune in un determinato gruppo etnico può essere molto raro in un'altra popolazione. Lo studio degli SNPs è molto utile poiché queste piccole variazioni possono influenzare diversi aspetti clinico-patologici. In ambito forense sono molto informativi, sia per la determinazione d'identità personale, per gli studi di parentele (*Fig.5*), per la determinazione dell'origine biogeografica, sia per studi fenotipici (6).

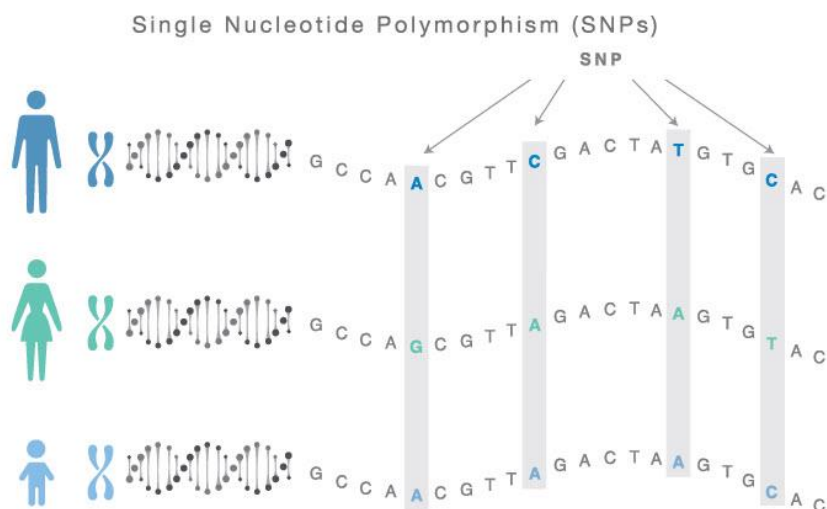


Fig.5 Rappresentazione grafica della trasmissione degli SNPs dai genitori ai figli.

Diversi studi hanno identificato un numero di SNPs utili per i test di identità umana. Inoltre gli SNPs possono servire come *markers* supplementari quando i risultati STRs non sono abbastanza definitivi.

Per quanto riguarda le indagini di parentela, il basso tasso mutazionale degli SNPs, rispetto agli STRs, mostra che essi possono essere molto utili quando si ottengono risultati ambigui con test basati sugli STRs. In alcuni casi SNPs, vicini tra loro, possono muoversi in blocco attraverso la trasmissione generazionale. Secondo lo stesso principio, gli SNPs possono essere usati per le indagini di ascendenza biogeografica. Il tasso mutazionale degli SNPs è dell'ordine di 10^8 generazioni, mentre i tassi di mutazione degli STRs sono circa uno su mille. È per questo che gli SNPs sono i *markers* migliori per la predizione di etnia (7).

Gli SNPs hanno mostrato buoni risultati anche negli studi fenotipici. Analizzando il profilo del DNA dell'individuo colpevole è possibile risalire ad alcuni aspetti fenotipici che possano aiutare a scoprire la sua identità. È stato sviluppato un *tool* informatico chiamato *Iris-Plex* per aiutare a prevedere colore degli occhi (8). Recentemente è stato ampliato diventando *HIris-Plex-S*, con l'aggiunta di SNPs capaci di prevedere anche colore dei capelli e della pelle (9).

Altre caratteristiche che individuano gli SNPs come buoni *markers* nella genetica forense sono le seguenti:

- Producono prodotti di PCR molto brevi, anche minori di 100 bp, per cui sono molto appropriati per l'analisi dei campioni degradati;
- Non producono picchi *stutter*;
- Presentano una maggiore potenzialità di essere amplificati con *multiplex PCR*;
- Presentano fasi di processazione e analisi del campione altamente automatizzabili (6).

1.5 Microplotipi del DNA

I microplotipi del DNA sono un nuovo tipo di marcatore genetico utilizzato nella genetica forense e negli studi di popolazione. In particolare essi sono dei *locus* in cui sono presenti due o più SNPs all'interno di una breve sequenza di DNA (circa 200-300 bp). I microplotipi o MH possono essere considerati come un unico marcatore poiché gli SNPs che li compongono non subiscono ricombinazione e vengono ereditati in blocco. L'ampiezza limitata (circa 200 bp) della regione genomica in cui mappano gli SNPs, che costituiscono un

microplotipo, è una delle caratteristiche fondamentali che li rende utili marcatori.

Lo scopo dello studio dei microplotipi in ambito forense è funzionale per l'identificazione individuale, per la determinazione di parentele e per la discriminazione di più contributori in un campione misto.

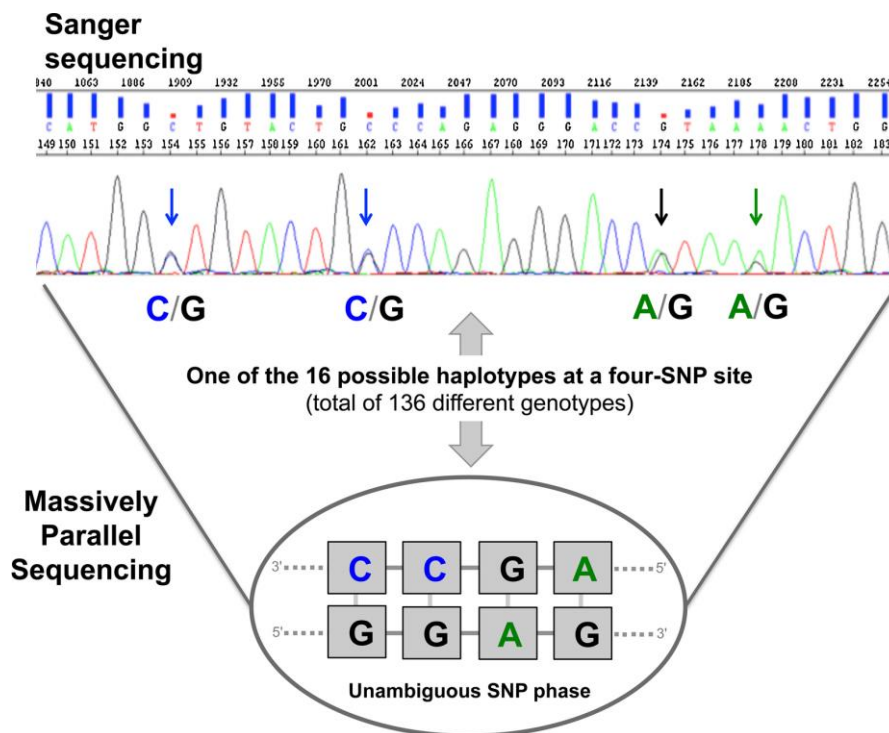


Fig.6 Sequenziamento di Sanger e Massive Parallel Sequencing (10).

Come mostrato in Fig.6, la tecnica migliore per analizzare i microplotipi è il sequenziamento del DNA tramite tecnologie *Massive Parallel Sequencing* (MPS). L'utilizzo delle tecniche MPS risulta essere altamente informativo nello caso dei microplotipi. In questo modo infatti non c'è bisogno di ricostruire le fasi *cis/trans* degli SNPs che compongono un singolo

microplotipo. La regione viene sequenziata in un'unica *reads*, il che permette di individuare e distinguere bene gli alleli relativi ad ogni SNPs, senza bisogno di *tools* predittivi. Questo avviene invece quando le tipizzazioni dei microplotipi vengono eseguite tramite tecniche in *Real-Time PCR* (11). Infatti in questo caso sono necessari dei *software*, come *PHASE*, per ricostruire le fasi ricomponendo la sequenza originale.

1.5.1 L'introduzione dei microplotipi del DNA

Il termine microplotipi del DNA venne introdotto dal Professor Kenneth Kay Kidd, Dipartimento di Genetica dell'Università di Yale. Nel suo articolo edito nel 2013 vengono descritti come:

“two or more SNPs within a molecular distance less than the read length of the desktop machines like the Personal Genome Machine (PGM)”

Kidd suggerisce che la regione genomica entro cui devono trovarsi gli SNPs di un microplotipo si aggira tra le 200-300 bp (12).

I microplotipi hanno un grado di eterozigosi globale molto alto, adatto per essere utilizzati nell'ottenimento di informazioni di discendenza e di relazioni di parentela tra individui (12).

Studi precedenti avevano già introdotto termini che descrivevano polimorfismi vicini ereditati in blocco. Negli anni '60 il Professor Roberto Ceppellini parlava di aplotipi, associati all'ereditarietà in blocco di alleli codificanti per *Human Leucocyte Antigen* (HLA) (1).

Successivamente, con l'inizio dello *Human Genome Project* (HGP), si ebbe un'ulteriore conferma della presenza di sequenze che vengono ereditate a blocchi, come struttura unica (13, 14 e 15).

1.5.2 Nomenclatura dei microaplotipi

Inizialmente quando si parlava di microaplotipi si utilizzavano i nomi degli SNPs implicati, utilizzando il numero rs di ognuno. Ovviamente si tratta di un metodo molto lungo e scomodo da utilizzare.

I primi lavori che nominavano i microaplotipi utilizzavano metodi di codifica diversi in base al laboratorio di origine.

Nel 2016 Kenneth Kidd propose un metodo standardizzato per la nomenclatura dei microaplotipi (16). Il metodo di nomenclatura proposto segue le raccomandazioni della *HUGO Gene Nomenclature Committee* (17). Tra i suggerimenti vi è l'utilizzo di un codice che abbia come radice uno o più

caratteri uguali per tutti i microplotipi. In particolare la nomenclatura proposta da Kidd nel 2017 è mostrata in *Fig.7*.

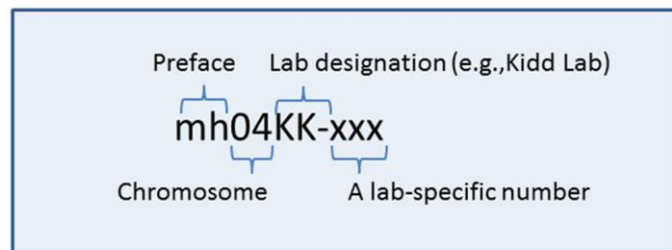


Fig. 7 Illustrazione grafica delle regole di nomenclatura proposte da Kenneth Kidd (16).

Le prime due lettere (mh o MH) identificano il codice come un microplotipo.

I due numeri successivi identificano il numero del cromosoma su cui si trova il microplotipo (da 01 a 22, per gli autosomi, e 0X o 0Y, per i cromosomi sessuali).

Il codice da 2 a 4 lettere che segue (es. KK) identifica il laboratorio in cui il microplotipo è stato scoperto e caratterizzato.

Il codice termina con un numero a tre cifre che rappresenta la combinazione di due o più SNPs. Quel codice è specifico per quel determinato microplotipo.

Facciamo un esempio con il microplotipo mh14KK-048. Esso identifica un microplotipo (“mh”) presente sul cromosoma 14 (“14”) studiato dal

laboratorio di Kenneth Kidd (“KK”). In particolare il codice finale “048” identifica il microplotipo composto da due SNPs: rs12717560 e rs12878166.

Il *database online ALFRED (ALlele FREquency Database)* racchiude la nuova nomenclatura dei microplotipi e tutte le caratteristiche di ognuno, compreso i codici identificativi utilizzati in precedenza e le frequenze alleliche su campioni di popolazione umana (18 e 19).

1.5.3 Vantaggi e svantaggi dell’utilizzo dei microplotipi

Nella pratica forense routinaria gli STRs identificati dal CODIS vengono usati per l’identificazione individuale; gli SNPs possono aumentare il potere discriminativo a causa del più basso tasso di mutazioni (10^{-8} - 10^{-6} per generazione) rispetto allo stesso valore relativo agli STRs (10^{-3} per generazione) (20).

Inoltre l’utilizzo di SNPs biallelici su regioni brevi aumenta molto la probabilità di identificazione individuale e, secondo recenti studi, permette anche l’acquisizione di informazioni più dettagliate, come la popolazione di provenienza (21) e in alcuni casi anche dei tratti fenotipici (22).

L’utilizzo di microplotipi nelle indagini forensi favorisce diversi aspetti.

La breve distanza tra gli SNPs che caratterizzano un microaplotipo implica un valore estremamente basso del tasso di ricombinazione.

Inoltre, dato che i microaplotipi sono multi-SNPs, possono fornire un insieme di informazioni più dettagliate e precise rispetto a un singolo SNP (10).

È importante sottolineare che l'assenza di una breve sequenza ripetuta in *tandem* (come negli STRs), impedisce lo slittamento della DNA polimerasi, che causa frammenti più brevi chiamati *stutter fragments*. Come già detto, si tratta semplicemente di frammenti di DNA in cui l'enzima è stato accecato dalla ripetizione della stessa sequenza, producendo un frammento con una ripetizione in meno rispetto all'originale. A differenza degli STRs, i microaplotipi non generano picchi *stutter* (10).

Un altro vantaggio dei microaplotipi rispetto agli STRs è che tutti gli alleli di un *locus* hanno la stessa lunghezza. Gli alleli STRs possono variare molto in lunghezza nucleotidica. Se il DNA è degradato o se il campione contiene inibitori della PCR, gli alleli più brevi possono essere preferibilmente amplificati rispetto agli alleli più lunghi dello stesso contributore, complicando la deconvoluzione del campione misto. Questo non avviene nel caso dei microaplotipi, i quali mantengono inalterata la lunghezza del frammento (11).

Uno svantaggio dei microplotipi rispetto agli STRs è che generalmente hanno meno alleli e di conseguenza è necessario un pannello più grande di marcatori per ottenere un potere di discriminazione comparabile a quello degli STRs.

Per quanto riguarda i microplotipi, le frequenze alleliche variano significativamente tra le popolazioni; in questo modo, le frequenze alleliche specifiche della popolazione sono necessarie per il calcolo della *Random Match Probability* (RMP) o il *Likelihood Ratio* (LR) in quanto questi valori cambieranno tra le popolazioni.

Un altro svantaggio della rilevazione dei microplotipi sono sicuramente le operazioni di laboratorio. Quando si parla di microplotipi, il metodo d'elezione è sicuramente la *Massive Parallel Sequencing*, tecnologia costosa e laboriosa da preparare. La procedura per l'individuazione degli alleli STRs è invece basata su una più semplice ed economica elettroforesi capillare (10).

1.5.4 Criteri per la selezione dei microplotipi

I microplotipi risultano essere altamente informativi per molte indagini forensi. In base a dei criteri specifici si possono scegliere i microplotipi migliori e più informativi per l'analisi.

Per la selezione dei microaplotipi ottimali vengono valutati due valori importanti: numero effettivo di alleli (A_e) e informatività (I_n).

Il primo valore (A_e) viene descritto come:

“effective number of alleles represents the number of equally frequent alleles that would generate the same heterozigosity as the locus with multiple alleles at very different frequencies” (24)

È importante valutare questo parametro poiché in questo modo si vede la variazione nel *locus* all'interno di una popolazione specifica e inoltre è associato alla capacità del marcatore di distinguere gli individui. Pertanto i microaplotipi con elevati valori di A_e sono correlati con una maggiore capacità di discriminare un individuo dall'altro, di rilevare campioni misti di DNA, di determinare il numero di contributori e di ricostruire le linee parentali (23).

Invece il valore I_n indica una misura della variazione di frequenze alleliche (aplotipo) osservata tra le diverse popolazioni testate (25). Di conseguenza, valori I_n più alti indicano che un microaplotipo ha alleli che sono frequenti in alcune popolazioni e rari in altre. Microaplotipi con alti valori di I_n possono aiutare a distinguere tra alcune delle popolazioni selezionate e prevedere la relativa ascendenza biogeografica.

Il *locus* ideale avrà valori alti sia di A_e che di I_n , il che significa che ogni popolazione presenta molti alleli e alcuni (o anche tutti) variano in frequenza tra le popolazioni.

1.5.5 Applicazioni di microaplotipi in genetica forense

I microaplotipi sono un nuovo strumento che può essere utilizzato per ottenere utili informazioni a scopo forense.

I set di microaplotipi possono essere ottimizzati in modo da risultare utili per: identificazione individuale, relazioni tra individui, informazioni sul fenotipo, informazioni sull'origine biogeografica, deconvoluzione di campioni misti, riconoscimento di cadaveri nei disastri di massa, identificazione di persone scomparse e identificazione DNA non umano (*Fig.8*) (26).

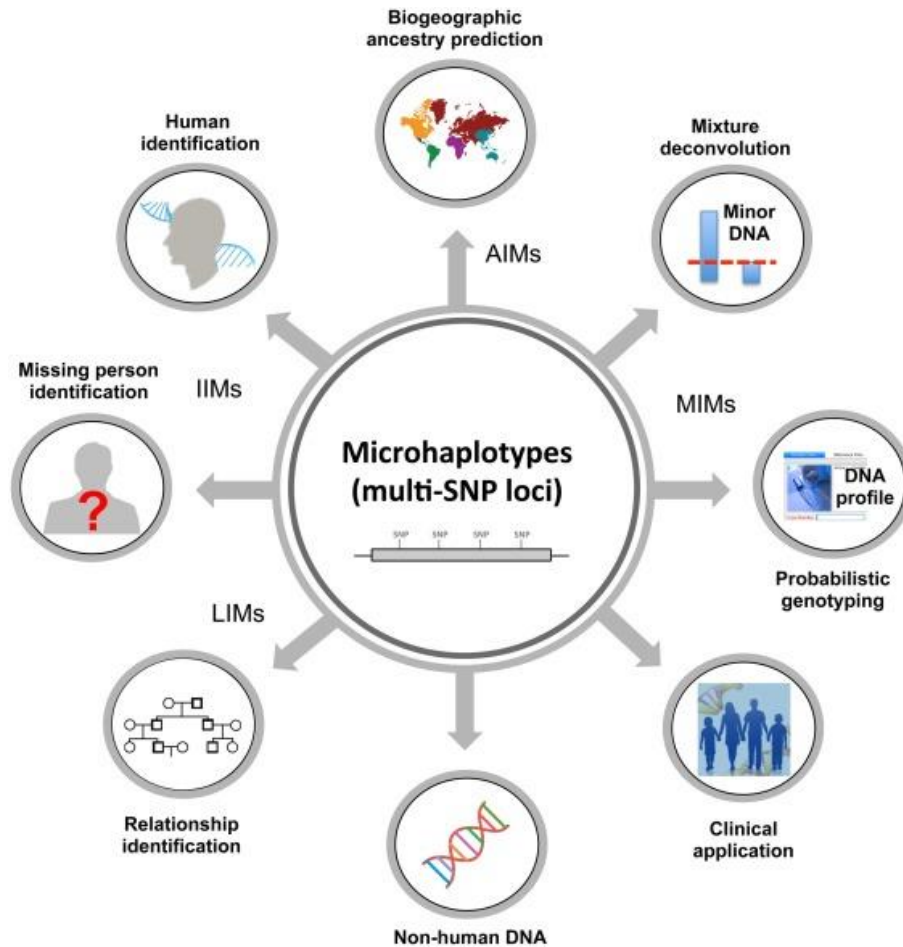


Fig. 8 Panoramica delle potenziali applicazioni dei microplotipi in campo forense (10).

Microplotipi nell'identificazione personale

Recenti studi dimostrano la capacità e la sensibilità delle tecniche MPS, utilizzando la piattaforma *Ion S5™* (27). In particolare vengono indagati 36 microplotipi comparando la tecnica MPS e l'elettroforesi capillare (CE). I risultati dimostrano come i microplotipi abbiano grande potenzialità nell'identificazione umana. L'applicazione in MPS di pannelli di

microplotipi ha prodotto un maggior numero di alleli per *locus* rispetto alle stesse analisi effettuate con CE, suggerendo la maggiore sensibilità associata alle tecniche di *Next Generation Sequencing* (27).

Precedenti valutazioni dei microplotipi, eseguite da Kidd su un set di 31 *loci*, forniscono la prova della loro utilità per l'identificazione individuale, l'origine biogeografica e la definizione delle parentele (11).

Un approccio alternativo è stato recentemente presentato da *Bose et al.* in uno studio in cui viene valutato l'utilizzo del MPS per la tipizzazione di campioni di DNA degradati e misti (28). È stato utilizzato un set composto da 411 *loci*: 375 SNPs e 36 microplotipi, composti da 76 SNPs. Questo set di *markers* genetici è stato selezionato per l'identificazione individuale, le relazioni di parentela e deconvoluzione di campioni misti. Complessivamente il saggio è risultato efficiente nella tipizzazione di campioni degradati artificialmente (da 200 bp a 75 bp). Ne consegue la possibile applicazione su campioni di DNA degradati comuni nella pratica forense e, di conseguenza, utili nei casi di identificazione personale (28).

Microplotipi nell'identificazione di parentele

L'identificazione di relazioni familiari è stata sempre eseguita utilizzando *locus* di SNPs multiallelici. La grande varietà di alleli e delle loro

combinazioni tende ad abbassare notevolmente la probabilità che due individui condividano lo stesso profilo. Quindi, maggiore sarà l'eterozigosi di un *locus* genico e maggiore sarà la probabilità che tale *locus* non sia comune nella popolazione. È però probabile che gli stessi alleli vengano osservati tra parenti stretti.

Un parametro molto utile per l'identificazione di parentela (ma anche per l'inferenza ancestrale) è l'*identity by descendent* (IBD). Si parla di IBD quando si ha un segmento di DNA più o meno lungo condiviso da due individui imparentati. La lunghezza del segmento dipende dal numero di generazioni che passano tra i due individui e il loro antenato comune per quel determinato *locus* condiviso. Minore è il numero di segmenti condivisi, maggiore è la distanza in generazioni dall'antenato comune per questo *locus*.

Pannelli di qualche centinaia di SNPs sono sufficienti per avere un grado di predizione di parentela tra due individui (29).

Per il pannello di 130 microaplotipi proposto da Kidd (30), solo 28 *loci* mostrano $A_e \geq 3.0$ e possono quindi essere di grande valore per l'identificazione di parentela.

Presto le ricerche si concentreranno anche sullo studio di microaplotipi per la corretta identificazione familiare utilizzando veri e propri *casework sample*,

per verificare se si ha lo stesso valore identificativo anche utilizzando campioni di DNA degradato (30).

Microplotipi nella deconvoluzione di campioni misti

È stato recentemente utilizzato un pannello a 74 microplotipi sulla piattaforma *Ion S5™* per testare la capacità dei MH di deconvoluzione di campioni misti (31). I risultati hanno confermato la possibilità di identificare e distinguere i profili di due persone in campioni misti, anche in presenza di squilibri allelici tra i due soggetti. Inoltre è stata indicata anche la possibilità di predire l'ascendenza biogeografica degli stessi individui (31).

Diversi autori hanno puntato l'attenzione su questo argomento. Nel 2018, sono stati prodotti in vitro dei campioni misti di DNA derivati da due soggetti (in proporzioni diverse). I risultati del test hanno messo in evidenza che, utilizzando un set di 26 microplotipi con valore di $A_e=3$, la probabilità di rilevare una miscela di due individui non imparentati raggiunge lo 0.9999998, (32).

Infine, *Voskoboinik et al.* (33) ha proposto l'uso di microplotipi altamente polimorfici ($A_e>15$), sempre con lo scopo di identificare campioni di DNA misto. Il team di ricerca di Voskoboinik ha selezionato un set di soli 10

microaplotipi contenenti più di 10 SNPs. Gli esiti ottenuti dalle prove indicano che le miscele complesse possono essere interpretate in modo accurato anche in presenza di basso *coverage* e elevati tassi di errore (33).

Microaplotipi nell'ascendenza biogeografica

Numerosi studi testimoniano che lo studio degli SNPs permette l'individuazione dell'ascendenza biogeografica degli individui. A riguardo sono stati creati dei pannelli di SNPs appositi, chiamati *AISNP panels* (*Ancestry Informative Single Nucleotide Polymorphism*) (21, 34 e 35).

I microaplotipi sono formati da due o più SNPs, quindi come questi ultimi, permettono la determinazione dell'origine biogeografica. Tra i valori più significativi dei microaplotipi abbiamo I_n (*informativeness*). Questo valore misura la variazione della frequenza allelica nelle diverse popolazioni, quindi *loci* con alto I_n mostrano una maggiore variazione nella popolazione. È importante sottolineare che il valore di I_n varia in base alla popolazione considerata. Infatti alcuni microaplotipi sono specifici solo di alcune popolazioni. Diversi studi lo confermano (35, 36, 37 e 38).

Microaplotipi in genetica medica

Un'applicazione non forense per cui sono stati proposti i microaplotipi ricade nel campo del trapianto di cellule staminali ematopoietiche. Prima di sottoporsi a trapianto i pazienti sono sottoposti a un trattamento chiamato condizionamento mieloablativo. Tale procedura che consiste nel ridurre le cellule staminali ematopoietiche del paziente. Lo scopo di questo passaggio è:

- Sopprimere il sistema immunitario del ricevente prevenendo il rigetto del nuovo midollo osseo;
- Creare spazio nel midollo osseo per far crescere le cellule staminali del donatore;
- Uccidere le cellule cancerogene.

Dopo il trapianto i pazienti otterranno un *pattern* di cellule midollari a metà tra se stessi e donatore.

Per determinare l'efficacia del trapianto e valutare il potenziale rigetto, viene eseguito il test del chimerismo studiando i convenzionali STRs dal sangue del ricevente (39 e 40).

Il rapporto tra alleli del donatore e quelli del paziente viene valutato nel tempo a intervalli stabiliti e viene utilizzato per formulare una prognosi e

sviluppare un piano di trattamento in caso di rigetto. La diagnosi precoce del rigetto migliora la prognosi, consentendo l'intervento precoce (41).

Allo stesso modo, sono state suggerite applicazioni non forensi dei microplotipi per il test parentale non invasivo di cellule fetali (e DNA libero da cellule) che circolano nel sangue materno (42).

Un'ulteriore promettente applicazione di microplotipi è per il DNA non umano. Alcuni autori hanno riportato l'uso di un pannello di microplotipi per determinare l'origine di alcuni pesci e alghe di mare.

Il pannello identificato ha mostrato un grado di eterozigosi notevolmente superiore rispetto quello ottenuto con gli SNPs. L'uso dei microplotipi potrebbe essere esteso ad altre specie e piante.

2 - LA GENETICA FORENSE:

LIMITI E POTENZIALITÀ

Le indagini genetiche forensi sono complesse perché molto spesso le analisi si basano su quantità esigue di campioni o prelievi non ripetibili, quindi occorre avere a disposizione sistemi in grado di dare informazioni utili a partire da piccole quantità di DNA.

Per la fase preliminare di estrazione del DNA dal campione esistono kit commerciali che riducono notevolmente i tempi tecnici e le attività manuali, a volte a scapito però della resa finale. Inoltre, a differenza delle tecniche tradizionali, i nuovi metodi di estrazione non includono l'utilizzo di reagenti chimici pericolosi, come fenolo e cloroformio che sono invece alla base della tecnica classica di estrazione del DNA (3). Per la corretta esecuzione dell'analisi è necessario introdurre anche un bianco d'estrazione, importante per verificare l'assenza di contaminanti esterni nei reattivi usati per l'estrazione.

Esistono prodotti commerciali specifici per la quantizzazione del DNA che si basano sulla *Real-Time PCR*. Questi sono molto utili in ambito forense poiché permettono di amplificare e, allo stesso tempo, di quantizzare ridotte quantità di DNA.

Per quanto riguarda la fase di amplificazione sono in commercio numerosi kit che, tramite tecniche *PCR multiplex*, vanno ad analizzare più *loci* contemporaneamente e con alto grado di sensibilità. Diverse ditte specializzate hanno infatti prodotto kit che consentono l'amplificazione contemporanea dei microsatelliti del CODIS, in poco tempo e partendo da meno di 1 ng di DNA stampo. Recentemente questi kit sono stati perfezionati e implementati con l'inserimento di altri microsatelliti per aumentare ulteriormente il potere informativo dell'analisi.

Ovviamente è fondamentale allestire un bianco anche durante l'amplificazione, che spesso viene chiamato controllo negativo. Si tratta di una provetta che in cui viene allestita una reazione di PCR come per gli altri campioni, ma in questo caso al posto dell'estratto di DNA viene aggiunta dell'acqua nucleasi-free. Questo bianco serve per verificare che non ci siano contaminazioni nei reagenti usati per l'amplificazione.

Per quanto riguarda la fase di elettroforesi capillare, le ditte forniscono un *ladder* allelico insieme al kit commerciale d'elettroforesi, per la corretta genotipizzazione del campione. Per *ladder* allelico (*allelic ladder*) si intende una miscela artificiale di tutti gli alleli più comuni presenti nella popolazione, prodotto con gli stessi *primers* presenti nel kit e che serve da riferimento per l'assegnazione allelica del campione, oltre a uno standard di lunghezza (*size*

standard), ovvero una miscela di frammenti di lunghezza nota, colorati con un fluorocromo differente, che viene fatto correre in elettroforesi insieme al campione. Il *size standard* è necessario per attribuire a ogni punto del tracciato elettroforetico una lunghezza espressa in paia di basi (3).

Nel caso in cui si vadano a analizzare campioni con esigue quantità di DNA, è possibile riscontrare delle problematiche a livello dei picchi elettroforetici. In particolare si possono ottenere i seguenti artefatti: *allele drop-in*, *allele drop-out*, *locus drop-out*, sbilanciamento dell'eterozigosi e formazione di *stutter*.

Nel caso di *allele drop-in* si ha la rilevazione di un nuovo allele. Si tratta della sintesi, durante la reazione di PCR, di uno o più alleli diversi da quelli biologicamente presenti. Questo errore può essere dovuto alla fase iniziale di amplificazione o a una contaminazione (Fig.9).

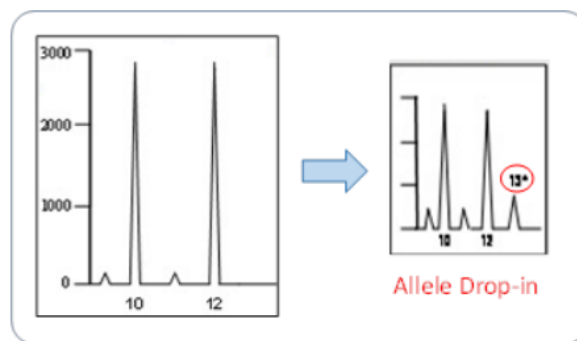


Fig.9 Drop-in allelico (ADI).

Per quanto riguarda gli individui eterozigoti può verificarsi un fenomeno chiamato *drop-out* allelico (ADO). In questo caso si intende la perdita della rilevazione di uno dei due alleli; si otterrà quindi una condizione di omozigosi

per un allele che in realtà è eterozigote. Questa situazione si verifica di frequente nei campioni degradati, soprattutto in corrispondenza degli alleli ad alto peso molecolare (*Fig.10*).

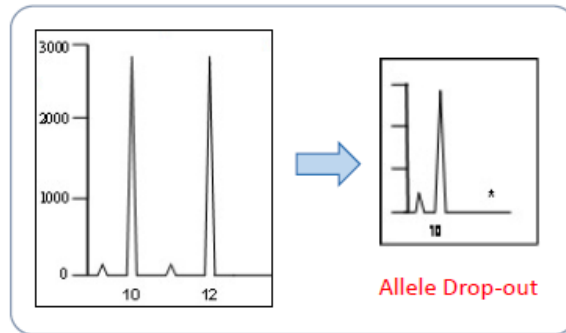


Fig.10 Drop-out allelico (ADO).

Può accadere anche un fenomeno simile che porta alla perdita di entrambi gli alleli dello stesso *locus*. Quindi non verrà assegnato nessun genotipo (omozigote o eterozigote) per il *locus* per cui avviene questo fenomeno (*Fig.11*).

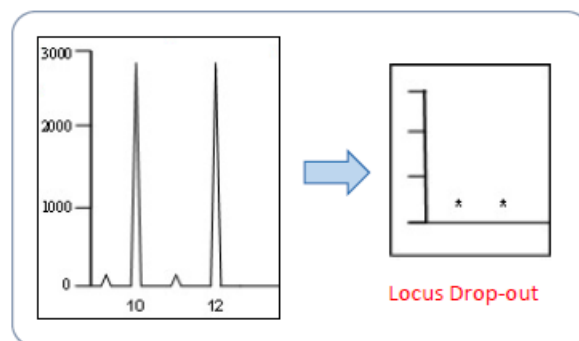


Fig.11 Drop-out del locus.

L'amplificazione preferenziale di uno degli alleli in un *locus* eterozigote determina invece lo sbilanciamento allelico. La sintesi quantitativamente squilibrata degli alleli, durante la reazione di PCR, provoca problemi di interpretazioni dei profili, come mostrato in *Fig.12*.

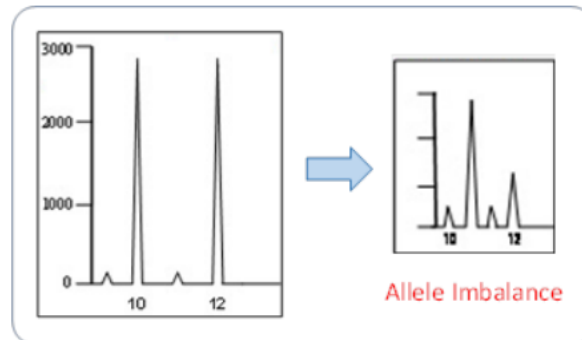


Fig.12 Sbilanciamento allelico.

Infine, quando la quantità iniziale di DNA è bassa, è possibile la formazione di *stutter bands* (SB) o picchi *stutter*. Come già detto, si tratta di artefatti di polimerizzazione durante la PCR. Normalmente, il picco *stutter* è circa il 10% rispetto al picco dell'allele principale (*Fig.13*).

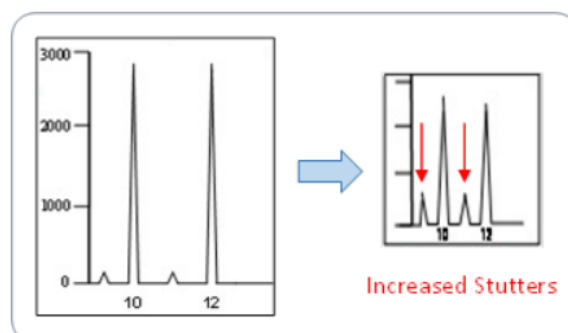


Fig.13 Picchi stutters (indicati in rosso).

Se si dispone di materiale a sufficienza sarebbe bene ripetere le analisi per ridurre al minimo le possibilità di ottenere un profilo non corretto (6).

2.1 Degradazione del DNA

In genetica forense molto spesso si ha a che fare con campioni degradati. Sebbene sia sempre possibile estrarre DNA da un basso numero di cellule, non sempre la qualità del DNA permette di ottenere un profilo sufficientemente informativo.

Il campione può arrivare al laboratorio di genetica in stato già degradato.

La degradazione del DNA può essere determinata da fattori fisici, quali acqua (umidità), luce, ossigeno, temperatura e radiazioni, o da fattori enzimatici, come l'attivazione di nucleasi (45).

Tra i primi agenti degradativi del DNA ci sono le nucleasi endogene. Questo gruppo di enzimi idrolizza i legami fosfodiesterici tra i nucleotidi degli acidi nucleici. Sono distinte in endonucleasi, le quali agiscono all'interno della molecola di DNA, e in esonucleasi, che agiscono alle due estremità della molecola. Ovviamente non solo le nucleasi sono implicate nella degradazione del DNA, infatti, a seguito della morte cellulare, vi è il rilascio di alcuni enzimi lisosomiali, come idrolasi, lipasi e proteasi. Queste ultime agiscono

sulle proteine istoniche, le quali sono alla base della struttura cromatinica. La rimozione degli istoni favorisce l'azione casuale delle nucleasi endogene (43).

La degradazione può anche essere esogena. In alcuni casi infatti il DNA cellulare può venir degradato dagli enzimi dei microrganismi che colonizzano il tessuto. La maggior parte di questi microrganismi, chiamati decompositori, possiede la capacità di decomporre gli acidi nucleici e alcuni loro prodotti di degradazione (44).

Le reazioni enzimatiche sono seguite da processi di degradazione del DNA di tipo spontaneo e non enzimatico, più lenti ma persistenti (45). In primis abbiamo le reazioni idrolitiche. Il legame glicosidico (*Fig.14*) tra la base azotata e il deossiribosio della catena carboniosa del *double strand* è il legame più suscettibile all'idrolisi. Questa reazione porta alla perdita della base azotata.

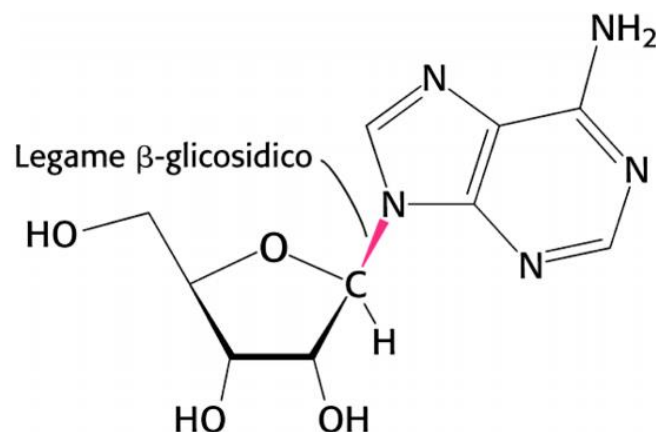


Fig.14 Legame glicosidico tra ribosio (zucchero) e adenina (base azotata).

Un'altra modificazione del DNA è il *crosslinkage*. È possibile infatti la formazione di legami crociati: tra il DNA e proteine o tra lo zucchero con anello aperto, causato dall'idrolisi, e un gruppo amminico situato sullo *strand* opposto. La reazione di *crosslinkage* tra i due filamenti di DNA è temperatura dipendente ed è più lenta a basse temperature (43).

Altra causa importante della degradazione del DNA è l'ossidazione. Le specie derivate dall'ossigeno, come il radicale superossido (O_2^-) e l'acqua ossigenata (H_2O_2), vengono generate a seguito di radiazioni ionizzanti o dal metabolismo di alcuni microrganismi aerobici che colonizzano i tessuti. L'area più suscettibile ai radicali derivati dall'ossigeno è la base azotata. In particolare le pirimidine (soprattutto la timina) sono più sensibili al danno ossidativo rispetto alle purine. Per danno ossidativo si intende l'insieme di processi di ossidazione dei residui di zucchero, ossidazione delle basi e rimozione delle stesse (43).

L'ultima causa che favorisce la degradazione del DNA sono le radiazioni. Esse possono produrre una varietà di lesioni tra cui danno ossidativo, rotture di *strand* singoli e doppi, modifiche di base, distruzione di zuccheri, formazione di legami crociati e formazione di dimeri (43).

Il risultato dell'analisi genetica deriva comunque dalla quantità iniziale di DNA, ma a volte è possibile che l'amplificazione, fase cruciale di tutto il processo, venga intaccata dalla presenza di inibitori della PCR (46).

Esistono molti inibitori della PCR. In *Tab.2* ne vengono riportati alcuni.

INIBITORI DELLA PCR	POSSIBILE FONTE
Eme, emoglobina, lattoferrina, IgG	Sangue
Melanina ed eumelanina	Tessuti epiteliali, capelli e formazioni pilifere
Polisaccaridi	Feci e materiale vegetale
Sali biliari	Feci
Urea	Urine
Proteinasi e ioni calcio	Latte
Colorante tessile (indaco)	Indumenti (blue jeans)
Collagene	Tessuti
Mioglobina	Tessuto muscolare
Ioni calcio	Ossa

Tab.2 Esempi di inibitori di PCR e delle possibili fonti (6).

Tra i più comuni inibitori della PCR presenti nei campioni forensi abbiamo l'emoglobina e alcuni coloranti tessili. Questi inibitori possono:

- Interferire con la fase di lisi cellulare necessaria per l'estrazione del DNA;
- Legarsi all'acido nucleico evitando il legame dell'enzima polimerasi;
- Legarsi all'acido nucleico degradandolo fisicamente;
- Inibire l'attività della polimerasi (47, 48 e 49).

I campioni che contengono degli inibitori della PCR spesso producono dei profili parziali, che assomigliano molto ai profili ottenuti con campioni di DNA degradato.

Esistono comunque diverse strategie che permettono di rendere i campioni compatibili con la PCR anche se contengono inibitori (47 e 49).

2.2 Controlli di qualità e prevenzione della contaminazione

Come qualsiasi analisi scientifica, la tipizzazione del DNA deve essere eseguita con la massima cura. È un processo tecnico che deve essere eseguito da personale qualificato ed efficacemente addestrato per garantire che i risultati ottenuti siano accurati e interpretati correttamente.

Due valori su cui porre particolare attenzione sono: la garanzia della qualità e il controllo di qualità.

La garanzia della qualità (*Quality Assurance* o QA) si riferisce a quelle azioni pianificate necessarie per fornire un'adeguata sicurezza del risultato, ottenuto dalla tecnica applicata.

Il controllo di qualità (*Quality Control* o QC) si riferisce alle tecniche operative routinarie e alle attività impiegate per soddisfare i requisiti di qualità (6).

La procedura preliminare e di base necessaria in ogni laboratorio è l'utilizzo di DPI (dispositivi di protezione individuale) che permettono la protezione dell'operatore e la protezione del campione in lavorazione. Utilizzando guanti monouso sterili, camici, mascherine e copricapo si evita la contaminazione del campione, che in medicina legale è fondamentale. Importante è sicuramente anche l'uso di reagenti e consumabili esenti da contaminazioni. Anche le più piccole tracce di DNA contaminanti potrebbero portare a risultati errati. Sono in vendita prodotti denominati “*Forensic DNA Grade*” ovvero reagenti o consumabili *RNA*-free, *DNA*-free, *PCR*-inhibitor-free e *humanDNA*-free.

Un continuo controllo di qualità contribuisce in modo significativo alla certezza dei risultati. È per questo che molti laboratori di genetica forense partecipano a degli esercizi collaborativi che permettono di monitorare tutte le attività. In particolare vengono tenute sotto controllo la sicurezza e la correttezza di tutte le attività di laboratorio.

3- SISTEMI DI RILEVAMENTO DEI MARCATORI GENETICI

Esistono numerose tecniche per il rilevamento dei marcatori genetici. La tecnica attualmente più utilizzata per l'analisi degli STRs è l'elettroforesi capillare (6).

3.1 Elettroforesi Capillare (CE)

L'elettroforesi capillare è una tecnica di separazione elettroforetica che si basa sull'utilizzo di un capillare, riempito da un tampone di corsa, all'interno del quale viene fatto scorrere il campione grazie a una differenza di potenziale applicata alle due estremità del capillare. I vari segmenti di DNA vengono separati, in base alla loro carica e alla loro dimensione, e vengono contemporaneamente rilevati da un laser di rilevazione che cade direttamente sul capillare (*on column*).

Le caratteristiche che rendono la tecnica molto interessante in campo forense sono:

- Versatilità d'applicazione;
- Alto potere risolutivo;

- Possibilità di eseguire separazione basate su principi diversi utilizzando la stessa strumentazione di base;
- Necessità di piccoli volumi di campione per allestire l'analisi, nell'ordine dei nanolitri;
- Ridotto consumo di reagenti;
- Semplicità e basso costo della strumentazione.

3.2 Sequenziamento del DNA: metodo di Sanger

L'altra importante tecnica per la rilevazione dei *markers* genetici è il sequenziamento del DNA.

La tecnica messa a punto da Sanger ha preso piede in molti laboratori di genetica e tuttora viene impiegata. I vantaggi del metodo risiedono nella facilità d'uso e nell'impiego di sonde a fluorescenza (sostitute delle precedenti sonde radioattive) che permettono di rendere il processo automatizzabile. Infatti utilizzando sonde con fluorescenze diverse è possibile analizzare un campione con una sola corsa.

La tecnica di Sanger si basa sull'utilizzo di nucleotidi terminatori di catena, ovvero dideossinucleotidi (ddNTP), questi rispetto ai nucleotidi usati dalle cellule per la sintesi del DNA sono privi del gruppo 3'-OH dello zucchero,

necessario per formare il legame fosfodiesterico con il nucleotide successivo. Quindi, una volta incorporati nella catena nascente di DNA, i ddNTP causano l'interruzione della sintesi.

Per sequenziare un frammento di DNA (filamento stampo), il metodo di Sanger sfrutta la capacità della DNA polimerasi di sintetizzare un nuovo filamento complementare al filamento stampo. Nella miscela di reazione sono presenti sia nucleotidi normali (dNTP), in eccesso, sia terminatori di catena (ddNTP), i quali sono presenti in difetto per evitare la formazione di catene troncate dopo poche basi. Al termine della reazione, si avrà una miscela di filamenti con lunghezza diversa, a seconda di dove è avvenuto l'inserimento del ddNTP terminatore di catena. Grazie all'elettroforesi su gel i diversi frammenti vengono ordinati in base alla loro lunghezza: la disposizione dei frammenti e della fluorescenza svelerà la sequenza dei nucleotidi.

3.3 *Next Generation Sequencing*

Il nuovo millennio ha portato un grande avanzamento tecnologico per quanto riguarda il sequenziamento del DNA: la *Next Generation Sequencing* (NGS), anche nota come *Massive Parallel Sequencing* (MPS).

In tempi brevissimi, sono state sviluppate diverse piattaforme tecnologiche basate su questo nuovo sistema di sequenziamento e via via perfezionate, dando origine ad una vera e propria rivoluzione nel campo del sequenziamento degli acidi nucleici, sia per quanto riguarda la quantità dei dati prodotti sia per la velocità con cui essi vengono generati.

Inoltre, negli ultimi anni si è assistito ad una considerevole diminuzione dei costi di processamento dei campioni, fattore che ha favorito l'ampia diffusione di questi sistemi. Dai seguenti grafici (*Fig.15 e 16*) è palese il drastico calo dei prezzi col passare degli anni.

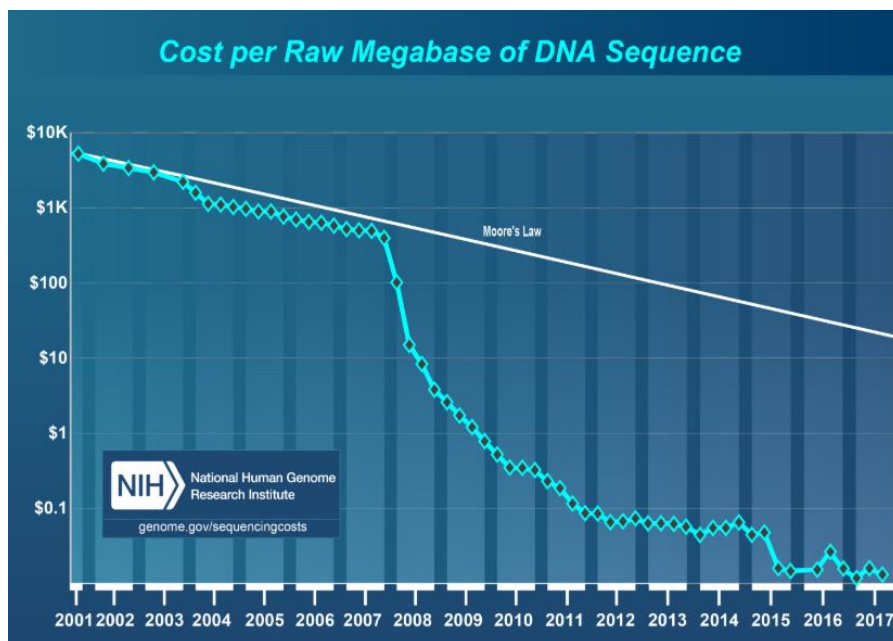


Fig.15 Costo per il sequenziamento di un Mb dal 2001 al 2017 (50).

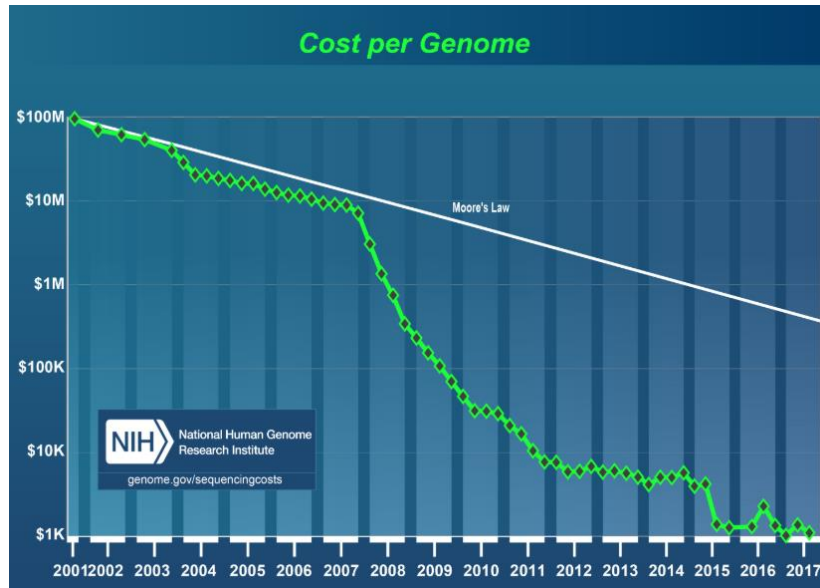


Fig. 16 Costo per il sequenziamento di un intero genoma dal 2001 al 2017 (50).

In Tab.3 vengono mostrati i dati sotto forma di tabella.

DATA	COSTO PER MB	COSTO PER GENOMA
Sep-01	\$5.292,39	\$95.263.072
Sep-02	\$3.413,80	\$61.448.422
Oct-03	\$2.230,98	\$40.157.554
Oct-04	\$1.028,85	\$18.519.312
Oct-05	\$766,73	\$13.801.124
Oct-06	\$581,92	\$10.474.556
Oct-07	\$397,09	\$7.147.571
Oct-08	\$3,81	\$342.502
Oct-09	\$0,78	\$70.333
Oct-10	\$0,32	\$29.092
Oct-11	\$0,09	\$7.743
Oct-12	\$0,07	\$6.618
Oct-13	\$0,06	\$5.096
Oct-14	\$0,06	\$5.731
Oct-15	\$0,014	\$1.245
Oct-16	\$0,015	\$1.356
Jul-17	\$0,012	\$1.121

Tab.3 Decremento annuale del costo necessario per il sequenziamento di un Mb e di un intero genoma umano dal 2001 al 2017 (50).

Attualmente il *gold standard* è rappresentato dal sequenziamento di tipo Sanger, tuttavia l'utilizzo della tecnologia NGS offre diversi vantaggi quali la produzione di molte Gb di sequenze in poco tempo e costi nettamente inferiori.

Con le tecniche NGS è possibile miniaturizzare l'analisi; nella tecnologia a capillari usata per il sequenziamento di Sanger è invece presente un limite fisico, dato dalla struttura stessa dei capillari, che non può essere superato. Inoltre, il passaggio di amplificazione dei frammenti nella tecnologia NGS non prevede l'utilizzo di plasmidi per il clonaggio ma viene eseguito mediante un processo automatizzato, come l'*emulsion PCR* o la *bridge PCR*, più economiche e veloci.

Un'altra caratteristica importante che diversifica le tecnologie *Next Generation* dal Sanger è la capacità di determinare in maniera accurata la presenza di alleli a bassa frequenza. Il dato prodotto da un sequenziatore a capillare deriva dalla rilevazione dell'intensità di fluorescenza emessa globalmente da tutti i frammenti di pari lunghezza. Quindi non è possibile la rilevazione del segnale di un eventuale allele meno rappresentato.

Nelle tecniche NGS ogni frammento di DNA viene sequenziato a sé in un microreattore e l'analisi globale viene fatta prendendo in considerazione tutte

le sequenze ottenute. Quindi è possibile la determinazione di tutte le varianti presenti in un determinato campione.

Un ulteriore vantaggio offerto dalle NGS è la possibilità di processare diversi campioni in una singola corsa attraverso la marcatura delle sequenze (*barcoding*), che, nella fase di analisi informatica, permette di contraddistinguere le sequenze di ciascun campione.

Le tecniche NGS presentano però anche alcune limitazioni.

Una prima barriera è la lunghezza dei frammenti (*reads*) che generalmente si aggira intorno alle 200-300 bp, contro le 1000 bp di un sequenziatore a capillari. Questo aspetto complica notevolmente l'assemblaggio delle sequenze contigue.

Ulteriore fattore limitante è l'aumentata percentuale di errore (falsi positivi e falsi negativi), se paragonata al sistema Sanger. Tale limite, in gran parte, viene superato attraverso l'aumento del livello di *coverage* (quantità di sequenze analizzate relative ad un frammento di DNA) e attraverso l'utilizzo di polimerasi ad altissima efficienza.

Inoltre è importante considerare l'aspetto bioinformatico. Infatti è necessario lo sviluppo di algoritmi bioinformatici che consentano di analizzare con efficacia i dati prodotti. Inoltre, l'interpretazione dei dati e la gestione di un

sistema informatico così complesso prevede competenze specialistiche e richiede la presenza di figure professionali specifiche.

3.3.1 Tecnologia *Ion Torrent*: la piattaforma *Ion PGM™ System*

Esistono diversi tipi di sequenziatori NGS, spesso identificati con il nome della compagnia che li ha sviluppati. La piattaforma utilizzata in questo studio è l'*Ion PGM™ System (Personal Genome Machine)*.

Il protocollo prevede l'amplificazione clonale del campione attraverso *emulsion PCR* ma il processo di sequenziamento presenta due aspetti essenziali che lo caratterizzano rispetto alle altre tecnologie:

- Utilizzo dei semiconduttori come elemento strutturale dei supporti in cui viene dispensato il campione per il sequenziamento (*Ion Chips*);
- Impiego di un sistema di rilevamento non basato su reazioni luminescenti ma su variazioni di potenziale.

In particolare *Ion PGM™ System* sfrutta la tecnologia dei semiconduttori per rilevare processi biochimici già noti. In tutte le cellule, quando un singolo nucleotide è incorporato dall'enzima DNA polimerasi nella catena nascente di DNA a doppia elica, viene rilasciato uno ione idrogeno (H^+) come sottoprodotto. Infatti la formazione di ogni nuovo legame fosfodiesterico

determina l'idrolisi del gruppo trifosfato del nucleotide entrante con la liberazione di un protone (*Fig.17*).

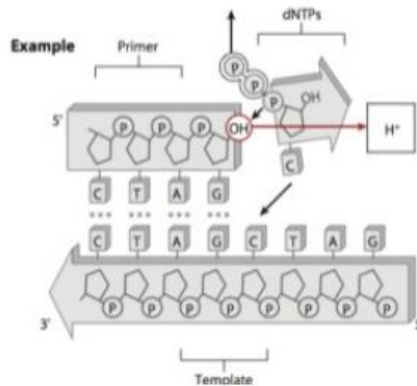


Fig.17 Reazione di incorporazione dNTP, con rilascio di PP_i e H^+ .

La tecnologia *Ion Torrent* è incentrata sull'impiego di un particolare chip (*Ion Chip*) con micropozzetti all'interno dei quali avviene ciclicamente il processo di incorporazione nucleotidica (e conseguente rilascio di ioni H^+).

L'*Ion Chip* è costituito da uno strato superiore, in cui sono ricavati dei pozzetti di grandezza tale da poter accogliere solo una biglia e i reagenti per il sequenziamento. Questo primo strato superficiale poggia su un secondo costituito da semiconduttori, che permette la trasmissione dei segnali allo strato ancora sottostante. Quest'ultimo è strutturato come una piastra di sensori, uno corrispondente a ciascun pozzetto, che hanno la capacità di registrare le piccolissime variazioni di pH che avvengono all'interno del pozzetto durante il sequenziamento, trasformandole in differenze di potenziale (*Fig.18*).

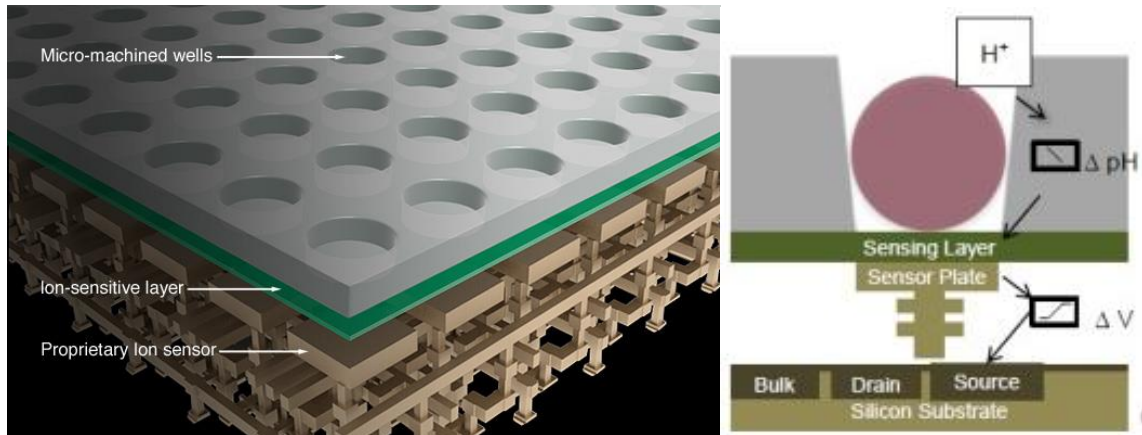


Fig.18 Schema della struttura a tre strati dell' Ion Chip (51).

Il sistema consente l'aggiunta di un nucleotide noto alla volta: se il nucleotide aggiunto al sistema non viene incorporato nella molecola di DNA neosintetizzata, in quanto non complementare al filamento stampo adesivo alle *beads*, non verrà prodotta nessuna modifica del pH e di conseguenza nessuna modifica del voltaggio. Un ciclo del processo dura pochi secondi poiché non richiede una rilevazione ottica, come accade invece negli altri sistemi di *Next Generation Sequencing* (52).

È importante sottolineare come la variazione di pH registrata dal sensore sia direttamente proporzionale al numero delle basi incorporate nella sequenza nascente. Una serie di adenine consecutive risulterà visibile graficamente come un picco più alto. L'intensità della variazione di potenziale registrata è tanto superiore quante più sono le basi identiche consecutive (Fig.19).

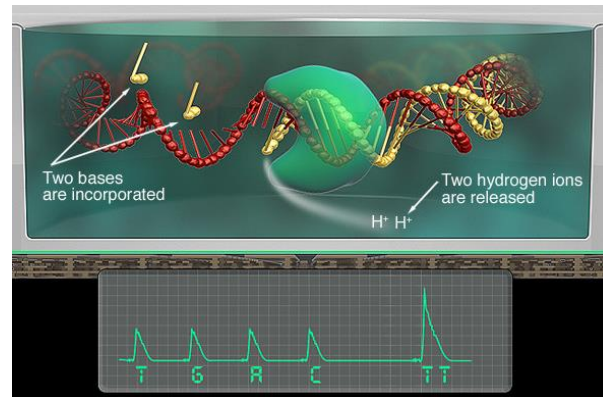


Fig.19 Reazione di incorporazione di dNTP e relativo segnale prodotto.

Il processo totale di generazione e misura del segnale (da quando viene riversata la soluzione contenente i nucleotidi sulla superficie del chip fino alla avvenuta misura del pH) dura poco più di 4 s. Successivamente vi è una fase di lavaggio necessaria per eliminare dai pozzetti i reagenti già utilizzati: così il chip viene preparato per un successivo flusso. Questo processo viene ripetuto ciclicamente per 500 volte consentendo la lettura di *reads* lunghe circa 200 bp.

A differenza delle tecnologie che sfruttano i segnali di fluorescenza, il sistema *Ion Torrent* non prevede le reazioni chimiche necessarie per la eliminazione dei fluorofori legati ai nucleotidi e ciò favorisce la rapidità del processo che caratterizza questa tecnologia.

Un altro aspetto vantaggioso offerto dalla tecnologia *Ion Torrent* è la disponibilità di chip con diverse capacità che consente di utilizzare chip specifici per ogni esigenza. Nel più piccolo dei chip attualmente disponibili

per *PGM™ (Ion 314™ Chip v2)* si contano circa $1,2 \times 10^6$ pozzetti, ciascuno del diametro di $3.5 \mu\text{m}$, per una totale capacità di sequenziamento che oscilla tra le 30 e le 100 Mb, mentre i chip a più alta processività ($\sim 12 \times 10^6$ pozzetti) consentono il sequenziamento dalle 600 Mb alle 2 Gb (*Ion 318™ Chip v2*).

Prima della messa in funzione del *PGM™* è necessario effettuare la preparazione del campione. In particolare si effettua una fase preliminare di *emulsion PCR* delle librerie tramite l'*Ion OneTouch 2™ Instrument* (Fig.20).



Fig.20 *Ion One Touch 2™* (destra) e *Ion OneTouch™ ES* (sinistra).

In pratica si allestisce una amplificazione in emulsione dei frammenti di DNA barcodati. Successivamente, con il sistema *Ion OneTouch™ ES* si procede con l'arricchimento/purificazione del campione amplificato.

L'*Ion One Touch 2™* è in grado di creare un'emulsione di acqua e olio che favorisce la formazione di micelle in cui possono essere facilmente incorporati una biglia (*bead*) e un singolo amplicone. Nella micella avviene l'amplificazione della molecola di DNA incorporata (*Fig.21*).

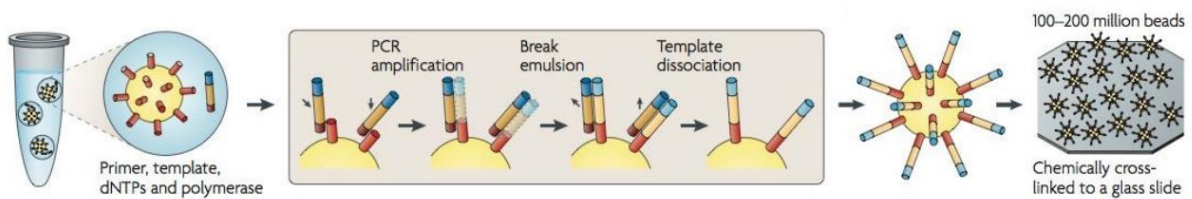


Fig.21 Emulsion PCR (51).

Dopo l'amplificazione clonale, ciascuna particella sferica (ISP) contiene tante copie dello stesso frammento di DNA, che da un lato è legato alla ISP, dall'altro invece è biotinilato.

L'arricchimento con il sistema *Ion OneTouch™ ES*, sfrutta l'affinità tra biotina e streptavidina. In particolare viene poi utilizzato un secondo tipo di biglie, a cui è adesa la streptavidina, in grado di selezionare solo le *beads* positive (monoclonali o policlonali). In questo modo, tutte le ISPs vuote, che non contengono alcun frammento di DNA, vengono eliminate (*Fig.22*).

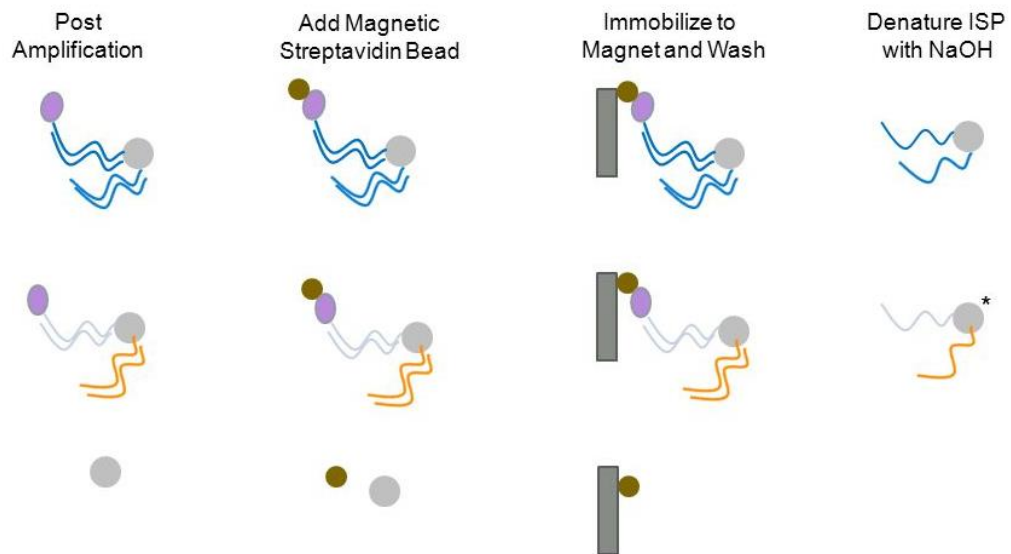


Fig. 22 Illustrazione grafica delle tre possibili situazioni. Beads monoclonale, beads policlonale, beads che non ha legato nessun frammento di DNA.

Il campione finale è composto da ISPs a cui sono adesi i frammenti di DNA. Questo viene poi caricato su un chip (*Ion Chip*) per la fase di sequenziamento con *PGM™*.

Capitolo 4

SCOPO DELLO STUDIO

L'obiettivo della presente ricerca è quello di valutare la performance di un pannello di 29 microaplotipi su DNA degradato tramite l'utilizzo della piattaforma *Ion PGM™ System*.

Sono stati applicati dei criteri per la scelta dei 29 microaplotipi a partire da un pannello di 89 MH già testati su *Ion PGM™ System* (24). La scelta è ricaduta su microaplotipi molto informativi con un alto valore A_e , indice della variazione di un determinato *locus* genico all'interno di una popolazione.

Lo studio si è concentrato esclusivamente sui microaplotipi proprio per evidenziare i vantaggi di questi nuovi *markers* rispetto agli STRs. Uno tra i principali vantaggi è proprio la ristretta ampiezza della regione genica amplificabile. Questo comporta la possibilità di utilizzare tali *markers* anche su campioni con DNA moderatamente o fortemente degradato, cosa impossibile se si parla di STRs. Pertanto il pannello è stato testato su campioni di DNA diversi.

La valutazione ha previsto delle prove con DNA non degradato per verificare la performance del pannello di *primers* disegnato e il suo corretto

funzionamento. Inoltre sono state eseguite prove di sensibilità, utilizzando sei diluizioni seriali di un campione di DNA di controllo. Lo scopo prefissato era quello di evidenziare fino a che punto si spingevano l'affidabilità e la riproducibilità del pannello valutato. Dopodiché sono stati analizzati dei campioni di DNA degradato artificialmente per verificare l'effettiva applicabilità del pannello su campioni di DNA non ottimali. Infine è stata eseguita la stessa procedura di preparazione delle librerie e di sequenziamento anche a campioni forensi reali.

La valutazione dei risultati ottenuti tramite sequenziamento con tecnologia MPS è stata effettuata tramite un *tool* bioinformatico che ha permesso di confrontare gli esiti ottenuti con il genoma hg19 di riferimento. È stata scelta proprio la *Massive Parallel Sequencing* poiché risulta essere una tecnica molto avanzata che produce una grande quantità di dati.

Capitolo 3

MATERIALI E METODI

3.1 Selezione dei loci d'interesse

I microaplotipi sono stati selezionati a partire dagli 89 MH precedentemente tipizzati e studiati in un campione popolazionistico italiano (24). Tutti i *loci* studiati sono presenti in ALFRED, *ALlele FREquency Database* (19).

La scelta è stata fatta in base ai seguenti criteri:

- Microaplotipi composti da 2 SNPs a 5 SNPs;
- Microaplotipi con valori di $A_e \geq 1,8$;
- Microaplotipi amplificabili con lunghezze inferiori alle 180 bp.

3.2 Disegno dei *primers*

I *primers* da utilizzare per la prima fase di amplificazione sono stati progettati *online* tramite *Ion AmpliSeq Designer* sul sito della *ThermoFisher® Scientific* (56). Nella sezione *Ion AmpliSeq Designer* è possibile creare un pannello su misura immettendo tutti i parametri richiesti, come il nome del pannello, il

tipo di acido nucleico, il genoma di riferimento e il file CSV con i 72 microplotipi selezionati.

Si è posta particolare attenzione sulla scelta della grandezza dell'amplicone, poiché questo studio punta all'applicazione del pannello su campioni con DNA fortemente degradato.

La scelta cade sull'opzione "FFPE" ("*Formalin-Fixed Paraffin-Embedded*") in cui la lunghezza massima dei frammenti è di 175 bp.

Di seguito sono riportati i pannelli proposti dalla *ThermoFisher*[®] seguendo i criteri scelti (*Fig.23*).

Available M-F 9 AM - 5 PM EST
LIVE CHAT
Contact

8 My Designs

Home

Chip Calculator

Notifications

Help

Switch design: 72_Microhaps Copy Targets

Start a new design

72_Microhaps

Amplicon distribution Hide solutions -

IAD176720 - Results ready

Solution ID	Solution Type	DNA Type	Instrument & Application	Amplicon Range	Pools (Input DNA)	Amplicons	Misused (bp)	Coverage (%)
IAD176720_152	High Specificity	cdDNA	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530,540,550*) Ion S5, S5 XL (510,520,530,540,550*) PGM Proton 	125 - 140 bp	2 (1-20 ng)	138	1,289	85.55
IAD176720_167	High Specificity	FFPE	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (530,540,550*) Ion S5, S5 XL (530,540,550*) 	125 - 175 bp	2 (1-20 ng)	125	481	94.61
IAD176720_182	High Specificity	Standard DNA (275 bp)	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530,540,550*) PGM Proton 	125 - 275 bp	2 (1-20 ng)	81	0	100
IAD176720_196	High Specificity	Standard DNA (375 bp)	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530) PGM 	125 - 375 bp	1 (1-10 ng)	72	0	100

*550 chip is supported only on Ion S5 XL, Ion GeneStudio S5 Plus and Ion GeneStudio S5 Prime.

Recommended for:

Ion GeneStudio S5, S5 Plus, S5 Prime (530,540,550*)
Ion S5, S5 XL (530,540,550*)

94.61% Coverage

See how to [increase coverage](#) up to 100%

Chip Calculator

2 (1-20 ng)
Pools (Input DNA)

Pool1: 66 amplicons | Pool2: 59 amplicons

125 - 175 bp

Amplicon Range

12.79 kb

Panel Size

Fig.23 Pannelli proposti dalla ThermoFisher® con i 72 microaplotipi.

Come si vede dalla *Fig.23* il pannello scelto non ha il 100% di *coverage*, ma solo il 94,61% (evidenziato in rosso). Andando a vedere meglio gli ampliconi proposti (*Fig.24*) si evince che già tra i primi 10 ampliconi, ce n'è uno (evidenziato in rosso) che perde delle basi (30) e quindi ha un *coverage* più basso (84,46%).

Pipeline version 7.06 Genome reference Human (hg19) Exon padding 5 bp

Type	Name	Chrom	Start	End	Target (bp)	Missed (bp)	Covered (%)	
Region	mh01KK-001	chr1	3743132	3743319	187	0	100	UCSC ¹²
Region	mh01KK-106	chr1	4227464	4227634	170	0	100	UCSC ¹²
Region	mh01KK-205	chr1	18722692	18722846	154	0	100	UCSC ¹²
Region	mh01KK-211	chr1	161955656	161955804	148	0	100	UCSC ¹²
Region	mh01KK-117	chr1	204633340	204633526	186	0	100	UCSC ¹²
Region	mh10KK-163	chr10	3162410	3162659	259	0	100	UCSC ¹²
Region	mh10KK-169	chr10	14208510	14208652	142	0	100	UCSC ¹²
Region	mh11KK-180	chr11	1690791	1690984	193	30	84,46	UCSC ¹²
Region	mh11KK-036	chr11	5109946	5110068	122	0	100	UCSC ¹²
Region	mh11KK-037	chr11	72269214	72269347	133	0	100	UCSC ¹²
Region	mh11KK-187	chr11	78947596	78947777	181	0	100	UCSC ¹²
Region	mh11KK-191	chr11	99880163	99880352	189	0	100	UCSC ¹²

Fig.24 Primi 10 ampliconi proposti con il pannello a 72 microaplotipi.

Inoltre molti di questi microaplotipi vengono divisi in due ampliconi separati. Questi microaplotipi devono essere necessariamente eliminati poiché amplificare due regioni separatamente non permette di risolvere le fasi *cis/trans* tra gli alleli che compongono un microaplotipo.

Viste le incongruenze rilevate per alcuni microaplotipi, si è giunti alla scelta di un pannello che presenta 29 MH. Questo pannello è stato scelto e approvato poiché possiede tutte le caratteristiche richieste (Fig.25).

iontorrent
by Thermo Fisher Scientific

29_Microhaps

Switch design: 29_Microhaps

IAD176925 - Results ready

Solution ID	Solution Type	DNA Type	Amplicon Range	Instrument & Application	Pools (Input DNA)	Amplicons	Missed (bp)	Coverage (%)	
IAD176925_152	High Specificity	ctDNA	125 - 140 bp	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530,540,550*) Ion S5, S5 XL (510,520,530,540,550*) PGM Proton 	germline and somatic germline and somatic germline and somatic	2 (1-20 ng)	41	175	91,38
IAD176925_166	High Specificity	FFPE	125 - 175 bp	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530,540,550*) Ion S5, S5 XL (510,520,530,540,550*) PGM Proton 	germline and somatic germline and somatic germline and somatic	1 (1-10 ng)	29	0	100
IAD176925_181	High Specificity	Standard DNA (275 bp)	125 - 275 bp	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530,540,550*) Ion S5, S5 XL (510,520,530,540,550*) PGM Proton 	germline germline germline	1 (1-10 ng)	29	0	100
IAD176925_196	High Specificity	Standard DNA (375 bp)	125 - 375 bp	<ul style="list-style-type: none"> Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530) Ion S5, S5 XL (510,520,530) PGM 	germline germline germline	1 (1-10 ng)	29	0	100

*550 chip is supported only on Ion S5 XL, Ion GeneStudio S5 Plus and Ion GeneStudio S5 Prime.

Recommended for:

 Ion GeneStudio S5, S5 Plus, S5 Prime (510,520,530,540,550*)
 Ion S5, S5 XL (510,520,530,540,550*)

Also supported with:

 PGM Proton

100.00% Coverage *i*
 See how to [increase coverage](#) up to 100%

Chip Calculator *i*

1 (1-10 ng)
 Pools (Input DNA) *i*
 Pool1: 29 amplicons

125 - 175 bp
 Amplicon Range

3.41 kb
 Panel Size *i*

Fig.25 Pannelli proposti dalla ThermoFisher® con i 29 microaplotipi.

Il pannello IAD176925_166 prevede:

- Range di lunghezza degli ampliconi tra 125-175 bp;
- Un solo pool di primer;
- Una quantità di DNA richiesta di 1-10 ng;
- 29 ampliconi;
- 0 bp perse;
- Coverage del 100%.

Di seguito vengono riportati i dati relativi ad ogni amplicone (Tab.4).

Amplicon_ID	MH	GENOMA	Chr	Amplicon _start	Insert _Start	Insert _Stop	Amplicon _Stop
AMPL7163692463	mh02KK-073	hg19	chr2	46586978	46587003	46587129	46587152
AMPL7159737231	mh02KK-134	hg19	chr2	161079339	161079362	161079486	161079509
AMPL7163692466	mh02KK-136	hg19	chr2	228092338	228092364	228092485	228092511
AMPL7163692467	mh02KK-213	hg19	chr2	101096835	101096858	101096974	101096997
AMPL7157278743	mh03KK-007	hg19	chr3	46352280	46352310	46352428	46352454
AMPL7163692444	mh03KK-009	hg19	chr3	113890728	113890748	113890869	113890892
AMPL7163692470	mh04KK-010	hg19	chr4	1986895	1986918	1986993	1987021
AMPL7163692471	mh04KK-011	hg19	chr4	37858838	37858866	37858979	37859009
AMPL7161217348	mh04KK-013	hg19	chr4	68444062	68444089	68444206	68444235
AMPL7163692474	mh04KK-074	hg19	chr4	56323561	56323584	56323708	56323731
AMPL7157278785	mh05KK-022	hg19	chr5	9619823	9619848	9619972	9619996
AMPL7157301183	mh05KK-062	hg19	chr5	6844941	6844965	6845085	6845115
AMPL7165692997	mh05kk-170	hg19	chr5	2448025	2448049	2448165	2448187
AMPL7157278802	mh06KK-026	hg19	chr6	16399591	16399619	16399739	16399764
AMPL7157278786	mh07KK-031	hg19	chr7	96733931	96733956	96734080	96734104
AMPL7157278787	mh09KK-034	hg19	chr9	12672180	12672209	12672322	12672354
AMPL7163692481	mh09KK-153	hg19	chr9	103969718	103969739	103969854	103969881
AMPL7157278796	mh11KK-039	hg19	chr11	113211274	113211296	113211404	113211434
AMPL7161217297	mh12KK-043	hg19	chr12	3708464	3708489	3708614	3708638
AMPL7157278762	mh12KK-046	hg19	chr12	118889453	118889478	118889593	118889623
AMPL7161217304	mh13KK-218	hg19	chr13	54060755	54060778	54060892	54060912
AMPL7163692506	mh13KK-225	hg19	chr13	66712704	66712731	66712830	66712855
AMPL7161217316	mh16KK-302	hg19	chr16	7587710	7587733	7587854	7587884
AMPL7157278801	mh17KK-053	hg19	chr17	45695784	45695807	45695933	45695958
AMPL7157278768	mh17KK-054	hg19	chr17	47287036	47287059	47287157	47287187
AMPL7161217321	mh18KK-293	hg19	chr18	76089846	76089868	76089986	76090012
AMPL7156695641	mh21KK-320	hg19	chr21	43062901	43062925	43063046	43063074
AMPL7161217337	mh21KK-324	hg19	chr21	46714605	46714624	46714760	46714779
AMPL7154029350	mh22KK-060	hg19	chr22	19951146	19951166	19951279	19951298

Tab.4 Ampliconi proposti dal pannello ThermoFisher® e relative caratteristiche.

3.3 Campioni utilizzati per la validazione del pannello

Inizialmente sono stati selezionati 6 campioni di DNA, estratti da sangue fresco (*Tab.5*), già testati nel pannello originale di 89 MH (24), per confermare la funzionalità del *pool* di *primer* disegnato e la concordanza dei risultati finali ottenuti con il presente pannello.

CODICE	TIPO DI CAMPIONE	CONCENTRAZIONE
29MH_ASAN056	SANGUE FRESCO	20 ng/μL
29MH_ASAN057	SANGUE FRESCO	20 ng/μL
29MH_ASAN090	SANGUE FRESCO	20 ng/μL
29MH_ASAN091	SANGUE FRESCO	20 ng/μL
29MH_ASAN102	SANGUE FRESCO	20 ng/μL
29MH_ASAN115	SANGUE FRESCO	20 ng/μL

Tab.5 Informazioni sui campioni di DNA non degradato di controllo.

Per il test di sensibilità sono state utilizzate 6 diluizioni seriali, riportate in *Tab.6*, del DNA di controllo (2800M) a concentrazione nota (10 ng/μL).

CODICE	CONCENTRAZIONE	QUANTITA' DI DNA	QUANTITA' DI H2O	VOLUME TOTALE	DILUIZIONE
29MH_D1	5 ng/μL	10 μL	10 μL	20 μL	1:2
29MH_D2	1 ng/μL	5 μL	20 μL	25 μL	1:5
29MH_D3	500 pg/μL	10 μL	10 μL	20 μL	1:2
29MH_D4	100 pg/μL	5 μL	20 μL	25 μL	1:5
29MH_D5	50 pg/μL	10 μL	10 μL	20 μL	1:2
29MH_D6	25 pg/μL	10 μL	10 μL	20 μL	1:2

Tab.6 Tabella diluizioni campione di controllo 2800M.

Sono state effettuate due prove di amplificazione della regione target, a 21 e a 25 cicli, utilizzando in entrambi i casi 1 μL delle diluizioni in tabella.

Per i test su DNA degradato sono stati utilizzati campioni di DNA degradato artificialmente a concentrazione nota.

Nella tabella successiva (*Tab.7*) sono riportate le informazioni relative ai tre campioni degradati tramite depurinazione.

CODICE	TIPO DI CAMPIONE	CONCENTRAZIONE	INDICE DI DEGRADAZIONE
29MH_TS22	DEPURINATO	loq	nc
29MH_TS26	SANGUE FRESCO (campione di riferimento)	0,574 ng/ μL	0,8
29MH_TS29	DEPURINATO	0,462 ng/ μL	nc

Tab.7 Informazioni sui campioni di DNA degradato artificialmente.

Anche con i campioni di DNA degradato sono state effettuate delle prove a 21 e a 25 cicli di amplificazione della regione *target*. L'allestimento della reazione di amplificazione, in questo caso, ha comportato l'utilizzo di quantità diverse del campione in base alla concentrazione iniziale.

Sono stati infine testati dei campioni forensi reali (casework samples), contenenti DNA degradato. La quantità di DNA contenuta in ogni campione è descritta nella tabella successiva (*Tab.8*).

CODICE	TIPO DI CAMPIONE	CONCENTRAZIONE	INDICE DI DEGRADAZIONE
29MH_BO04	TESSUTO INCLUSO IN PARAFFINA	8,214 ng/μL	40
29MH_BO05	TESSUTO INCLUSO IN PARAFFINA	3,592 ng/μL	nc
29MH_TS04	TESSUTO OSSEO	loq	nc
29MH_TS07	TESSUTO OSSEO	0,689 ng/μL	5

Tab.8 Informazioni sui campioni forensi reali.

In questo caso è stato utilizzato un solo protocollo di amplificazione a 25 cicli. Anche in questo caso sono state utilizzate quantità diverse del campione in base alla loro concentrazione iniziale.

3.3 Estrazione e quantificazione del DNA

Per questo studio sono stati utilizzati campioni di DNA precedentemente estratti e quantizzati con metodiche diverse da altri laboratori (*Tab.9*).

CAMPIONE	METODO D'ESTRAZIONE	METODO DI QUANTIZZAZIONE
29MH_ASAN056	QIAamp® DNA Micro	E-Gel™ Agarose Gels with SYBR Safe™
29MH_ASAN057	QIAamp® DNA Micro	E-Gel™ Agarose Gels with SYBR Safe™
29MH_ASAN090	QIAamp® DNA Micro	E-Gel™ Agarose Gels with SYBR Safe™
29MH_ASAN091	QIAamp® DNA Micro	E-Gel™ Agarose Gels with SYBR Safe™
29MH_ASAN102	QIAamp® DNA Micro	E-Gel™ Agarose Gels with SYBR Safe™
29MH_ASAN115	QIAamp® DNA Micro	E-Gel™ Agarose Gels with SYBR Safe™
29MH_TS22	Fenolo/Cloroformio	Quantifiler® Trio DNA Quantification
29MH_TS26	Fenolo/Cloroformio	Quantifiler® Trio DNA Quantification
29MH_TS29	Fenolo/Cloroformio	Quantifiler® Trio DNA Quantification
29MH_BO04	QIAamp® DNA Investigator	Quantifiler® Trio DNA Quantification
29MH_BO05	QIAamp® DNA Investigator	Quantifiler® Trio DNA Quantification
29MH_TS04	Fenolo/Cloroformio	Quantifiler® Trio DNA Quantification
29MH_TS07	Fenolo/Cloroformio	Quantifiler® Trio DNA Quantification

Tab.9 Tecniche utilizzate per l'estrazione e la quantizzazione del DNA.

3.4.1 Estrazione DNA con *QIAamp*® DNA Micro

L'estrazione del DNA genomico tramite *QIAamp*® DNA Micro della *QIAGEN*® (*QIAamp*® DNA Micro Handbook, Third Edition, December 2014, *QIAGEN*®) (54), prevede una fase iniziale di lisi cellulare, necessaria per il rilascio dell'acido nucleico, che avviene in condizioni denaturanti ad alta temperatura; il passaggio del lisato attraverso una colonnina e il legame del DNA alla matrice silicea della colonnina. Dopodiché si effettuano diversi lavaggi e infine si eluisce il DNA con un buffer o semplice acqua distillata.

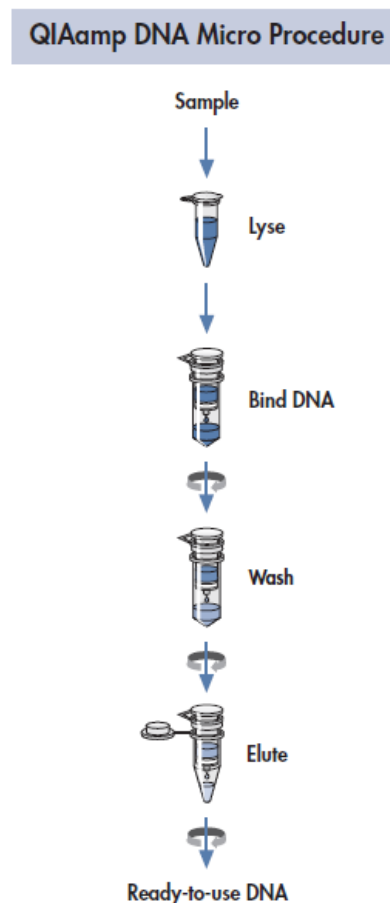


Fig.27 Protocollo d'estrazione del DNA con *QIAamp*® DNA Micro (54).

La procedura prevede l'utilizzo di sangue e anticoagulante (EDTA, citrato o eparina) come materiale di partenza.

- Aggiungere 100 μ L di *Buffer ATL* a 2 gocce di sangue;
- Aggiungere 10 μ L di *Proteinasi K*;
- Aggiungere 100 μ L di *Buffer AL*;
- Vortexare per 15 secondi;
- Incubazione a 56°C per 10 minuti in agitazione su *Eppendorf Thermomixer*;
- Breve centrifugazione;
- Aggiungere 50 μ L di etanolo (96-100%);
- Vortexare per 15 secondi;
- Incubazione a temperatura ambiente per 3 minuti;
- Breve centrifugazione;
- Preparare una *QIAamp MinElute column* per ogni campione (ogni colonnina è già dotata di una *collection tube* per la raccolta del liquido in eccesso);
- Trasferire l'intero lisato (dopo la breve centrifugazione) all'interno della colonnina, senza toccare la membrana della colonnina;
- Centrifugazione a 8000 rpm per 1 minuto;

- Trasferire la colonnina all'interno di una nuova *collection tube* e gettare quella che contiene il liquido di scarto.
- Aggiungere 500 μL di *Buffer AW1*, senza toccare la membrana della colonnina;
- Centrifugazione a 8000 rpm per 1 minuto;
- Trasferire la colonnina all'interno di una nuova *collection tube* e gettare quella che contiene il liquido di scarto.
- Aggiungere 500 μL di *Buffer AW2*, senza toccare la membrana della colonnina;
- Centrifugazione a 8000 rpm per 1 minuto;
- Trasferire la colonnina all'interno di una nuova *collection tube* e gettare quella che contiene il liquido di scarto.
- Centrifugazione a 14000 rpm per 3 minuti, per asciugare la membrana;
- Trasferire la colonnina in una provetta da 1,5 mL (in cui viene eluito il DNA);
- Aggiungere 50 μL di *Buffer AE*, al centro della membrana della colonnina;
- Incubare a temperatura ambiente per 1 minuto;
- Centrifugare a 14000 rpm per 1 min;
- Gettare la colonnina, il DNA eluito si trova nella provetta da 1,5 mL.

3.4.2 Estrazione con fenolo/cloroformio

Il metodo classico di estrazione del DNA prevede l'uso di solventi organici come fenolo e cloroformio e consta di diverse fasi.

La prima è quella della lisi cellulare necessaria per liberare il DNA e l'RNA dall'involucro nucleare. La reazione avviene utilizzando detergenti (SDS, EDTA...), che permettono la lisi delle pareti cellulari, ed enzimi (proteasi K), che permettono l'allontanamento delle proteine istoniche.

Dopodiché si procede con la fase di separazione e recupero dell'acido nucleico. In pratica al lisato cellulare si aggiunge una soluzione di fenolo e cloroformio, si vortexa e si centrifuga fino ad ottenere tre fasi:

- superiore, che contiene la soluzione di acidi nucleici;
- interfase, di proteine denaturate;
- inferiore, fase fenolica contenente lipidi e proteine ricche di amminoacidi idrofobici.

Viene sfruttata l'affinità degli acidi nucleici con soluzioni le acquose.

Dopodiché si raccoglie la superiore (fase acquosa), alla quale viene aggiunto etanolo, necessario per far precipitare gli acidi nucleici. Sequestrando l'acqua, l'etanolo permette la formazione di un pellet (DNA o RNA). Il surnatante (etanolo) si elimina, mentre gli acidi nucleici si risospendono in acqua.

3.4.3 Estrazione con *QIAamp*® *DNA Investigator*

Il protocollo di estrazione *QIAamp*® *DNA Investigator* (55), segue i passaggi principali dell'estrazione *QIAamp*® *DNA Micro*. Le uniche differenze sono riscontrabili negli ultimi passaggi.

Dopo il secondo lavaggio con il *Buffer AW2*, viene effettuato un ulteriore lavaggio della colonnina con 700 µL di etanolo, eliminando lo scarto dopo la centrifugazione. La fase finale di eluizione del DNA viene effettuata come nel protocollo *QIAamp*® *DNA Micro* utilizzando 20-100 µL di *Buffer ATE* o acqua distillata.

3.4.4 Protocollo di depurinazione

Alcuni campioni presi in esame sono stati degradati artificialmente tramite depurinazione. Per depurinazione si intende l'idrolisi del legame glicosidico tra lo zucchero e la base azotata e sostituzione con una molecola d'acqua.

La metodica utilizzata segue studi precedenti basati su depurinazione del DNA in condizioni di blanda acidità (56).

La depurinazione è stata effettuata aggiungendo al campione di DNA il tampone Na-citrato a pH 5. La reazione è avvenuta in soluzione acquosa a 70°C per 5 ore. In condizioni acquose il processo idrolitico avviene molto più

lentamente ed è possibile monitorare meglio il grado di degradazione del campione.

Una volta finito il tempo d'incubazione, ogni campione è stato subito centrifugato in colonne filtranti (*Ultracel 3K Amicon Ultra Columns, Millipore USA*) per 30 minuti a 12.000 rpm. In questo modo vi è l'eliminazione delle molecole con peso molecolare inferiore a 3.000 Daltons. Le molecole di DNA ritenute su filtro vengono successivamente lavate ed eluite con acqua tramite centrifugazione a 12.000 rpm per 20 minuti.

3.4.5 Quantizzazione DNA con *Quantifiler® Trio DNA Quantification Kit*

La quantizzazione fatta con *Quantifiler® Trio DNA Quantification Kit* (57) della *Applied Biosystem®* è una metodica di quantizzazione del DNA tramite *Real-Time PCR*. In particolare vengono amplificate regioni *target* umane costituite da più copie disperse sui cromosomi (*Small Autosomal Target, Large Autosomal Target* e *Y-chromosome*) e lo scopo è quello di ottenere una migliore sensibilità di rilevamento e di determinare l'indice di degradazione di un campione di DNA (57).

Il controllo interno di PCR (IPC) è composta da:

- Templato IPC (una sequenza sintetica di DNA non trovata in natura);
- *Primer* per l'amplificazione del DNA del template IPC (130 bp);
- Sonde per il rilevamento *TaqMan® probe dye-quencher—JUN® dye*.

L'IPC viene inserito in ogni campione e, dato che contiene una sequenza di DNA sintetico, fornisce una conferma che la reazione di PCR è avvenuta correttamente. Questo controllo interno è molto utile per confermare la validità dei risultati negativi. È utile anche per identificare campioni che contengono inibitori della PCR (*Tab.10*).

TARGET	LUNGHEZZA AMPICONE
Human Target, Small Autosomal (SA)	80 bp
Human Target, Large Autosomal (LA)	214 bp
Human Male Target (Y)	75 bp
Internal PCR Control (IPC)	130 bp

Tab.10 Target del Quantifiler® Trio DNA Quantification Kit.

Per effettuare una corretta quantizzazione, questa tecnica utilizza 5 diluizioni di uno standard a concentrazione nota (100 ng/μL). Dopo aver preparato queste 5 diluizioni, utilizzando il *Quantifiler® THP DNA Dilution Buffer* come diluente, si procede con la preparazione della mix di PCR (*Tab.11*).

COMPONENTE	VOLUME
Quantifiler® Trio Primer Mix	8 μL
Quantifiler® THP PCR Reaction Mix	10 μL
Campione (o Standard)	2 μL

Tab.11 Mix di amplificazione Real-Time.

Le provette vengono caricate su *Rotor-Gene 6000 PCR System*® con le seguenti condizioni (Tab.12).

PASSAGGIO	TEMPERATURA	TEMPO	CICLI
Attivazione dell'enzima	95°C	2 minuti	
Denaturazione	95°C	9 secondi	per 40 cicli
Appaiamento ed estensione	60°C	30 secondi	

Tab.12 Condizioni di amplificazione.

I risultati ottenuti da questo tipo di quantizzazione ci permettono di determinare il grado di degradazione del campione grazie al “*Degradation Index*”. Esso viene calcolato automaticamente dal software tramite la seguente formula:

$$\frac{\text{Small autosomal target DNA conc. (ng/}\mu\text{L)}}{\text{Large autosomal target DNA conc. (ng/}\mu\text{L)}}$$

L'indice di degradazione (DI) può essere influenzato da:

- Grado di degradazione della *Large Autosomal Target*;
- Presenza di inibitori della PCR.

Se ci sono inibitori della PCR, essi influenzano negativamente il *Large Autosomal Target* (LA) rispetto alla *Small Autosomal Target* (SA), causando un'amplificazione meno efficiente e valori di C_T più alti per LA.

Per determinare l'indice di qualità (QI), si deve valutare l'indice di degradazione (DI) in combinazione con il C_T dell'IPC (Tab.13).

RILEVAZIONE CT DELL'IPC	DEGRADATION INDEX	INTERPRETAZIONE DEL QUALITY INDEX
NO	<1	DNA non degradato o inibito
	Tra 1 e 10	DNA poco degradato (possibile presenza di inibitori di PCR, che non hanno inibito completamente l'amplificazione dell'IPC)
	>10 o nc	DNA molto degradato (possibile presenza di inibitori di PCR, che non hanno inibito completamente l'amplificazione dell'IPC)
SI	<1	Risultato possibile, ma molto improbabile (se gli inibitori della PCR inibiscono l'IPC, dovrebbero inibire anche LA)
	>1 o nc	DNA degradato o con inibitori della PCR

Tab.13 Determinazione del Quality Index (QI) (57).

3.4.6 Quantizzazione con *E-Gel™ Agarose Gels with SYBR Safe™*

La quantizzazione con *E-Gel™ Agarose Gels with SYBR Safe™* (58) prevede l'allestimento di un'elettroforesi su gel d'agarosio al 2%, per la determinazione della taglia e per la quantizzazione delle regioni amplificate.

Per ottenere buoni risultati, viene suggerito di caricare 20-100 ng di DNA per banda. Eventuali diluizioni del campione possono essere preparate in acqua deionizzata o tampone.

Il volume totale del campione per il caricamento è di 20 µL.

Per allestire la corsa è necessario anche il *DNA molecular weight marker*. A differenza dei campioni, per questo viene fatta una diluizione 1:20 e vengono

caricati sul gel solo 10 μ L. Il programma da impostare sulla cameretta elettroforetica (*E-Gel™ iBase™ device*) è “*E-Gel 0.8-2%*”. La corsa dura 26 minuti. Al termine della corsa si possono visualizzare le bande accendendo il transilluminatore (*E-Gel® Safe Imager™ Real-Time Transilluminator*), che mette in evidenza gli ampliconi.

3.4.7 Quantizzazione con *Agilent 2200 TapeStation System*

La quantizzazione con la *Agilent 2200 TapeStation System* (59) prevede una elettroforesi su gel d'agarosio. Il gel è contenuto in piccoli capillari all'interno di una *ScreenTape*, per cui si ottiene una separazione di frammenti di DNA, precedentemente amplificati, in modo molto rapido.

La procedura standard prevede l'allestimento di una provetta da 0,2 ml per ogni campione e una per il *ladder*. Il *ladder* è una miscela di DNA con peso molecolare noto, che viene inserito nelle corse elettroforetiche per permettere all'operatore di determinare le dimensioni dei frammenti processati. In ogni provetta vengono aggiunti 4 μ l di *D1000 HS Buffer* e 2 μ l di campione, o di *ladder*. Dopo aver caricato la *ScreenTape* ed i campioni nello strumento, si avvia la corsa (*Fig.26*).

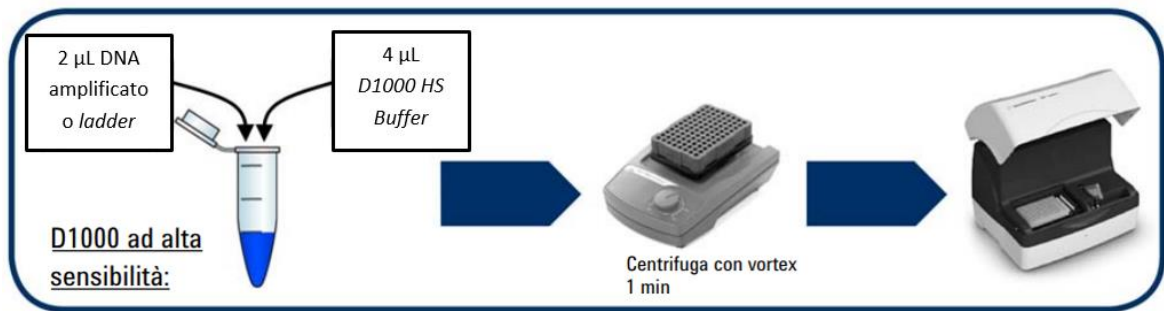


Fig.26 Allestimento elettroforesi capillare con Agilent 2200 TapeStation System (59).

Il risultato finale è dato da un'immagine del gel d'agarosio ad alta risoluzione, da un elettroferogramma e da una tabella in cui viene riportata la taglia dell'amplicone analizzato e una stima della sua concentrazione (Fig.27).

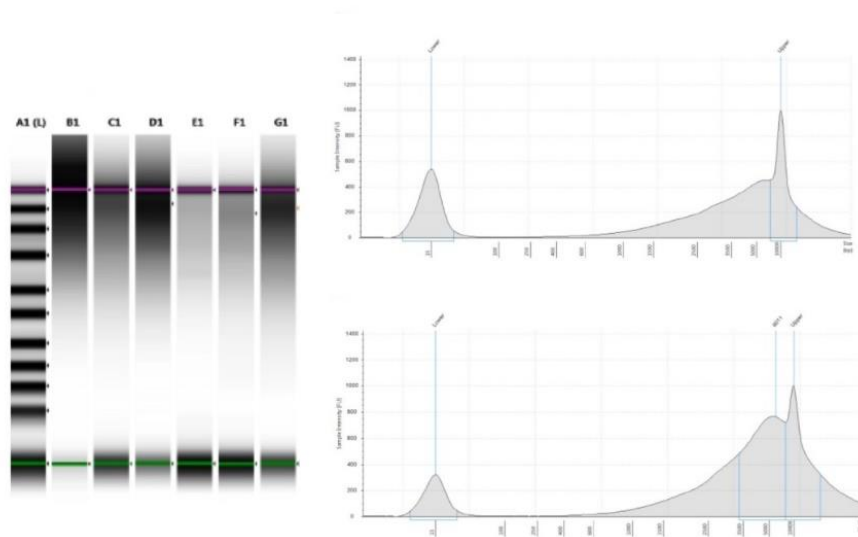


Fig.27 Risultati Agilent 2200 TapeStation System (59).

3.4.8 Massive Parallel Sequencing

Il sequenziamento è stato eseguito seguendo il protocollo *Precision ID Panels with the Ion PGM™ System* (60).

Amplificazione delle regioni target e preparazione delle librerie

La prima e importante fase è quella dell'amplificazione della regione di nostro interesse. Questo avviene grazie all'utilizzo del pool di *primer* disegnato come al punto 3.2 ("Disegno dei primers").

Allestire la PCR come in *Tab.14*, facendo attenzione a tenere in ghiaccio *5X Ion AmpliSeq™ HiFi Mix*.

COMPONENTE	VOLUME
5X Ion AmpliSeq™ HiFi Mix	4 µL
Precision ID Identity Panel	10 µL
DNA (1 ng)	x µL
Acqua nucleasi-free	6-x µL
TOTALE	20 µL

Tab.14 Mix di PCR per l'amplificazione della regione target.

La quantità di DNA è variabile in base alla concentrazione del DNA di partenza. La concentrazione ottimale suggerita dal kit è di 1 ng. Nel nostro caso sono state utilizzate quantità diverse di acqua e di ogni campione, le quali sono riportate nella seguente tabella (*Tab.15*).

CAMPIONE	QUANTITA' DI CAMPIONE	QUANTITA' DI ACQUA	CICLI DI AMPLIFICAZIONE
29MH_ASAN056	1 µL	5 µL	21 cicli
29MH_ASAN057	1 µL	5 µL	21 cicli
29MH_ASAN090	1 µL	5 µL	21 cicli
29MH_ASAN091	1 µL	5 µL	21 cicli
29MH_ASAN102	1 µL	5 µL	21 cicli
29MH_ASAN115	1 µL	5 µL	21 cicli
29MH_D1	1 µL	5 µL	21 e 25 cicli
29MH_D2	1 µL	5 µL	21 e 25 cicli
29MH_D3	1 µL	5 µL	21 e 25 cicli
29MH_D4	1 µL	5 µL	21 e 25 cicli
29MH_D5	1 µL	5 µL	21 e 25 cicli
29MH_D6	1 µL	5 µL	21 e 25 cicli
29MH_TS22	6 µL	----	21 e 25 cicli
29MH_TS26	2,16 µL	3,84 µL	21 e 25 cicli
29MH_TS29	2 µL	4 µL	21 e 25 cicli
29MH_BO04	1 µL	5 µL	25 cicli
29MH_BO05	2 µL	4 µL	25 cicli
29MH_TS04	6 µL	----	25 cicli
29MH_TS07	6 µL	----	25 cicli

Tab.15 Quantità di acqua e campioni utilizzate nell'amplificazione del target.

Le condizioni di amplificazione sono riportate in *Tab.16* e impostate su *Veriti™ 96-Well Thermal Cycler*.

PASSAGGIO	TEMPERATURA	TEMPO	CICLI
Attivazione dell'enzima	99°C	2 minuti	
Denaturazione	99°C	15 secondi	per 25 cicli
Appaiamento ed estensione	60°C	4 minuti	
	10°C	∞	

Tab.16 Condizioni di amplificazione.

Come riportato in *Tab.15*, sono stati utilizzati due protocolli di amplificazione: a 21 e 25 cicli, per testare quale dei due porta a un risultato di genotipizzazione finale migliore.

Il protocollo a 25 cicli è risultato il più performante dei due.

Parziale digestione degli ampliconi

A questo punto si procede con una parziale digestione dei campioni amplificati che viene effettuata con un reagente chiamato *FuPa Reagent*.

In particolare a questo punto abbiamo un amplicone *double strand* composto dalla regione di nostro interesse e dai due primer, *forward* e *reverse*. 2 μL di *FuPa Reagent* vengono aggiunti direttamente al prodotto di amplificazione, raggiungendo un volume totale di $\sim 22 \mu\text{L}$. Le provette vengono caricate su *Veriti™ 96-Well Thermal Cycler* con le seguenti condizioni (*Tab.17*).

TEMPERATURA	TEMPO
50°C	10 minuti
55°C	10 secondi
60°C	20 minuti
10°C	max 1h

Tab.17 Condizioni della reazione di digestione.

Ligazione dei barcode

È necessario diluire i *barcode* prima dell'uso. Ogni *barcode* deve essere diluito come in *Tab.18*.

COMPONENTE	QUANTITA'
Ion P1 Adapter	1 μ L
Ion Xpress™ Barcode X	1 μ L
Nuclease-free Water	2 μ L
TOTALE	4 μ L

Tab.18 Diluizione dei barcode. La X rappresenta il codice del barcode scelto.

I *barcode* così diluiti possono essere stoccati a -20°C fino a 10 giorni.

Successivamente i *barcode* diluiti vengono ligati a una estremità degli ampliconi. I *barcode* sono necessari al momento dell'analisi dei risultati per risalire al campione di partenza. In questo passaggio è necessario tenere la DNA ligasi in un porta provette ghiacciato. La preparazione della reazione di ligazione avviene come in *Tab.19*, facendo attenzione a seguire l'ordine dei reagenti da aggiungere agli ampliconi.

COMPONENTE	QUANTITA'
Switch Solution	4 μ L
Barcode X diluito	2 μ L
DNA ligasi	2 μ L
TOTALE	~30 μ L

Tab.19 Mix di PCR per la ligazione dei barcode.

Il caricamento delle provette su *Veriti™ 96-Well Thermal Cycler* segue le seguenti condizioni (*Tab.20*).

TEMPERATURA	TEMPO
22°C	30 minuti
72°C	10 minuti
10°C	max 1h

Tab.20 Condizioni della reazione di ligazione.

Purificazione delle librerie

Ora si procede con la purificazione delle librerie. In particolare si fanno aderire gli ampliconi a una resina, la *Agencourt™ AMPure™ XP Reagent*, per permettere l'eliminazione di tutte le componenti in eccesso usate nei passaggi precedenti. Gli step da seguire sono i seguenti:

- Trasferire la libreria in provette da 1,5 mL (il rack magnetico che viene utilizzato per la purificazione richiede questo tipo di provette);
- Vortexare per almeno un minuto la resina (*Agencourt™ AMPure™ XP Reagent*) per risospendere le biglie magnetiche;
- Aggiungere 45 µL di resina ad ogni campione;
- Incubare a temperatura ambiente per 5 minuti;

Intanto preparare etanolo al 70% per i lavaggi successivi con: 230 µL di etanolo e 100 µL di acqua nucleasi-free per ogni campione.

- Trasferire le provette con campione e resina su rack magnetico;
- Incubare a temperatura ambiente per 2 minuti (o finché la soluzione non risulta limpida);
- Rimuovere il surnatante limpido, evitando di toccare il precipitato;
- Aggiungere 150 μL di etanolo al 70% appena preparato;
- Ruotare le provette sul rack magnetico per 5-6 volte (controllare che il precipitato migri correttamente);
- Rimuovere il surnatante, evitando di toccare il precipitato;
- Aggiungere 150 μL di etanolo al 70% appena preparato;
- Ruotare le provette sul rack magnetico per 5-6 volte (controllare che il precipitato migri correttamente);
- Rimuovere il surnatante, evitando di toccare il precipitato;
- Fare una centrifuga veloce, facendo attenzione a posizionare il precipitato sulla parte esterna della centrifuga;
- Riposizionare le provette sul rack magnetico;
- Eliminare l'etanolo in eccesso che si è depositato sul fondo della provetta;
- Incubare a temperatura ambiente con il tappo della provetta aperto per 2 minuti (o finché l'etanolo non è completamente evaporato);
- Aggiungere 50 μL di TE facendo in modo di sciogliere il precipitato;

- Vortexare
- Incubare le provette a temperatura ambiente per 5 minuti;
- Centrifugare velocemente le provette;
- Posizionare le provette su rack magnetico;
- Incubare per 2 minuti;
- Prelevare il surnatante e trasferirlo su provette nuove, già codificate con il nome del campione.

Quantizzazione delle librerie

Dopo aver purificato ed eluito le librerie è necessario quantizzarle tramite PCR quantitativa (qPCR). Il kit usato per la quantizzazione è *Ion Library TaqMan® Quantitation Kit*.

Prima di procedere con l'allestimento della mix per la qPCR è necessario fare una diluizione 1:100 della libreria. Si preparano delle provette con 198 μL di acqua nucleasi free e 2 μL di libreria purificata. Queste diluizioni possono essere conservate a -20°C .

A questo punto si procede con la preparazione delle diluizioni seriali dello standard. Il kit contiene uno standard, *E. coli DH10B Ion Control Library* (~68 pM), che viene diluito alle concentrazioni elencate in *Tab.21*.

STANDARD	VOLUME LIBRERIA DI CONTROLLO	VOLUME ACQUA	CONCENTRAZIONE
Std1	5 μ L (<i>E. coli DH10B Ion Control Library</i>)	45 μ L	6,8 pM
Std2	5 μ L di Std1	45 μ L	0,68 pM
Std3	5 μ L di Std2	45 μ L	0,068 pM

Tab.21 Diluizioni del controllo *E. coli DH10B Ion Control Library* per la quantizzazione.

Preparare la mix di PCR per il numero totale di campioni, per il controllo negativo (1) e per gli standard appena diluiti (3) come da Tab.22.

COMPONENTE	QUANTITA'
Ion Library TaqMan® qPCR Mix	10 μ L
Ion Library TaqMan® Quantitation Assay, 20X	1 μ L
TOTALE	11 μ L

Tab.22 Mix di PCR per la quantizzazione delle librerie.

Aliquotare 11 μ L della mix nelle provette per la quantizzazione e aggiungere 9 μ L campione (standard, controllo negativo o librerie diluite). Il totale per ogni provetta è 20 μ L.

Le condizioni impostate su *Rotor Gene 6000 PCR System* sono le seguenti (Tab.23).

TEMPERATURA	TEMPO	CICLI
50°C	2 minuti	
95°C	20 secondi	
95°C	3 secondi	per 40 cicli
60°C	32 secondi	

Tab.23 Condizioni della reazione di quantizzazione.

Amplificazione aggiuntiva delle librerie con Platinum™

Una libreria che produce meno della concentrazione raccomandata può essere recuperata tramite una ulteriore amplificazione della libreria. Ovviamente è necessario effettuare anche una purificazione prima di procedere con il sequenziamento. In questo studio, 6 campioni hanno dato risultati bassi alla quantizzazione e quindi sono stati riamplificati con *Platinum™* (29MH_D1-1, 29MH_D2-1, 29MH_D3-1, 29MH_D5-1, 29MH_D4-2 e 29MH_TS22-2).

Unire 25 µL della libreria ottenuta dopo purificazione (*purificazione delle librerie*) con 72 µL di *Platinum™ PCR SuperMix HiFi* e 3 µL di *Library Amplification Primer Mix del Precision ID Library Kit*.

Fare vortex e centrifuga delle provette e poi caricare su *Veriti™ 96-Well Thermal Cycler* con le seguenti condizioni (*Tab.24*).

TEMPERATURA	TEMPO	CICLI
98°C	2 minuti	
98°C	15 secondi	per 40 cicli
64°C	1 minuto	
10°C	32 secondi	max 24h

Tab.24 Condizioni della reazione di amplificazione.

Purificazione delle librerie riamplicate con Platinum™

A questo punto si procede con un doppio step di purificazione utilizzando la resina *Agencourt™ AMPure™ XP Reagent*.

Il primo step di purificazione prevede un rapporto biglie/volume campione di 0,5X. In questo modo il DNA ad alto peso molecolare rimane adeso alla resina, mentre i gli ampliconi e i *primer* rimangono nel surnatante, che viene recuperato.

Nel secondo step di purificazione cambia il rapporto biglie/volume campione, che diventa di 1,2X. Così gli ampliconi sono legati alla resina e i *primer* rimangono in soluzione. Il protocollo è il seguente.

Primo step di purificazione:

- Inserire ~50 µL del campione in una provetta da 1,5 mL;
- Aggiungere 25 µL di *Agencourt™ AMPure™ XP Reagent* e miscelare bene;
- Incubare a temperatura ambiente per 5 minuti;
- Trasferire su rack magnetico;
- Incubare a temperatura ambiente per 5 minuti;
- Trasferire il surnatante in un'altra provetta da 1,5 mL

(Il surnatante contiene gli ampliconi desiderati)

Secondo step di purificazione:

- Aggiungere 60 μL di *Agencourt*TM *AMPure*TM *XP Reagent* e miscelare bene;
- Incubare a temperatura ambiente per 5 minuti;

Intanto preparare etanolo al 70% per i lavaggi successivi con: 230 μL di etanolo e 100 μL di acqua nucleasi-free per ogni campione.

- Trasferire su rack magnetico;
- Incubare a temperatura ambiente per 3 minuti;
- Rimuovere il surnatante limpido, evitando di toccare il precipitato;

(Il pellet contiene gli ampliconi desiderati)

- Aggiungere 150 μL di etanolo al 70%;
- Ruotare le provette sul rack magnetico per 5-6 volte (controllare che il precipitato migri correttamente);
- Rimuovere il surnatante (etanolo), evitando di toccare il precipitato;
- Aggiungere 150 μL di etanolo al 70% appena preparato;
- Ruotare le provette sul rack magnetico per 5-6 volte (controllare che il precipitato migri correttamente);
- Rimuovere il surnatante (etanolo), evitando di toccare il precipitato;
- Centrifuga veloce;

- Riposizionare le provette sul rack magnetico;
- Eliminare l'etanolo in eccesso;
- Incubare a temperatura ambiente con il tappo della provetta aperto per 2 minuti (o finché l'etanolo non è completamente evaporato);
- Aggiungere 50 μL di TE facendo in modo di sciogliere il precipitato;
- Vortexare
- Incubare le provette a temperatura ambiente per 2 minuti;
- Posizionare le provette su rack magnetico;
- Incubare a temperatura ambiente per 2 minuti;
- Prelevare il surnatante e trasferirlo su provette nuove.

(Il surnatante contiene gli ampliconi desiderati)

Una volta che i campioni sono stati riamplicati e purificati, è necessario quantizzare nuovamente la libreria.

3.4.9 *Emulsion PCR - Ion OneTouch 2™*

Prima di iniziare allestire *emulsion PCR* con *Ion OneTouch 2™*, è necessario portare tutte le librerie alla stessa concentrazione. In base alla concentrazione molare ottenuta a seguito della quantizzazione *Real-Time* con *Rotor-Gene*

6000 PCR System®, si allestiranno delle diluizioni specifiche per ogni diluizione. La concentrazione a cui portare le librerie è 8 pM.

A questo punto si crea un *pool* equimolare delle librerie mettendo 3 µL di libreria di ogni campione da analizzare (8pM) nella stessa provetta.

Dopo aver preparato l' *Ion OneTouch 2™* per iniziare la corsa (61), si procede con:

- Preparazione del campione come in *Tab.25*:

ORDINE	REAGENTE	VOLUME
1	Nuclease-free Water	25 µL
2	Ion PGM™ Hi-Q™ View Enzyme Mix	50 µL
3	Diluted library	25 µL
4	Ion PGM™ Hi-Q™ View ISPs	100 µL
	TOTALE	~1000µL

Tab.25 Ordine dei reagenti per la preparazione del campione per l'emulsion PCR.

- Vortex e centrifuga;
- Pipettare la soluzione appena preparata (~1000µL) nella *sample port* dell' *Ion OneTouch™ Reaction Filter*;
- Pipettare 850 µL di olio *Ion OneTouch™ Reaction Oil* nella *sample port* dell' *Ion OneTouch™ Reaction Filter*;
- Cambiare puntale;
- Pipettare 850 µL di olio *Ion OneTouch™ Reaction Oil* nella *sample port* dell' *Ion OneTouch™ Reaction Filter*;

- Invertire l'*Ion OneTouch™ Reaction Filter* e inserirlo nell'apposito alloggiamento;
- Avviare la corsa.

Una volta terminata la seduta di *emulsion PCR* è necessario recuperare il campione. In particolare il risultato finale dell'*emulsion PCR* è costituito da biglie a cui sono adese tante copie dello stesso amplicone. Ogni biglia possiede un solo tipo di amplicone sulla sua superficie. Tutti gli ampliconi sono da una parte adesi alla biglia e dall'altra legati a una molecola di biotina, fondamentale per lo step successiva di arricchimento.

Una volta terminata la seduta è necessario rimuovere subito le *Ion OneTouch™ Recovery Tubes* dalla centrifuga dello strumento e gettare l'*Ion OneTouch™ Recovery Router*.

- Eliminare il tutto il sovrnatante tranne 100 μL da ogni *Ion OneTouch™ Recovery Tubes*;
- Aggiungere 500 μL di *Ion OneTouch™ Wash Solution* ad ogni *Ion OneTouch™ Recovery Tubes* e pipettare per mettere le ISPs in sospensione;
- Recuperare le ISPs risospese e metterle in un'unica provetta;
- Centrifuga a 15.500 rcf per 2,5 minuti;
- Rimuovere tutto il sovrnatante tranne 100 μL .

3.4.10 Arricchimento/purificazione delle librerie - *Ion OneTouch™ ES*

Il campione ottenuto dall'*emulsion PCR* viene ora preparato per il caricamento su *Ion OneTouch™ ES*. Questo strumento lavora su una strip a 8 pozzetti (*Fig.28*) in cui è necessario mettere le soluzioni che servono per l'arricchimento.

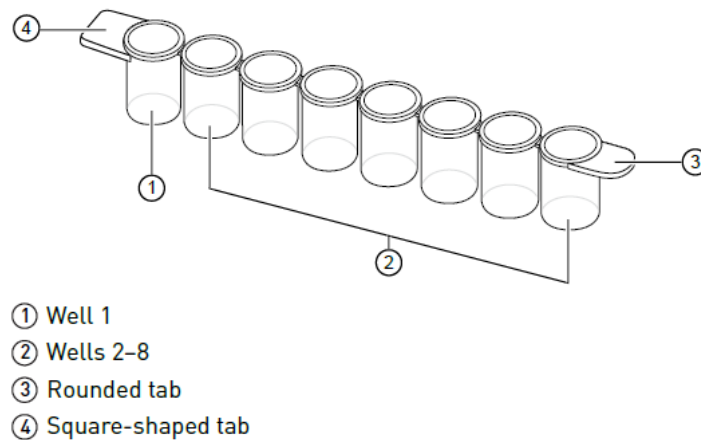


Fig.28 Strip a 8 pozzetti da caricare su Ion OneTouch™ ES (61).

- Preparare la *Melt Off Solution* come descritto in *Tab.26*:

ORDINE	REAGENTE	VOLUME
1	Tween™ Solution	280 µL
2	NaOH 1 M	40 µL
	Totale	320 µL

Tab.26 Preparazione della Melt Off Solution.

- Vortex per almeno 30 secondi della resina *Dynabeads™ MyOne™ Streptavidin C1 Beads*, per risospendere le biglie;

- Centrifuga per 2 secondi della resina *Dynabeads™ MyOne™ Streptavidin C1 Beads*;
- Trasferire 13 µL di *Dynabeads™ MyOne™ Streptavidin C1 Beads* in una provetta da 1,5 mL;
- Posizionare la provetta su rack magnetico per 2 minuti;
- Eliminare il surnatante, senza toccare il pellet;
- Aggiungere 130 µL di *MyOne™ Beads Wash Solution*;
- Rimuover la provetta dal rack magnetico;
- Vortex per 30 secondi;
- Centrifuga per 2 secondi;
- Riempire i pozzetti della strip da 8 pozzetti come in *Tab.27*:

POZZETTO	REAGENTE	VOLUME
1	Campione	~100µL
2	Dynabeads™ MyOne™ Streptavidin C1 Beads risospese in MyOne™ Beads Wash Solution	130 µL
3	Ion OneTouch™ Wash Solution	300 µL
4	Ion OneTouch™ Wash Solution	300 µL
5	Ion OneTouch™ Wash Solution	300 µL
6	Vuoto	----
7	Melt-Off Solution	300 µL
8	Vuoto	----

Tab.27 Ordine di inserimento dei reagenti nei pozzetti.

- Posizionare la strip da 8 pozzetti nell'apposito alloggiamento nell'*Ion OneTouch™ ES*;

- Preparare una provetta da 0,2 mL con 10 μ L di *Neutralization Solution*;
- Posizionare la provetta da 0,2 mL nell'alloggiamento a lato;
- Posizionare il nuovo puntale;
- Dare il via alla corsa di arricchimento/purificazione;
- Terminata la corsa, chiudere la provetta da 0,2 mL e invertirla per 5 volte;
- Gettare la strip da 8 pozzetti e il puntale.

3.4.11 Sequenziamento con *Ion PGM™ System*

Per sequenziare tramite *Ion PGM™ System* è necessario preparare il *plan* di lavoro on-line, sul sito della *ThermoFisher®*.

Il protocollo *Ion PGM™ Hi-Q™ View sequencing Kit* (62) indica che prima di iniziare una corsa di sequenziamento bisogna effettuare dei lavaggi dello strumento. In particolare si esegue un lavaggio con acqua 18 M Ω e uno con clorito. Questi lavaggi sono essenziali per stabilire un corretto valore di pH, necessario affinché il sequenziamento avvenga correttamente.

A questo punto è necessario preparare la piattaforma con l'inizializzazione. Si preparano le bottiglie per i lavaggi (3 *Wash Bottle*) e i quattro dNTP.

Dopodiché si procede con il caricamento del campione sull'*Ion Chip*.

- Scongellare il *Sequencing Primer* in ghiaccio;
- Vortex e centrifuga veloce di *Control Ion Sphere™ Particles*;
- Aggiungere 5 μL di *Control Ion Sphere™ Particles* direttamente nella provetta con il campione, uscita dalla *Ion OneTouch™ ES*;
- Centrifuga per 2 minuti a 15.500 rcf;
- Eliminare il surnatante lasciando $\sim 15 \mu\text{L}$ nella provetta;
- Aggiungere 12 μL di *Sequencing Primer*;
- Assicurarsi che il volume totale sia di $\sim 27 \mu\text{L}$ (eventualmente aggiungere *Annealing Buffer*);
- Caricare la provetta su termociclatore con le condizioni descritte in *Tab.28*:

TEMPERATURA	TEMPO
95°C	2 minuti
37°C	2 secondi

Tab.28 Condizioni della reazione.

- Aggiungere 3 μL *Ion PGM™ Hi-Q™ View Sequencing Polymerase* al campione (dopo esser stato su termociclatore);
- Incubare a temperatura ambiente per 5 minuti;
- Eliminare la maggior quantità possibile di liquido dal chip posizionando il chip a 45° e poggiando il puntale nella *loading port*;
- Centrifuga del chip;

- Riempire il chip con il campione ($\sim 30 \mu\text{L}$) a una velocità di $\sim 1 \mu\text{L}/\text{sec}$ (evitare l'introduzione di aria nel chip);
- Centrifuga del chip;
- Pipettare il campione dentro e fuori dal chip per favorire la dispersione delle ISPs su tutta la superficie del chip (pipettare lentamente per evitare la formazione di bolle);
- Centrifuga del chip;
- Caricare l'*Ion Chip* sul *Ion PGM™ System*.

Preparare il *plan online* e selezionarlo al momento dell'inizio del sequenziamento.

Capitolo 4

RISULTATI

4.1 Selezione dei *marker* e disegno dei *primer*

Il pannello MPS disegnato permette un'amplificazione in *multiplex* di 29 microplotipi (MH). Il pannello è composto da 15 MH formati da 2 SNPs, 6 MH da 3 SNPs, 7 MH da 4 SNPs e un solo microplotipo da 5 SNPs.

In alcuni casi sono stati eliminati degli SNPs poco informativi dal *locus* del microplotipo, poiché la lunghezza dell'amplicone superava il *range* prescelto.

Il totale di SNPs analizzati è quindi di 74 SNPs dispersi su 15 autosomi umani. Il *range* di lunghezza dell'amplicone prodotto va da 126 a 174 bp.

Tutti i microplotipi con le rispettive caratteristiche sono riportate in *Tab.29*.

Gli SNPs eliminati dal pannello sono in evidenza.

MH	Ae	PI	N° ALLELI	N° SNPs PER MH	N° ORIGINALE DI SNPs	CR	SNPs
mh02KK-073	2,74	0,22	3	2		chr2	rs1374748/rs7583554
mh02KK-134	6,81	0,05	12	3	4	chr2	rs12469721/rs3101043/ rs3111398/ rs72623112
mh02KK-136	4,53	0,09	6	3		chr2	rs6714835/rs6756898/ rs12617010
mh02KK-213	2,46	0,23	3	3		chr2	rs7568519/rs7577785/ rs1519654
mh03KK-007	2,58	0,22	3	2		chr3	rs4513489/rs6441961
mh03KK-009	1,98	0,35	3	2		chr3	rs3732783/rs6280
mh04KK-010	3,11	0,16	4	2		chr4	rs3135123/rs495367
mh04KK-011	2,57	0,23	3	2		chr4	rs6855439/rs6531591
mh04KK-013	4,72	0,09	6	3	5	chr4	rs13131164/rs3775866/ rs11725922/ rs3775867/ rs17088476
mh04KK-074	2,04	0,36	3	2		chr4	rs11932595/rs17085763
mh05KK-022	2,98	0,18	3	2		chr5	rs41461/rs41462
mh05KK-062	2,22	0,32	3	2		chr5	rs870348/rs870347
mh05KK-170	5,47	0,06	9	3	4	chr5	rs74865590 /rs438055/ rs370672/rs6555108
mh06KK-026	2,59	0,25	5	3		chr6	rs4565296/rs4431439/ rs179939
mh07KK-031	1,80	0,39	3	2		chr7	rs17168174/rs10246622
mh09KK-034	2,21	0,26	3	2		chr9	rs1408800/rs1408801
mh09KK-153	2,40	0,24	6	3		chr9	rs10125791/rs2987741/ rs7047561
mh11KK-039	2,34	0,26	3	2		chr11	rs2288159/rs10891537
mh12KK-043	2,12	0,30	4	3		chr12	rs11613749/rs11062734/ rs17780102
mh12KK-046	2,66	0,22	4	2		chr12	rs1503767/rs11068953
mh13KK-218	8,34	0,03	14	3	4	chr13	rs1927847/rs9536429/ rs7492234/ rs9536430
mh13KK-225	3,77	0,12	5	3		chr13	rs4884651/rs9529023/ rs7329287
mh16KK-302	1,92	0,32	5	4		chr16	rs1395579/rs1395580/ rs1395582/rs9939248
mh17KK-053	2,25	0,28	3	2		chr17	rs3760370/rs3760371
mh17KK-054	2,55	0,24	3	2		chr17	rs2233362/rs634370
mh18KK-293	2,05	0,32	6	4		chr18	rs621320/rs621340/ rs678179/rs621766
mh21KK-320	5,97	0,05	7	3	4	chr21	rs2838081 /rs2838082/ rs78902658/rs2838083
mh21KK-324	5,94	0,06	12	3	4	chr21	rs6518223 /rs2838868/ rs7279250/rs8133697
mh22KK-060	2,38	0,26	3	2		chr22	rs4818/rs4680

Tab.29 Microplotipi analizzati, valore di Ae, valore di Matching Probability e SNPs.

4.2 Risultati della quantizzazione delle librerie

I 6 campioni di DNA di controllo (non degradati) sono stati trattati secondo il protocollo *Precision ID Panels*, fino alla purificazione ed eluizione delle librerie. La quantizzazione delle librerie viene effettuata con *Agilent 2200 TapeStation System*®, che consente, a differenza della quantizzazione in *Real-Time PCR*, di stimare anche l'ampiezza delle regioni del DNA amplificate. In questo modo è stato effettuato anche un controllo della taglia degli ampliconi, secondo quanto atteso dal disegno del pannello. In *Tab.30* sono riportate le concentrazioni molari e le taglie degli ampliconi ottenuti dopo l'amplificazione con *Precision ID Panels*.

CAMPIONE	MOLARITÀ	TAGLIA
29 MH_ANAS056	2580 pg/μL	225bp
29 MH_ANAS057	1800 pg/μL	224bp
29 MH_ANAS090	1110 pg/μL	228bp
29 MH_ANAS091	1880 pg/μL	214bp
29 MH_ANAS102	1160 pg/μL	220bp
29 MH_ANAS115	2250 pg/μL	223bp

Tab.30 Risultati della quantizzazione delle librerie dei campioni a concentrazione nota.

La quantizzazione dei campioni per il test di sensibilità, per il test su DNA degradato e per i casework sample è stata effettuata in *Real-Time PCR* con *Rotor-Gene 6000 PCR System*®. In questo modo si ottengono valori molto più precisi rispetto alla quantizzazione con *Agilent 2200 TapeStation System*®. Inoltre ogni campione viene analizzato in doppio per avere

un'ulteriore conferma. In *Tab.31* sono riportate le concentrazioni ottenute. Il valore medio delle due concentrazioni viene moltiplicato per 100 perché prima della quantizzazione è stata eseguita una diluizione delle librerie 1:100.

CAMPIONE	qPCR-1 (ng)	qPCR-1 (ng)	MEDIA	CONCENTRAZIONE MEDIA CAMPIONE (ng/ul)
29 MH_D1-1	0,078	0,108	0,093	9,3228
29 MH_D2-1	0,032	0,039	0,036	3,5547
29 MH_D3-1	0,023	0,025	0,024	2,3714
29 MH_D4-1	0,116	0,143	0,130	12,9524
29 MH_D5-1	0,085	0,071	0,078	7,81245
29 MH_D6-1	0,134	0,138	0,136	13,5653
29MH_ TS22-1	0,115	0,110	0,113	11,2524
29MH_ TS26-1	0,238	0,229	0,233	23,34425
29MH_ TS29-1	0,391	0,277	0,334	33,41825
29MH_D1-2	3,964	4,902	4,433	443,2868
29MH_D2-2	1,086	1,208	1,147	114,7391
29MH_D3-2	0,929	0,580	0,754	75,44955
29MH_D4-2	0,077	0,059	0,068	6,81685
29MH_D5-2	0,157	0,122	0,140	13,9643
29MH_D6-2	0,150	0,140	0,145	14,50865
29MH_ TS22-2	0,091	0,105	0,098	9,7913
29MH_ TS26-2	1,814	1,976	1,895	189,51005
29MH_ TS29-2	0,763	0,747	0,755	75,48765
29MH_BO04	0,599	0,302	0,450	45,0379
29MH_BO05	0,183	0,203	0,193	19,2912
29MH_TS04	0,306	0,290	0,298	29,8348
29MH_TS07	1,088	1,394	1,241	124,1379

Tab.31 Risultati della quantizzazione delle librerie dei campioni per il test di sensibilità, dei campioni per il test su DNA degradato e dei casework sample.

4.3 Risultati della quantizzazione delle librerie riamplicate e purificate con *Platinum*TM

Alcune librerie hanno ottenuti bassi valori di DNA alla quantizzazione con *Rotor-Gene 6000 PCR System*[®]. Tutti i campioni che hanno ottenuto valori inferiori a 10 ng/ μ L sono stati riamplicati e purificati. Dopo la purificazione e l'eluizione delle librerie, i campioni sono stati nuovamente quantizzati. I risultati ottenuti dai campioni quantizzati con *Agilent 2200 TapeStation System*[®] sono riportati in *Tab.32* I risultati dei campioni quantizzati con *Rotor-Gene 6000 PCR System*[®] sono riportati in *Tab.33*.

CAMPIONE	TAGLIA	MOLARITA'
29 MH_D1-1	208 bp	1620 pg/ μ L
29 MH_D2-1	223 bp	643 pg/ μ L
29 MH_D3-1	236 bp	100 pg/ μ L
29 MH_D5-1	178 bp	2 pg/ μ L

Tab.32 Risultati della quantizzazione con Agilent 2200 TapeStation System[®].

CAMPIONE	qPCR-1 (ng)	qPCR-1 (ng)	MEDIA	CONCENTRAZIONE MEDIA CAMPIONE (ng/ μ l)
29 MH_D4-2	0,077	0,059	0,807	80,77 ng/ μ L
29MH_ TS22-2	0,091	0,105	1,009	100,95 ng/ μ L

Tab.33 Risultati della quantizzazione con Rotor-Gene 6000 PCR System[®].

Le librerie ottenute con questa seconda fase di riamplicazione sono state utilizzate per il caricamento su *Ion Chip*.

4.4 Risultati del sequenziamento e analisi del *coverage*

È stata eseguita un'unica corsa di sequenziamento su *Ion PGM™ System*, utilizzando un *Ion 318™ Chip v2*. Il *coverage* medio ottenuto è 4307,68. Il *report* della corsa è riportato in Fig.29.

Run Summary

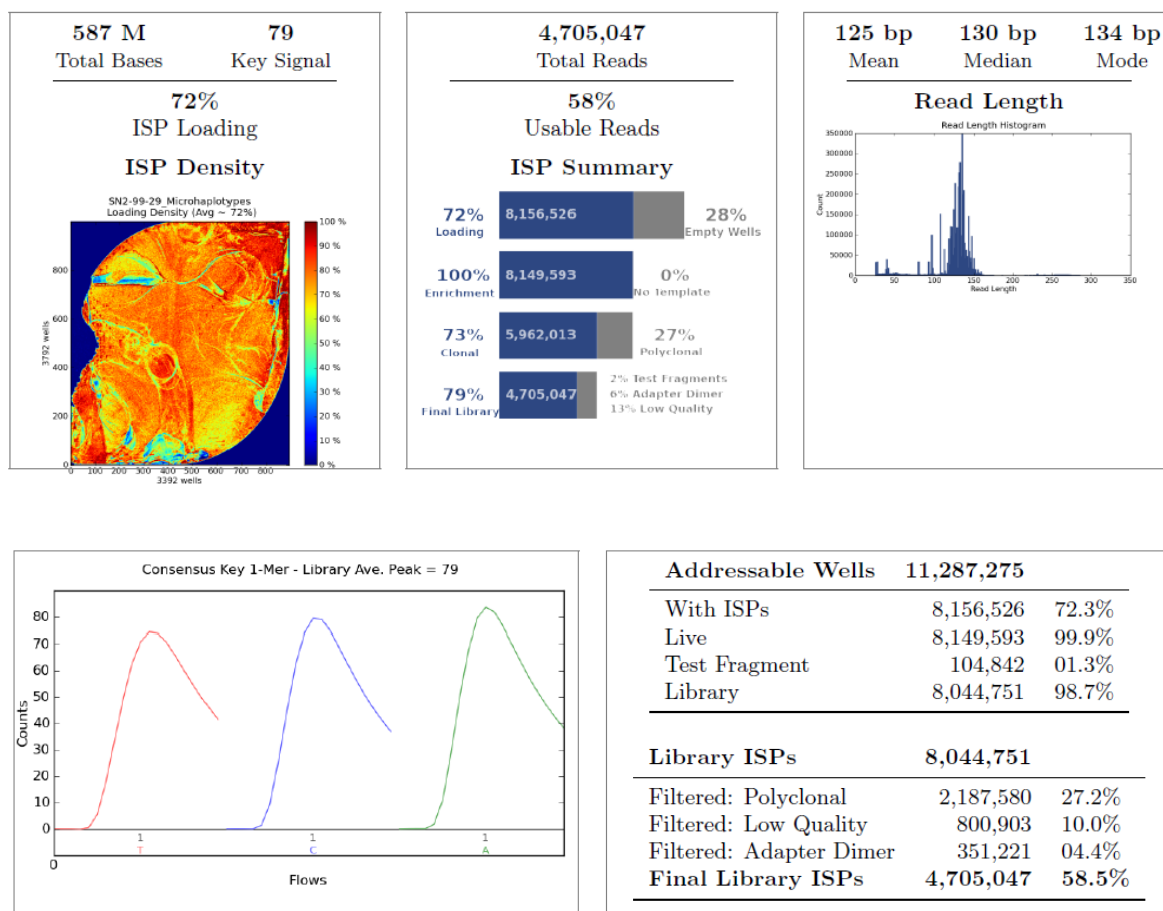


Fig.29 Run summary della corsa di sequenziamento.

Dal *Run Summary* (Fig.29), nel primo riquadro, è possibile vedere il totale di basi sequenziate (587 Mb), la percentuale di caricamento dell'*Ion Chip* espressa come densità di ISPs (72%) e la rappresentazione grafica del

caricamento del chip, in cui le zone rosse rappresentano aree ad alta densità di ISPs. Nel secondo riquadro sono riportati i valori relativi alle *reads*. Il totale di *reads* è 4.705.047, di queste solo il 58% è utilizzabile. Infatti circa 8 milioni di *reads* sono state caricate correttamente, ma alcune sono state scartate poiché risultavano essere policlonali. Per *reads* policlonale si intende una *read* sequenziata in una ISP in cui non è presente un singolo tipo di amplicone, ma più ampliconi diversi. Delle *reads* rimaste (circa 5 milioni) è necessario eliminare tutte quelle a bassa qualità e i dimeri degli adattatori, utilizzati per creare la libreria. Nel terzo riquadro è riportato un grafico che rappresenta la dispersione delle lunghezze di tutte le *reads*. Il *target* è mantenuto poiché la media delle *reads* che ci si aspetta è di 135,48 bp. Il valore della media ottenuto è 125 bp, mentre la moda è 134 bp: in linea con le aspettative. Successivamente è riportato il grafico di incorporazione che mostra le letture medie del segnale per i flussi delle basi T, C e A nella libreria. Nella tabella a lato sono invece riportati i dettagli di caricamento del chip e delle librerie adese alle ISPs. Di seguito vengono elencati i *barcode* utilizzati, il numero di basi sequenziate, il numero di basi sequenziate *on target*, numero di *reads* e lunghezza media degli ampliconi di ogni campione (Tab.34).

BARCODE	CAMPIONE	BASI	$\geq Q20$	READS	MEAN READ LENGTH
IonXpress_017	29MH_D1-1	16,970,456	16,476,052	133,707	127 bp
IonXpress_018	29MH_D2-1	14,909,778	14,467,894	115,312	129 bp
IonXpress_019	29MH_D3-1	9,357,898	9,105,087	73,329	128 bp
IonXpress_020	29MH_D4-1	3,145,234	3,053,373	25,788	122 bp
IonXpress_021	29MH_D5-1	1,599,454	1,554,944	13,338	120 bp
IonXpress_022	29MH_D6-1	1,772,274	1,722,740	15,009	118 bp
IonXpress_023	29MH_TS22-1	2,350,010	2,170,370	34,654	68 bp
IonXpress_024	29MH_TS26-1	16,865,568	16,412,929	133,52	126 bp
IonXpress_025	29MH_TS29-1	8,563,358	8,235,868	91,669	93 bp
IonXpress_026	29 MH_D1-2	23,140,647	22,481,227	173,112	134 bp
IonXpress_027	29 MH_D2-2	32,393,819	31,506,318	252,686	128 bp
IonXpress_028	29 MH_D3-2	18,831,011	18,327,134	148,745	127 bp
IonXpress_029	29 MH_D4-2	15,989,907	15,571,069	124,898	128 bp
IonXpress_030	29 MH_D5-2	14,064,907	13,681,471	112,182	125 bp
IonXpress_031	29 MH_D6-2	17,312,347	16,850,588	138,92	125 bp
IonXpress_032	29MH_TS22-2	389,464	376,739	3,597	108 bp
IonXpress_033	29MH_TS26-2	11,807,956	11,473,365	90,607	130 bp
IonXpress_034	29MH_TS29-2	9,489,115	9,188,489	84,381	112 bp
IonXpress_035	29MH_BO04	20,967,068	20,396,705	168,295	125 bp
IonXpress_036	29MH_BO05	13,768,270	13,406,112	118,659	116 bp
IonXpress_037	29MH_TS04	21,087,402	20,138,375	146,684	144 bp
IonXpress_038	29MH_TS07	18,869,235	18,374,707	148,869	127 bp
IonXpress_049	29MH_ASAN056	21,990,454	21,390,848	172,609	127 bp
IonXpress_050	29MH_ASAN057	13,796,076	13,415,117	109,341	126 bp
IonXpress_051	29MH_ASAN090	21,733,272	21,157,840	171,773	127 bp
IonXpress_052	29MH_ASAN091	13,945,242	13,575,191	109,892	127 bp
IonXpress_053	29MH_ASAN102	12,684,333	12,309,537	100,238	127 bp
IonXpress_054	29MH_ASAN115	13,170,839	12,797,933	103,512	127 bp

Tab.34 Numero di basi e di reads per ogni campione.

Di seguito sono riportati i valori di *coverage* ottenuti dal sequenziamento di ogni campione (*Tab.35*).

CAMPIONE	BARCODE	MAPPED READS	ON TARGET	MEAN DEPTH	UNIFORMITY
29MH_D1-1	IonXpress_017	133395	99.54%	4553	96.66%
29MH_D2-1	IonXpress_018	115111	95.87%	3794	96.66%
29MH_D3-1	IonXpress_019	73152	99.08%	2498	96.57%
29MH_D4-1	IonXpress_020	25601	94.86%	827.8	96.66%
29MH_D5-1	IonXpress_021	13245	92.85%	418.7	100.00%
29MH_D6-1	IonXpress_022	14844	90.07%	457.9	93.12%
29MH_TS22-1	IonXpress_023	34075	24.36%	277.1	89.50%
29MH_TS26-1	IonXpress_024	133285	99.11%	4520	92.99%
29MH_TS29-1	IonXpress_025	91080	54.79%	1695	100.00%
29 MH_D1-2	IonXpress_026	172849	99.71%	5911	89.50%
29 MH_D2-2	IonXpress_027	252300	99.59%	8630	89.50%
29 MH_D3-2	IonXpress_028	148505	99.65%	5042	92.99%
29 MH_D4-2	IonXpress_029	124656	99.47%	4280	86.07%
29 MH_D5-2	IonXpress_030	111684	97.10%	3708	89.59%
29 MH_D6-2	IonXpress_031	138309	96.94%	4619	92.99%
29MH_TS22-2	IonXpress_032	3576	78.02%	94.56	82.38%
29MH_TS26-2	IonXpress_033	90441	99.54%	3088	92.84%
29MH_TS29-2	IonXpress_034	84141	82.58%	2317	92.99%
29MH_BO04	IonXpress_035	167919	98.80%	5530	96.33%
29MH_BO05	IonXpress_036	118382	91.91%	3470	92.99%
29MH_TS04	IonXpress_037	62482	24.98%	528.6	92.99%
29MH_TS07	IonXpress_038	147656	98.87%	4946	96.33%
29MH_ASAN056	IonXpress_049	172308	99.74%	5884	100.00%
29MH_ASAN057	IonXpress_050	109160	99.70%	3725	100.00%
29MH_ASAN090	IonXpress_051	171531	99.59%	5846	100.00%
29MH_ASAN091	IonXpress_052	109717	99.72%	3757	100.00%
29MH_ASAN102	IonXpress_053	100064	99.71%	3417	100.00%
29MH_ASAN115	IonXpress_054	103322	99.74%	3539	100.00%

Tab.35 Risultati di coverage del sequenziamento di tutti i campioni.

Nello specifico sono riportati i *barcode* utilizzati per ogni campione e altri quattro valori: *mapped reads*, *on target*, *mean depth* e *uniformity*. Per *mapped reads* si intendono tutte quelle basi che mappano all'interno del genoma di riferimento (hg19). Il valore *on target* indica il numero di *reads* che, tra le *mapped reads*, ricadono all'interno del *target* scelto. Dopodiché viene

indicato il *mean depth*, che rappresenta il *coverage* medio di tutti i 29 ampliconi per ogni campione. Infine l'*uniformity* indica che il *target* è stato sequenziato in maniera uniforme. Qui è riportata una tabella con i *coverage* specifici di ogni amplicone; vengono evidenziati quelli ≤ 50 (Tab.36).

CAMPIONE	29MH_ ASAN056	29MH_ ASAN057	29MH_ ASAN090	29MH_ ASAN091	29MH_ ASAN102	29MH_ ASAN115
BARCODE	49	50	51	52	53	54
mh02KK-073	3212	1559	2866	2356	1954	2190
mh02KK-213	9262	5709	9319	5836	5216	5943
mh02KK-134	7634	4178	6266	4378	4380	4568
mh02KK-136	5376	3147	5724	3215	3067	3240
mh03KK-007	5457	3833	5632	3938	3531	3705
mh03KK-009	7076	4566	8306	4852	4002	4608
mh04KK-010	5845	4078	6298	3326	3415	2810
mh04KK-011	3142	1791	3009	2054	1785	1946
mh04KK-074	2584	1598	2480	1564	1454	1823
mh04KK-013	4985	3393	7403	3694	3004	3533
mh05KK-170	3936	3633	3837	2940	3304	2631
mh05KK-062	8272	4565	8183	4952	4278	4412
mh05KK-022	9166	5559	9298	5410	4681	5383
mh06KK-026	7261	4605	6262	4466	4266	4269
mh07KK-031	6397	4404	5347	4254	4116	4526
mh09KK-034	2958	1287	3588	1875	1270	1423
mh09KK-153	3471	2299	3435	2429	2023	2012
mh11KK-039	6729	4388	6623	4063	3918	3947
mh12KK-043	7946	5339	8220	5623	4987	5441
mh12KK-046	3745	2428	3961	2081	2334	2336
mh13KK-218	6590	3828	5702	3630	3257	3991
mh13KK-225	5620	3568	4877	3478	3324	2712
mh16KK-302	3516	2512	4961	2265	2358	2334
mh17KK-053	7399	4562	6990	5165	4218	5117
mh17KK-054	9067	5008	7709	5630	5102	4694
mh18KK-293	6570	4788	6426	3723	3890	3659
mh21KK-320	5154	3627	4652	3892	3362	2552
mh21KK-324	8829	6133	8331	5578	4974	4723
mh22KK-060	4659	2443	5125	2743	2303	2521

Tab.36A Valori di coverage dei campioni non degradati di controllo.

CAMPIONE	29MH_ D1-1	29MH_ D2-1	29MH_ D3-1	29MH_ D4-1	29MH_ D5-1	29MH_ D6-1	29MH_ TS22-1	29MH_ TS26-1	29MH_ TS29-1
BARCODE	17	18	19	20	21	22	23	24	25
mh02KK-073	3166	2566	3043	917	211	375	307	3368	1448
mh02KK-213	7147	5558	3869	1127	808	779	540	7651	2651
mh02KK-134	1242	920	622	235	114	91	45	853	453
mh02KK-136	4824	3465	3016	951	489	240	156	4059	2012
mh03KK-007	5221	4241	5039	400	418	542	145	4456	2295
mh03KK-009	6685	5135	3233	1365	453	659	298	5912	2443
mh04KK-010	3873	2710	1249	1120	502	564	534	5552	2266
mh04KK-011	2541	2319	1485	834	339	138	174	2919	886
mh04KK-074	2907	2822	1665	297	142	302	330	2534	1104
mh04KK-013	5154	4735	3731	859	348	559	323	4673	2300
mh05KK-170	1840	1197	488	727	245	487	136	3257	1050
mh05KK-062	7370	6457	4886	1575	921	556	317	6733	2050
mh05KK-022	5886	4890	2019	809	342	254	148	5434	1911
mh06KK-026	5745	5027	3015	1066	627	697	450	6167	1933
mh07KK-031	5040	3990	1990	1241	568	825	213	4544	2025
mh09KK-034	865	718	696	136	89	38	37	722	409
mh09KK-153	1958	1769	1036	314	184	305	70	2312	800
mh11KK-039	5105	4893	3077	709	493	560	309	6132	1854
mh12KK-043	6834	7001	3759	1136	591	946	457	7577	2359
mh12KK-046	4033	3790	3055	672	468	355	231	3481	1501
mh13KK-218	4076	3588	2398	883	446	335	414	4022	1473
mh13KK-225	3802	2943	1653	811	370	297	522	4333	1437
mh16KK-302	4116	2458	1918	427	356	392	142	3553	1013
mh17KK-053	7160	6345	4224	1686	963	784	560	5849	2841
mh17KK-054	8188	6401	3597	1168	685	545	592	6476	3455
mh18KK-293	2148	1943	1326	512	248	322	53	2361	738
mh21KK-320	4160	3502	2525	561	220	468	231	3570	1455
mh21KK-324	7754	5368	2201	1079	495	735	394	9358	2065
mh22KK-060	3938	3601	1662	667	163	220	173	4236	1672

Tab.36B Valori di coverage di ogni microaplotipo dei campioni per il test di sensibilità e dei campioni degradati artificialmente a 21 cicli di amplificazione.

CAMPIONE	29 MH _D1-2	29 MH _D2-2	29 MH _D3-2	29 MH _D4-2	29 MH _D5-2	29 MH _D6-2	29MH_ TS22-2	29MH_ TS26-2	29MH_ TS29-2
BARCODE	26	27	28	29	30	31	32	33	34
mh02KK-073	4620	8423	3825	2828	1696	3352	8	2569	1503
mh02KK-213	10532	14545	8409	4169	2518	6456	173	5894	3695
mh02KK-134	337	508	305	409	206	295	6	183	194
mh02KK-136	6580	9419	4032	4895	4907	4324	120	3191	2091
mh03KK-007	7584	9707	5567	11673	3041	6539	121	4184	3054
mh03KK-009	9523	14322	8388	4141	6629	5070	137	4646	3360
mh04KK-010	5432	8757	7245	3530	5487	4344	127	2762	4828
mh04KK-011	2942	4076	2521	3034	2051	2073	51	1563	1372
mh04KK-074	3281	5123	3111	2512	2334	1616	49	1554	1622
mh04KK-013	7312	10151	5156	7738	5504	4810	117	3786	2445
mh05KK-170	3414	5784	3680	719	2353	3446	55	1643	1681
mh05KK-062	9080	12017	7669	6107	6169	5371	247	4973	3481
mh05KK-022	7195	11756	7075	2038	5772	8882	133	3891	2573
mh06KK-026	7803	11008	7200	5557	3754	8326	125	3319	2676
mh07KK-031	6259	10488	5151	3717	5748	5646	96	3792	2135
mh09KK-034	1179	1016	405	500	358	542	10	1677	416
mh09KK-153	2471	3781	2528	1567	527	1390	8	1274	1336
mh11KK-039	6758	9645	6100	5487	3518	5450	54	3576	2692
mh12KK-043	10272	12291	8201	7989	6425	9957	116	4707	3010
mh12KK-046	5093	5793	3648	5699	3671	3020	122	2621	2468
mh13KK-218	5595	9533	4998	1200	3634	3668	88	2616	2457
mh13KK-225	3879	4882	4271	3075	2831	2028	99	2477	3144
mh16KK-302	4404	5773	3230	3370	4147	3724	62	2465	1572
mh17KK-053	9605	15760	7177	17347	4394	6295	154	5657	2893
mh17KK-054	10563	15551	9044	4929	6321	9407	226	5084	5744
mh18KK-293	1029	1291	1098	520	944	1259	16	516	556
mh21KK-320	4689	6580	3750	2607	3172	3019	82	2527	1399
mh21KK-324	9244	14709	9136	4123	7872	10754	146	4526	2533
mh22KK-060	5679	8583	5067	2521	2462	3011	42	2348	2552

Tab.36C Valori di coverage di ogni microaplotipo dei campioni per il test di sensibilità e dei campioni degradati artificialmente a 25 cicli di amplificazione.

CAMPIONE	29MH_BO04	29MH_BO05	29MH_TS04	29MH_TS07
BARCODE	35	36	37	38
mh02KK-073	3081	952	454	3063
mh02KK-213	9660	6729	1379	9074
mh02KK-134	478	140	78	960
mh02KK-136	3790	1509	629	4440
mh03KK-007	7239	2671	859	5862
mh03KK-009	8184	6051	796	8298
mh04KK-010	8972	15674	473	6948
mh04KK-011	3960	2612	381	1925
mh04KK-074	3335	1558	256	2924
mh04KK-013	8548	2450	840	5656
mh05KK-170	3051	2432	285	3733
mh05KK-062	4933	2210	378	5493
mh05KK-022	6146	3987	300	6423
mh06KK-026	7201	5874	602	4257
mh07KK-031	6064	3008	517	5432
mh09KK-034	2354	513	98	1074
mh09KK-153	3767	2821	180	2547
mh11KK-039	8431	5297	441	6699
mh12KK-043	5259	3241	930	8203
mh12KK-046	3318	1434	368	3707
mh13KK-218	4656	1926	555	6107
mh13KK-225	8867	5515	242	4792
mh16KK-302	2187	1086	348	2956
mh17KK-053	8066	3033	976	8211
mh17KK-054	16357	10961	1587	9390
mh18KK-293	1350	908	166	2172
mh21KK-320	3857	986	394	3314
mh21KK-324	5898	998	433	5739
mh22KK-060	6903	12231	665	6584

Tab.36D Valori di coverage di ogni microplotipo dei campioni casework sample.

4.5 Risultati della genotipizzazione dei campioni

A questo punto si procede con la genotipizzazione vera e propria. Per farlo è stato utilizzato il software IGV (*Integrative Genomics Viewer*) (63).

Il *software* permette di allineare le sequenze ottenute dal sequenziamento con la sequenza di riferimento (nel nostro caso hg19) e fare la tipizzazione di ogni campione per ogni *locus* scelto. IGV necessita inoltre di un file BED contenente i riferimenti delle regioni cromosomiche analizzate (*target file*) e la lista degli SNPs tipizzati con le relative localizzazioni cromosomiche. Durante l'analisi dei dati di sequenziamento sono stati considerati sistemi eterozigoti solo quelli in cui la frequenza dell'allele minoritario superava il 20%. Il *coverage* minimo considerato era pari a 50.

Per quanto riguarda i 6 campioni di DNA non degradato, i risultati del sequenziamento sono i seguenti.

CAMPIONE	BARCODE	MAPPED READS	ON TARGET	MEAN DEPTH	UNIFORMITY
29MH_ASAN056	IonXpress_049	172308	99.74%	5884	100.00%
29MH_ASAN057	IonXpress_050	109160	99.70%	3725	100.00%
29MH_ASAN090	IonXpress_051	171531	99.59%	5846	100.00%
29MH_ASAN091	IonXpress_052	109717	99.72%	3757	100.00%
29MH_ASAN102	IonXpress_053	100064	99.71%	3417	100.00%
29MH_ASAN115	IonXpress_054	103322	99.74%	3539	100.00%

Tab.35A Risultati del coverage del sequenziamento dei campioni di DNA non degradato.

Come da Tab.35A, questi 6 campioni risultano avere elevati valori di *coverage* (da 3417 a 5884); il valore di *uniformity* è uguale per tutti i campioni ed è pari a 100%. Anche la percentuale di *reads on target* è in tutti i casi maggiore di 99%.

La genotipizzazione di questi campioni è riportata in *Tab.37*.

CAMPIONE	29MH_ ASAN056	29MH_ ASAN057	29MH_ ASAN090	29MH_ ASAN091	29MH_ ASAN102	29MH_ ASAN115
mh02KK-073	GT/GC	TT/TT	GT/GT	GC/GC	GT/GC	GC/TT
mh02KK-213	TGT/CGT	TGC/TGT	TGC/TGT	TGC/CGT	TGC/TGT	TGC/TGT
mh02KK-134	ACT/ACT	ACT/TTT	ACT/TTT	ATC/TTT	ATC/ACC	ACT/ACT
mh02KK-136	TCC/GTC	GTA/GTA	TTC/GTC	TCC/TCC	GTC/GTA	TTC/GTC
mh03KK-007	TT/TC	TC/CC	TT/TT	TT/CC	TT/TC	TT/TT
mh03KK-009	TT/TT	TC/TT	TC/TT	TC/TC	TT/CC	TT/TT
mh04KK-010	AG/AG	AA/AA	GA/GA	GA/GA	GA/GG	AA/AA
mh04KK-011	AC/AC	AC/AC	AT/GT	AC/AT	AC/AT	AT/GT
mh04KK-074	AT/AT	AT/AT	AT/AT	AT/AC	AT/GT	AT/AT
mh04KK-013	CGG/CGA	CGG/CGA	CGG/AAG	CGG/CGG	CGA/CGA	CGG/CAG
mh05KK-170	GGA/GGG	AGA/AGA	GGA/GGG	AGA/AGA	AGA/AGA	GGA/GAG
mh05KK-062	AA/TA	TA/TA	AA/TA	AA/TA	AA/AA	AA/AC
mh05KK-022	TC/CC	CC/CA	CC/CA	TC/CA	TC/TC	TC/CC
mh06KK-026	GCG/GTG	GTG/GTG	GCA/GCA	GCG/GTG	GCG/GTG	GCG/GCG
mh07KK-031	CA/CG	CA/CG	CA/CG	CA/CA	CG/CG	CA/CA
mh09KK-034	GA/GA	AA/AA	GA/AA	AA/AA	AA/AA	AA/AA
mh09KK-153	TAA/TAA	TAA/TAA	TAA/CAC	TAA/TAC	CAA/CAA	TAC/CAA
mh11KK-039	GG/GG	GT/GG	GT/GG	GT/GT	GT/GG	GG/TT
mh12KK-043	CTG/CTG	CTG/CTG	CTG/CTG	CTG/CTG	CTG/CTA	CTG/CTG
mh12KK-046	TA/GG	TA/TA	TG/TA	TG/TG	TG/TA	TA/TA
mh13KK-218	TTC/CTC	TTT/CTT	TTT/CTT	TTT/CTT	TTT/TTC	CCC/CCC
mh13KK-225	ACG/ACG	GAG/GCG	GAG/GAA	GAG/GCG	GAA/GCG	ACG/ACG
mh16KK-302	ACTT/ACTT	GCTT/ACTT	ACTT/ACTT	ACTT/ACTT	GTTT/ACTT	ACTT/ACTT
mh17KK-053	TT/TC	CT/TC	CT/TC	CT/TT	CT/TC	CT/CT
mh17KK-054	GG/GG	AG/GG	GG/GG	AA/AA	AA/GG	AG/AA
mh18KK-293	ATGA/AGAA	AGAA/GGAA	ATGA/AGAA	AGAA/GGAA	ATGA/AGAA	AGAA/AGAA
mh21KK-320	GCG/ACG	ACA/ACA	ACA/ACG	GCA/ACA	ACA/ACA	ACG/GCG
mh21KK-324	CTG/CTG	CTA/TTA	CTG/CAG	CTG/CTG	CTA/TTA	CTG/CTA
mh22KK-060	CA/CA	CA/GG	GG/GG	CA/GG	CA/CA	CA/GG

Tab.37 Genotipizzazione dei campioni non degradati.

Il test di sensibilità ha ottenuto risultati diversi a 21 e 25 cicli di amplificazione. In *Tab.35B* sono riportati i risultati del sequenziamento.

CAMPIONE	BARCODE	MAPPED READS	ON TARGET	MEAN DEPTH	UNIFORMITY
29MH_D1-1	IonXpress_017	133395	99.54%	4553	96.66%
29MH_D2-1	IonXpress_018	115111	95.87%	3794	96.66%
29MH_D3-1	IonXpress_019	73152	99.08%	2498	96.57%
29MH_D4-1	IonXpress_020	25601	94.86%	827.8	96.66%
29MH_D5-1	IonXpress_021	13245	92.85%	418.7	100.00%
29MH_D6-1	IonXpress_022	14844	90.07%	457.9	93.12%
29 MH_D1-2	IonXpress_026	172849	99.71%	5911	89.50%
29 MH_D2-2	IonXpress_027	252300	99.59%	8630	89.50%
29 MH_D3-2	IonXpress_028	148505	99.65%	5042	92.99%
29 MH_D4-2	IonXpress_029	124656	99.47%	4280	86.07%
29 MH_D5-2	IonXpress_030	111684	97.10%	3708	89.59%
29 MH_D6-2	IonXpress_031	138309	96.94%	4619	92.99%

Tab.35B Risultati del coverage del sequenziamento dei campioni di DNA non degradato.

Dalla tabella si vede che i valori di *coverage* (*meand depth*) sono tutti molto alti ad eccezione di quelli ottenuti per D5-1 e D6-1. A differenza dei risultati ottenuti per i campioni DNA non degradato, i valori di *uniformity* sono più bassi, ma comunque non scendono oltre la soglia dell'86%. Anche i valori delle *reads on target* sono molto alti (Tab.35B). Di seguito sono riportate la legenda (Tab.38) e le tabelle delle genotipizzazioni dei campioni per il test di sensibilità (Tab.39 e Tab.40).

	DROP-IN ALLELICO
	DROP-OUT ALLELICO
	NO CALL
	COVERAGE ≤ 50
	TRE ALLELI

Tab.38 Legenda colori della genotipizzazione.

CAMPIONE	29MH_D1-1	29MH_D2-1	29MH_D3-1	29MH_D4-1	29MH_D5-1	29MH_D6-1
mh02KK-073	TT/GC	TT/GC	TT/GC	TT/GC	TT/GC	TT/GC
mh02KK-213	CGT/TGT	CGT/TGT	CGT/TGT	CGT/TGT	CGT/TGT	CGT/TGC
mh02KK-134	TTT/ACT	TTT/ACT	TTT/ACT	TTT/ACT	TTT/ACT	TTT/ACT
mh02KK-136	GTC/TCC	GTC/TCC	GTC/TCC	GTC/TCC	GTC/TCC	GTC/TTC
mh03KK-007	TC/TC	TC/TC	TC/TC	TC/TC	TC/TC	TC/TT
mh03KK-009	TT/TT	TT/TT	TT/TT	TT/TT	TT/TT	TT/TC
mh04KK-010	GA/AA	GA/AA	GA/AA	GA/AA	GA/AA	GA/AA
mh04KK-011	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	GT/AT
mh04KK-074	GT/GT	GT/GT	GT/GT	GT/GT	GT/AT	GT/AT
mh04KK-013	CGG/CGG	CGG/CGG	CGG/CGG	CGG/CGG	CGA/CGG	CGA/CGG
mh05KK-170	AAG/AGA	AAG/AGA	AAG/AGA	AAG/AGA	AAG/AGA	AAG/AGA
mh05KK-062	TA/AA	TA/AA	TA/AA	TA/AA	TA/AA	TA/AA
mh05KK-022	CC/TC	CC/TC	CC/TC	CC/TC	CC/TC	CC/TC
mh06KK-026	GTG/GCA	GTG/GCA	GTG/GCA	GTG/GCA	GTG/GCA	GTG/GCA
mh07KK-031	CG/CG	CG/CG	CG/CG	CG/CG	CG/CG	CG/CG
mh09KK-034	AA/AA	AA/AA	AA/AA	AA/AA	AA/AA	
mh09KK-153	TAC/TAA	TAC/TAA	TAC/TAA	TAC/TAA	TAC/TAA	TAC/TAA
mh11KK-039	TT/GT	TT/GT	TT/GT	TT/GT	TT/GT	TT/GG
mh12KK-043	CTG/CTG	CTG/CTG	CTG/CTG	CTG/CTG	CTG/CTG	
mh12KK-046	TA/TA	TA/TA	TA/TA	TA/TA	TA/TA	TA/TA
mh13KK-218	CCC/TTT	CCC/TTT	CCC/TTT	CCC/TTT	CCC/TTT	CCC/TTT
mh13KK-225	ACG/GAA	ACG/GAA	ACG/GAA	ACG/GAA	ACG/GAA	
mh16KK-302	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT
mh17KK-053	TC/CT	TC/CT	TC/CT	TC/CT	TC/CT	TC/CT
mh17KK-054	AA/AA	AA/AA	AA/AA	AA/AA	AA/AA	AA/AA
mh18KK-293	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA
mh21KK-320	ACG/ACG	ACG/ACG	ACG/ACG	ACG/ACG	ACG/ACG	ACG/ACA
mh21KK-324	CTG/CAG	CTG/CAG	CTG/CAG	CTG/CAG	CTG/CAG	CAG/CAG
mh22KK-060	GG/CA	GG/CA	GG/CA	GG/CA	GG/CA	

Tab.39 Genotipizzazione dei campioni di controllo a diluizioni scalari (21 cicli).

CAMPIONE	29 MH_D1-2	29 MH_D2-2	29 MH_D3-2	29 MH_D4-2	29 MH_D5-2	29 MH_D6-2
mh02KK-073	TT/GC	TT/GC	TT/GC	TT/GC	TT/GC	TT/GC
mh02KK-213	CGT/TGT	CGT/TGT	CGT/TGT	CGT/TGT	CGT/TGC	CGT/TGC
mh02KK-134	TTT/ACT	TTT/ACT	TTT/ACT	TTT/ACT	TTT/ACT	
mh02KK-136	GTC/TCC	GTC/TCC	GTC/TCC	GTC/TCC	GTC/TCC	GTC/TCA
mh03KK-007	TC/TC	TC/TC	TC/TC	TC/TC	TC/TC	TC/TC
mh03KK-009	TT/TT	TT/TT	TT/TT	TT/TT	TT/TT	TT/TC
mh04KK-010	GA/AA	GA/AA	GA/AA	GA/AA	GA/AA	GA/AG
mh04KK-011	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	GT/AT
mh04KK-074	GT/GT	GT/GT	GT/GT	GT/GT	GT/GT	GT/AT
mh04KK-013	CGG/CGG	CGG/CGG	CGG/CGG	CGG/CGG	CGG/CGG	CGA/CGG
mh05KK-170	AAG/AGA	AAG/AGA	AAG/AGA	AAG/AGA	AAG/AGA	AAA/AGA
mh05KK-062	TA/AA	TA/AA	TA/AA	TA/AA	TA/AA	TA/AA
mh05KK-022	CC/TC	CC/TC	CC/TC	CC/TC	CC/TC	CC/TC
mh06KK-026	GTG/GCA	GTG/GCA	GTG/GCA	GTG/GCA	GTG/GCA	GTG/GCA
mh07KK-031	CG/CG	CG/CG	CG/CG	CG/CG	CA/CG	CA/CG
mh09KK-034	AA/AA	AA/AA	AA/AA	AA/AA	AA/AA	AA/AA
mh09KK-153	TAC/TAA	TAC/TAA	TAC/TAA	TAC/TAA	TAA/TAA	TAA/TAA
mh11KK-039	TT/GT	TT/GT	TT/GT	TT/GT	TT/GT	TT/GT
mh12KK-043	CTG/CTG	CTG/CTG	CTG/CTG	CTG/CTG	CTG/CTG	CTG/TCG
mh12KK-046	TA/TA	TA/TA	TA/TA	TA/TA	TA/TA	TA/TA
mh13KK-218	CCC/TTT	CCC/TTT	CCC/TTT	CCC/TTT	CCC/TTT	CCC/TTT
mh13KK-225	ACG/GAA	ACG/GAA	ACG/GAA	ACG/GAA	ACG/GAA	ACG/ACG
mh16KK-302	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT
mh17KK-053	TC/CT	TC/CT	TC/CT	TC/CT	TC/CT	CT/CT
mh17KK-054	AA/AA	AA/AA	AA/AA	AA/AA	AA/AA	AA/AG
mh18KK-293	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA	AGAA/AGAA	GGAA/AGAA
mh21KK-320	ACG/ACG	ACG/ACG	ACG/ACG	ACG/ACG	ACG/ACG	ACG/ACA
mh21KK-324	CTG/CAG	CTG/CAG	CTG/CAG	CTG/CAG	CTG/CAG	CTG/CAG
mh22KK-060	GG/CA	GG/CA	GG/CA	GG/CA	GG/CA	CG/CG

Tab.40 Genotipizzazione dei campioni di controllo a diluizioni scalari (25 cicli).

I risultati ottenuti per i campioni D1, D2, D3 e D4 sono concordi sia a 21 che 25 cicli di amplificazione, non vale lo stesso per D5 e D6.

Nel campione D5 sono stati utilizzati 50 pg di DNA nell'amplificazione e nella tipizzazione del campione amplificato a 21 cicli (D5-1) risultano essere presenti 2 microplotipi in cui si osserva *drop-in* allelico. Per tutti gli altri

sistemi la genotipizzazione coincide con i campioni a maggiore concentrazione di DNA iniziale (*Tab.39*). Per quanto riguarda lo stesso campione amplificato a 25 cicli (D5-2), sono state riscontrate 3 discordanze: 2 *drop-in* e 1 *drop-out* allelici. Le discordanze ottenute non ricadono sugli stessi loci (*Tab.40*).

Nel campione D6, in cui sono stati utilizzati soli 25 pg di DNA all'amplificazione, risultano esserci diversi errori di tipizzazione. A 21 cicli di amplificazione, il campione (D6-1) presenta 9 *drop-in* e 1 *drop-out* allelico. Inoltre si osservano 3 sistemi in cui non è avvenuta la chiamata allelica (*no call*) e un sistema in cui il *coverage* medio è <50. In questo caso l'assegnazione allelica non è affidabile, quindi non viene considerata. La tipizzazione ottenuta è comunque concorde con i controlli (*Tab.39*). Per quanto riguarda lo stesso campione a 25 cicli di amplificazione (D6-2) sono osservate molte discrepanze con il genotipo di riferimento. In particolare sono presenti 14 sistemi in cui si presenta un *drop-in* allelico, 3 sistemi in cui si ha *drop-out* e 1 sistema in cui non è avvenuta la chiamata allelica (*no call*) (*Tab.40*).

Il sequenziamento dei campioni degradati artificialmente hanno ottenuto i seguenti valori di *coverage* (*Tab.35C*).

CAMPIONE	BARCODE	MAPPED READS	ON TARGET	MEAN DEPTH	UNIFORMITY
29MH_TS22-1	IonXpress_023	34075	24.36%	277.1	89.50%
29MH_TS26-1	IonXpress_024	133285	99.11%	4520	92.99%
29MH_TS29-1	IonXpress_025	91080	54.79%	1695	100.00%
29MH_TS22-2	IonXpress_032	3576	78.02%	94.56	82.38%
29MH_TS26-2	IonXpress_033	90441	99.54%	3088	92.84%
29MH_TS29-2	IonXpress_034	84141	82.58%	2317	92.99%

Tab.35C Risultati del coverage del sequenziamento dei campioni di DNA degradati artificialmente.

I valori di *uniformity* risultano essere anche in questo caso abbastanza elevati. Sono il *coverage* medio e le *reads on target* che presentano valori discordanti. Infatti le *reads on target* del campione 29MH_TS22-1 è di appena 24,36% e anche nel caso del 29MH_TS29-2 poco più della metà delle *reads* sono *on target*.

Invece i valori medi del *coverage* risultano essere molto bassi per il campione 29MH_TS22, sia a 21 che a 25 cicli di amplificazione. È necessario sottolineare che questi valori sono bassi poiché sono stati riscontrati rispettivamente 2 e 7 locus con *coverage* inferiore a 50 (nei campioni 29MH_TS22-1 e 29MH_TS22-2). Questi valori vengono evidenziati in *Tab.36A* e *Tab.36B*. In entrambi i campioni 29MH_TS22 (1 e 2) vengono rilevati i valori di *uniformity* più bassi (rispettivamente 89,50% e 82,38%).

Riportiamo ora la genotipizzazione dei campioni degradati artificialmente amplificati a 21 e 25 cicli (*Tab.41* e *Tab.42*).

CAMPIONE	29MH_TS22-1	29MH_TS26-1	29MH_TS29-1
mh02KK-073	GC/GC	TT/GT	TT/GT
mh02KK-213	CGT/TGC	TGC/TGC	TGC/TGC
mh02KK-134		TCT/ACT	TCT/ACT
mh02KK-136	TTC/TCA	GTA/TTC	GTA/TTC
mh03KK-007	TT/TC	TT/TT	TT/TT
mh03KK-009	CT/CC	CC/TT	CC/TT
mh04KK-010	AG/GA	GA/GG	GA/GG
mh04KK-011	GT/AT	AT/AT	AT/AT
mh04KK-074	AT/AT	AT/AT	AT/AT
mh04KK-013	CGA/CGA	CGA/CGG	CGA/CGG
mh05KK-170	AGA/AGA	AGA/GAA	AGA/GAA
mh05KK-062	TA/AA	TA/AA	TA/AA
mh05KK-022	CC/TC	CC/TC	CC/TC
mh06KK-026	GCA/GCG	GTG/GCG	GTG/GCG
mh07KK-031	CG/CA	CG/CG	CG/CG
mh09KK-034		AA/AA	AA/AA
mh09KK-153	TAA/TAA	TAA/TAA	TAA/TAA
mh11KK-039	GG/GG	GT/GG	GT/GG
mh12KK-043	TCG/CTG	CTA/CTG	CTA/CTG
mh12KK-046	TA/TA	TA/TA	TA/TA
mh13KK-218	CTT/TTC	CTC/CTT	CTC/CTT
mh13KK-225	ACG/GAG	GCG/GAG	GCG/GAG
mh16KK-302	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT
mh17KK-053	CT/CT	CT/CT	CT/CT
mh17KK-054	AA/AG	AA/AA	AA/AA
mh18KK-293		GGAA/GGAA	GGAA/GGAA
mh21KK-320	ACA/ACG	ACA/ACG	ACA/ACG
mh21KK-324	TTA/TAG	CTG/TTA	CTG/TTA
mh22KK-060	CA/CG	CA/CA	CA/CA

Tab.41 Genotipizzazione dei campioni degradati artificialmente (21 cicli).

CAMPIONE	29MH_TS22-2	29MH_TS26-2	29MH_TS29-2
mh02KK-073		TT/GT	TT/GT
mh02KK-213	TGC/TGC	TGC/TGC	TGC/TGC
mh02KK-134		TCT/ACT	TCT/ACT
mh02KK-136	TTC/TCA	GTA/TTC	GTA/TTC
mh03KK-007		TT/TT	TT/TT
mh03KK-009	CC/TT	CC/TT	CC/TT
mh04KK-010	AG/GG	GA/GG	GA/GG
mh04KK-011	GT/AT	AT/AT	AT/AT
mh04KK-074		AT/AT	AT/AT
mh04KK-013	CGA/CGG	CGA/CGG	CGA/CGG
mh05KK-170	AGA/AGA	AGA/GAA	AGA/GAA
mh05KK-062	TA/TA	TA/AA	TA/AA
mh05KK-022	TC/TC	CC/TC	CC/TC
mh06KK-026	TRI	GTG/GCG	GTG/GCG
mh07KK-031	CG/CA	CG/CG	CG/CG
mh09KK-034		AA/AA	AA/AA
mh09KK-153		TAA/TAA	TAA/TAA
mh11KK-039	GG/GG	GT/GG	GT/GG
mh12KK-043	TCG/CTG	CTA/CTG	CTA/CTG
mh12KK-046	TA/TA	TA/TA	TA/TA
mh13KK-218	CTC/TTC	CTC/CTT	CTC/CTT
mh13KK-225	ACT/GAC	GCG/GAG	GCG/GAG
mh16KK-302	ACTT/ACTT	ACTT/ACTT	ACTT/ACTT
mh17KK-053	CT/CT	CT/CT	CT/CT
mh17KK-054	AA/AG	AA/AA	AA/AA
mh18KK-293		GGAA/GGAA	GGAA/GGAA
mh21KK-320	ACA/ACG	ACA/ACG	ACA/ACG
mh21KK-324	CAG/CAG	CTG/TTA	CTG/TTA
mh22KK-060		CA/CA	CA/CA

Tab.42 Genotipizzazione dei campioni degradati artificialmente (25 cicli).

I risultati ottenuti per i campioni di DNA degradato artificialmente vanno considerati separatamente. In primis andremo a valutare il 29MH_TS26 che rappresenta il campione di DNA di riferimento. In questo caso sono stati utilizzati 5 ng di DNA iniziale. La quantità è la stessa utilizzata nella diluizione D1. La genotipizzazione è risultata concorde sia a 21 che a 25 cicli.

Il campione 29MH_TS29 ha ottenuto buoni risultati di sequenziamento. Infatti la tipizzazione è concorde con quella ottenuta dal campione di riferimento 29MH_TS26.

Per quanto riguarda invece il campione 29MH_TS22 ci sono degli errori nel sequenziamento, sia nel campione amplificato a 21 e a 25 cicli di amplificazione. A 21 cicli di amplificazione vengono rilevati 15 sistemi in cui è avvenuto *drop-in* allelico, 2 sistemi con *coverage* inferiore a 50 e un *no call*. Nel campione amplificato a 25 cicli sono stati invece rilevati 10 *drop-in* allelici, 3 *drop-out* allelici, 8 sistemi non considerabili (7 con *coverage* inferiore a 50 e 1 *no call*).

Dopodiché sono stati testati i *casework sample* (campioni reali forensi). I risultati del sequenziamento sono di seguito riportati.

CAMPIONE	BARCODE	MAPPED READS	ON TARGET	MEAN DEPTH	UNIFORMITY
29MH_BO04	IonXpress_035	167919	98.80%	5530	96.33%
29MH_BO05	IonXpress_036	118382	91.91%	3470	92.99%
29MH_TS04	IonXpress_037	62482	24.98%	528.6	92.99%
29MH_TS07	IonXpress_038	147656	98.87%	4946	96.33%

Tab.35D Risultati del coverage del sequenziamento dei casework sample.

Dalla *Tab.35D* è chiaro che tutti i campioni hanno ottenuto buoni risultati, eccetto il 29MH_TS04. In questo caso il numero di *reads* e il valore di *reads on target* sono molto bassi (rispettivamente 62482 e 24,98%); anche il

coverage medio risulta decisamente inferiore agli altri, circa un decimo degli altri valori ottenuti. In *Tab.43* sono riportati i valori della genotipizzazione dei *casework sample*. In questo caso non si ha un controllo del genotipo.

CAMPIONE	29MH_BO04	29MH_BO05	29MH_TS04	29MH_TS07
mh02KK-073	GT/GT	GT/GC		GT/GT
mh02KK-213	CGT/TGC	CGT/TGC	CGT/TGT	TGT/TGC
mh02KK-134	TCT/ACT	TCT/ACT	ACT/ATC	ATC/ACC
mh02KK-136	GTC/TTC	GTC/TTC		GTA/CTA
mh03KK-007	CC/TC	CC/TC	TC/TT	TC/TT
mh03KK-009	TC/TT	TC/TT	TC/TT	TT/TT
mh04KK-010	AG/GA	AG/GA		AG/GA
mh04KK-011	AC/AC	AC/AC	GT/AT	GT/AC
mh04KK-074	AT/GT	AT/GT	AT/AC	GT/AT
mh04KK-013	AAG/CGG	AAG/CGA	CGA/CGG	CGG/CGG
mh05KK-170	AAG/AAG	AAG/AAG	AGA/AAA	AAG/AGG
mh05KK-062	AA/AA	AA/AA	TA/TA	AA/AA
mh05KK-022	CA/CC	CA/CC		CA/CC
mh06KK-026	GTG/GCG	GTG/GCG	GCG/GCG	GCG/GCG
mh07KK-031	CG/CA	CG/CA	CA/CG	CA/CA
mh09KK-034	GG/GA		AG/GA	GA/GA
mh09KK-153	CAA/TAA	CAA/TAA	TAA/TAA	TAA/TAC
mh11KK-039	GG/GT	GG/GT	GG/GT	TT/GT
mh12KK-043	TCG/CTG	TCG/CTG	TCG/CTA	CTG/CTG
mh12KK-046	TG/TA	TG/TA	TA/TA	TA/TG
mh13KK-218	CTT/TTT	CTT/TTT		TTT/TTT
mh13KK-225	GCG/GAG	GCG/GAG	ACG/GAA	GAA/GAG
mh16KK-302	ACTT/GCTT	ACTT/GCTT	ACTT/ACTT	GTAT/GTTT
mh17KK-053	TC/CT	TC/CT	TT/CT	CT/CT
mh17KK-054	AA/AG	AA/AG		AA/AG
mh18KK-293	GGAA/AGAA	GGAA/AGAA	GGAA/AGAA	GGAA/AGAA
mh21KK-320	ACA/ACG	ACA/ACG	ACG/ACA	GCG/ATA
mh21KK-324	CAG/CAG	CAG/CAG	TRI	CAG/CTG
mh22KK-060	GG/CG	GG/CG	CA/CG	GG/GG

Tab.43 Genotipizzazione dei casework sample.

Nei campioni 29MH_BO04, 29MH_BO05 e 29MH_TS07 i genotipi ottenuti non hanno dato particolari problemi di chiamate alleliche.

Il problema sorge invece con il campione 29MH_TS04. Dall'analisi con IGV sono stati analizzati dei sistemi in eterozigosi determinati da tre alleli diversi. In particolare sono stati rilevati 5 sistemi con tre alleli ciascuno e 3 sistemi che in cui non c'è stata chiamata allelica (*no call*). Si tratta di un campione molto degradato e in cui è stata utilizzata una quantità esigua di DNA.

Capitolo 5

DISCUSSIONE

L'applicazione dei microaplotipi nelle indagini forensi è risultata essere molto promettente negli ultimi anni. Numerosi studi confermano che la tipizzazione dei microaplotipi permette di aumentare notevolmente il potere discriminativo nelle analisi forensi (10). Infatti a differenza di quanto accade per gli STRs, il tasso mutazionale degli SNPs è molto inferiore e la valutazione di più SNPs in blocco permette di aumentare ancora di più il potere discriminativo.

La tecnologia innovativa MPS comprende metodi di sequenziamento del DNA ad alto rendimento o *high-throughput*. Le piattaforme in commercio elaborano milioni di *reads* in parallelo, invece di elaborare singoli ampliconi come avveniva in passato.

Recentemente è stato confermato che l'applicazione delle tecnologie MPS permette di ottenere risultati migliori soprattutto per i campioni di DNA degradato o a basse concentrazioni (64).

Infatti il presente studio si concentra proprio sulla valutazione della performance di un pannello di microaplotipi su campioni di DNA degradati

analizzati mediante tecnologia MPS. L'attenzione si è rivolta proprio sui microaplotipi poiché risultano essere *marker* genetici particolarmente informativi per campioni forensi.

L'applicazione dei microaplotipi risulta vincente per l'identificazione personale, per l'accertamento di relazioni parentali, nella deconvoluzione di campioni misti e nell'ascendenza biogeografica. Sono note applicazioni anche in genetica medica, umana e animale.

In questo studio viene dimostrata l'efficienza del pannello proposto anche su campioni molto degradati. Sono state effettuate prove su campioni di DNA di diversa tipologia e i risultati ottenuti dal sequenziamento tramite *Ion PGM™ System* hanno portato alle seguenti conclusioni.

Le prime prove sono state fatte su campioni di DNA non degradato. La genotipizzazione di questi 6 campioni ha prodotto i genotipi attesi. Infatti per questi campioni sono state confrontate le genotipizzazioni ottenute con il pannello a 29 microaplotipi e le genotipizzazioni ottenute da *Turchi et al* (24) con il pannello a 89 MH. Il confronto ha dato risultati pienamente concordanti. I valori di *coverage* ottenuti sono molto buoni ed in equilibrio tra di loro (barre verdi in *Fig.30*).

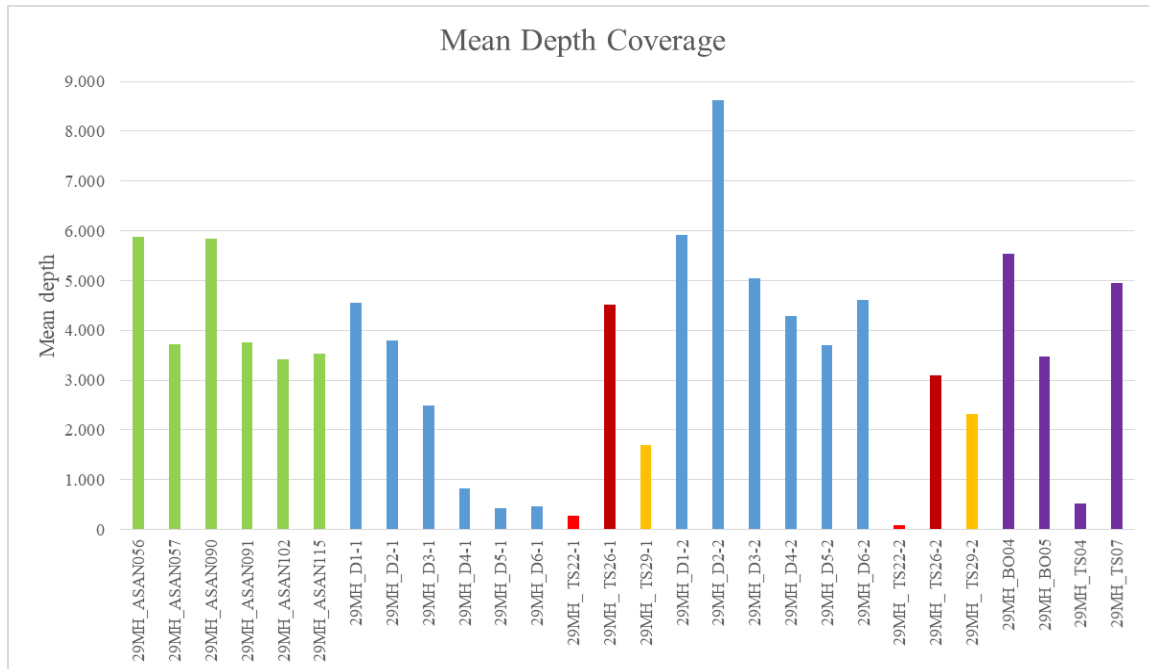


Fig.30 Istogramma dei valori di coverage di tutti i campioni.

I campioni utilizzati in queste prove presentavano condizioni ottimali per l'analisi, si trattava di campioni con buoni livelli di DNA non degradato. Questa prova è stata effettuata per valutare la riproducibilità dei risultati su microplotipi già studiati e per confermare che il disegno dei *primer* fosse corretto. Inoltre si conferma anche l'assenza di ampliconi che vengono amplificati preferenzialmente rispetto ad altri. Tutti risultano avere una buona efficienza di amplificazione. Questo ci permette quindi di affermare che il pannello a 29 microplotipi consegue ottimi risultati quando il DNA di partenza è in condizioni ottimali, di quantità e di degradazione. Il numero di cicli di amplificazione utilizzato (21 cicli) è stato sufficiente per la corretta amplificazione e preparazione di librerie.

Dato che lo scopo dello studio era quello di validare il pannello a 29 microplotipi su campioni forensi, sono stati effettuati dei test per la sensibilità, analizzando campioni a concentrazioni scalari di DNA di partenza. I risultati ottenuti per questi campioni ci permettono di affermare che la quantità del DNA di partenza è fondamentale per la corretta tipizzazione. Infatti i valori del *coverage* si abbassano con l'abbassarsi della concentrazione iniziale di DNA (barre azzurre in Fig.30). Questo fenomeno è molto più marcato nei campioni amplificati a 21 cicli. Infatti si passa da un *coverage* di 4553 per D1 (5000 pg) a un *coverage* molto più basso di 418,7 per D5 (50 pg) e di 457,9 per D6 (25 pg). I valori di *coverage* ottenuti dai campioni amplificati a 25 cicli sono migliori rispetto a quelli appena citati, ma questo non influisce sulla migliore tipizzazione.

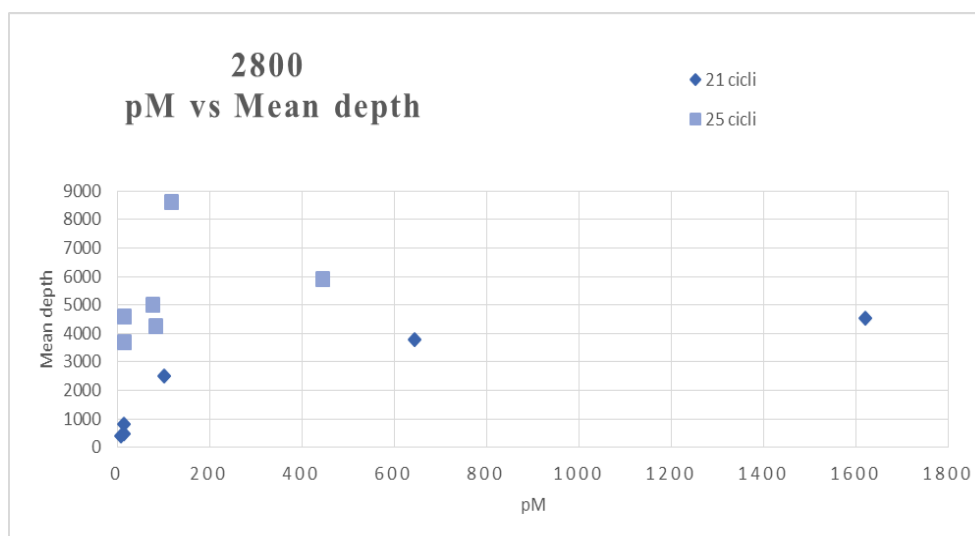


Fig.31 Relazione tra le concentrazioni in pM e il valore di coverage medio nei test per la sensibilità (a 21 e a 25 cicli).

In *Fig.31* è riportata la relazione che intercorre tra le concentrazioni in pM e il valore di *coverage* medio nei campioni per il test di sensibilità, a 21 e a 25 cicli di amplificazione. I valori di *reads on target* sono molto alti e sovrapponibili in entrambi i protocolli di amplificazione utilizzati. Di seguito è presente una rappresentazione grafica (*Fig.32*).

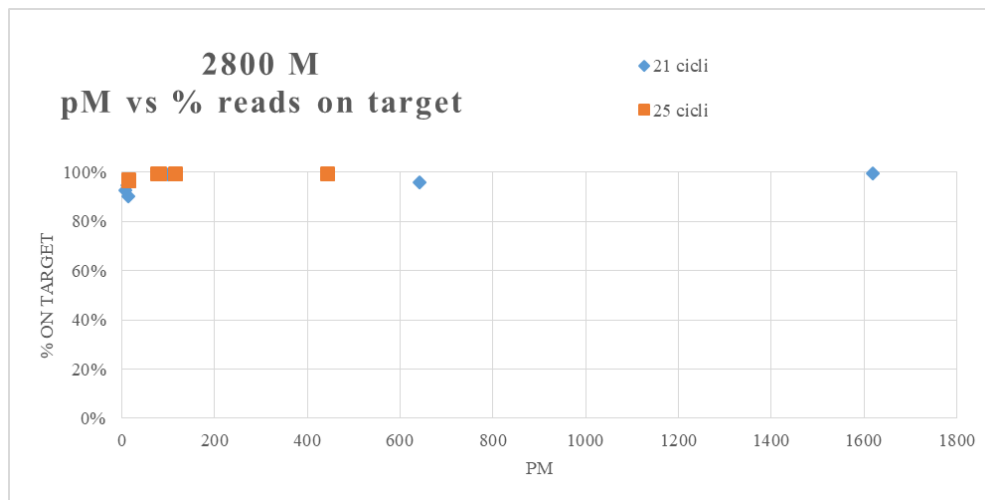


Fig.32 Relazione tra le concentrazioni in pM e il la percentuale di reads on target nei test per la sensibilità (a 21 e a 25 cicli).

È importante sottolineare che nei campioni D5-2 e D6-2 sono stati riscontrati dei sistemi con sbilanciamento degli eterozigoti. In questo caso si può affermare che si ottengono genotipi riproducibili e esatti fino a una concentrazione di 100 pg (D4), al di sotto di questa soglia i risultati della genotipizzazione non sono pienamente concordanti. Infatti nei due campioni D5 e D6 si ottengono diversi errori di tipizzazione, soprattutto *drop-in* allelici.

I risultati ottenuti a 25 cicli di amplificazione non hanno mostrato miglioramenti di genotipizzazione rispetto a quelli ottenuti a 21 cicli.

Per quanto concerne invece i campioni di DNA degradati artificialmente si hanno problemi di tipizzazione solo con 29MH_TS22 (1 e 2). Questo campione infatti è l'unico in cui è stata utilizzata una bassa concentrazione di DNA alla amplificazione (loq), inoltre il DNA risultava essere molto degradato (DI=nc).

Il campione 29MH_TS22 ottiene una genotipizzazione con molti errori, sia a 21 che a 25 cicli di amplificazione. Nel primo caso si ottengono 18 sistemi alterati (per *drop-in*, *drop-out* e *no call*) e nel secondo ne risultano alterati ben 21 sul totale di 29 sistemi considerati. Ne consegue che l'aumento del numero di cicli, soprattutto in campioni con DNA degradato, non porta miglioramenti di tipizzazione.

Nell'analisi dei risultati del campione 29MH_TS26 non sono stati riscontrati errori. Si tratta di un campione di riferimento a cui è stato assegnato senza riscontrare né problemi di interpretazione né differenze tra i due i campioni, a 21 e a 25 cicli. I livelli di *coverage* e di *reads on target* erano molto alti in entrambi i campioni.

Il 29MH_TS29, campione di sangue depurinato, presenta un alto indice di degradazione (nc). La quantità di DNA utilizzato per l'amplificazione è di

circa 1000 pg, quantità paragonabile con quella utilizzata per il campione di controllo (29MH_TS26). In questo caso si ottengono valori di *coverage* non molto alti, ma equilibrati tra loro. Questo permette comunque una buona tipizzazione. Il genotipo ottenuto è perfettamente sovrapponibile con quello di riferimento. Pertanto è possibile affermare che lo stato di degradazione del DNA non influisce sull'efficienza del pannello utilizzato per la tipizzazione.

Nel grafico in *Fig.33* sono rappresentate le correlazioni tra concentrazione molare (pM) dei campioni e il *coverage* ottenuto con il sequenziamento. I dati riguardano i campioni di DNA degradato artificialmente (■ e ◆), i campioni di DNA di controllo non degradato (✕) e i *casework sample* (▲).

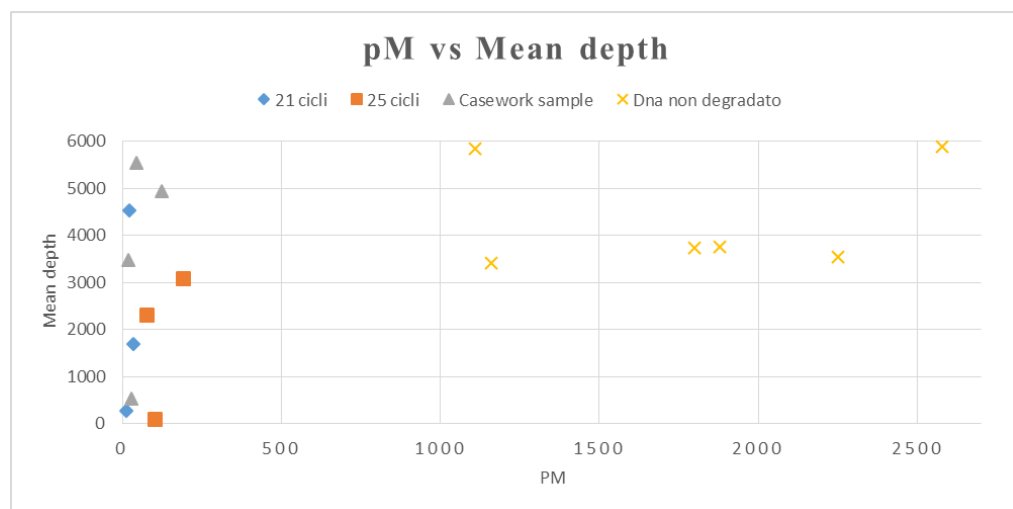


Fig.33 Relazione tra le concentrazioni in pM e il valore di *coverage* medio nei campioni di DNA degradato artificialmente (a 21 e a 25 cicli), nei *casework sample* e nei campioni di DNA non degradato.

Lo scopo dell'analisi di questi ultimi era quello di testare il pannello su campioni forensi reali con entità diverse di degradazione. I campioni analizzati risultavano essere infatti degradati e in alcuni casi anche a bassa concentrazione di DNA. Sono stati selezionati dei campioni forensi da tessuti inclusi in paraffina e tessuti ossei. Questo genere di campioni infatti sono quelli che, in genetica forense, ottengono profili parziali, determinati proprio dal grado di degradazione.

I tessuti in paraffina sono tessuti molto alterati, affinché sia possibile la buona riuscita di un preparato istologico, è necessario effettuare un passaggio di fissazione del preparato in formalina. Questo step permette di mantenere il più possibile inalterato la struttura dei tessuti. La formalina determina legami crociati tra il liquido fissativo e gruppi funzionali delle proteine, con mascheramento di molti siti antigenici. Questo consente al campione di sopportare gli stress fisici e chimici delle successive fasi di disidratazione, inclusione e sezionamento. A livello nucleare la formalina porta però anche alla denaturazione delle proteine istoniche, ne deriva una parziale degradazione del DNA.

Per quanto riguarda le genotipizzazioni dei campioni inclusi in paraffina non sono stati riscontrati particolari problemi, infatti solo un sistema analizzato ha ottenuto un errore di sequenziamento. In particolare si tratta di una mancata

chiamata allelica (*no call*), per cui è possibile che si tratti di un errore dovuto alla fase finale del processo, e non all'alterazione del campione.

L'analisi dei genotipi ottenuti dai 2 campioni di DNA da tessuto osseo ha mostrato problemi solo in un caso, probabilmente questo deriva dalle caratteristiche iniziali del campione. Si tratta del campione 29MH_TS04, il quale presenta bassa concentrazione iniziale di DNA (loq) e alto grado di degradazione (nc). La concentrazione del campione è inferiore al limite minimo di rilevazione del sistema di quantizzazione utilizzato, per cui non si ottiene un valore. Il *Quantifiler® Trio DNA Quantification* (57) è il kit utilizzato per la quantizzazione e possiede un'elevata sensibilità, determinata dall'utilizzo di *loci target* (Small Autosomal target e Large Autosomal target) a più copie. La ditta garantisce la possibilità di rilevare concentrazioni di DNA inferiori a 5pg/μL, ne deriva che la concentrazione iniziale del campione 29MH_TS04 è sicuramente inferiore a questo valore. L'indice di degradazione rilevato per questo campione è nc. Si tratta di un campione ad alto grado di degradazione, per cui non è possibile ottenere un rapporto tra la concentrazione della Small Autosomal target (SA) e la concentrazione della Large Autosomal target (LA).

La bassa concentrazione e l'alta degradazione del campione di partenza non ha impedito la genotipizzazione. Sono stati rilevate numerose incoerenze, ma

22 dei 29 sistemi studiati mostrano profili non alterati, a differenza dei valori di coverage che sono poco equilibrati (barre viola in *Fig. 30*). In particolare, in 5 sistemi è stata riscontrata eterozigosi triallelica e in 2 sistemi non c'è stata chiamata allelica (*no call*). L'elevato grado di degradazione può spiegare le discrepanze ottenute.

È infatti dimostrato che la lettura di sequenze alleliche su DNA danneggiato può comportare problemi durante la PCR, introducendo errori casuali nelle sequenze. È stato messo in evidenza (65), che errori di sequenziamento del DNA antico causati dalla presenza di nucleotidi danneggiati nella sequenza stampo sono più frequenti rispetto a quelli del DNA contemporaneo.

Oppure si può supporre che la presenza di un allele sovranumerario sia dovuto a contaminazione. Sicuramente un'ulteriore analisi sullo stesso campione può confermare la tipizzazione ottenuta.

Per quanto riguarda l'ultimo campione esaminato, il 29MH_TS07, non sono stati rilevati problemi di genotipizzazione. Si trattava di un campione parzialmente degradato (DI=5) e in cui è stata utilizzata una quantità esigua di DNA (0,689 ng/ μ L).

Ovviamente è di fondamentale importanza la valutazione di altri fattori, caso per caso. Se il campione di tessuto osseo deriva da un individuo deceduto, è importante stabilire il lasso di tempo che intercorre tra la morte e il momento

dell'analisi: maggiore è il periodo di tempo e maggiore è la probabilità di degradazione del campione. È necessario sottolineare però che l'età del campione di per sé non dà indicazioni assolute sul deterioramento del DNA. Inoltre bisogna valutare anche le condizioni di conservazione, ossia l'insieme di condizioni ambientali (temperatura, pH, umidità...) e dei processi cui è sottoposto l'organismo dopo la morte (disidratazione, essiccazione...). In ambito forense ci può trovare di fronte a resti cadaverici inumati, tumulati o rinvenuti in mare. In ogni caso ci sono variabili da considerare che influiscono sullo stato di degradazione del DNA. Il tipo di terreno e il necrobioma, ossia l'insieme di tutti i microrganismi decompositori che attaccano i tessuti, possono influire diversamente sullo stato di conservazione. Quindi il tasso di degradazione e la quantità di DNA che si degrada dipendono da numerosi processi *post-mortem*, che possono portare a livelli di degradazione più o meno marcata.

Non è possibile risalire a questo genere di parametri relativi ai campioni presi in esame nel presente studio, quindi è necessario basarsi sulle sole informazioni disponibili, ossia concentrazione e indice di degradazione.

Capitolo 6

CONCLUSIONI

L'introduzione della tecnologia *Massive Parallel Sequencing* offre nuove possibilità in genetica forense, sia per il maggior numero di informazioni che possono essere ottenute in tempi brevi, sia per la riduzione dei costi rispetto al tradizionale metodo di sequenziamento di Sanger. Inoltre questa tecnologia consente di esplorare ed analizzare *nuovi marker* genetici come i microplotipi. Per definizione un microplotipo è costituito da 2 o più SNPs geneticamente vicini all'interno di un breve segmento di DNA (<300 bp). I microplotipi sono molto informativi per scopi forensi quali identificazione personale, identificazione di parentele, deconvoluzione di miscele e determinazione dell'ascendenza biogeografica. Inoltre questi loci sono potenzialmente adatti per l'analisi di campioni di DNA degradati.

Nel presente studio di validazione sono stati analizzati 14 campioni differenti per matrice, quantità di DNA e indice di degradazione. Nonostante le diverse condizioni è stato possibile testare il pannello su tutti i campioni, dimostrando la sua versatilità di applicazione.

L'incremento del numero di cicli di PCR non sembra fornire un miglioramento nei risultati di tipizzazione nei campioni degradati con basse quantità di DNA, come accade invece per i campioni di DNA di controllo, tuttavia si riscontrano un maggior numero di eventi di *drop-in* seguendo il protocollo a 25 cicli di amplificazione.

Concludendo l'analisi permette di affermare che il pannello a 29 microplotipi produce risultati di tipizzazione buoni e riproducibili anche se si hanno alti livelli di degradazione; si ottengono artefatti nella genotipizzazione solo quando si parte da una quantità molto limitata di DNA iniziale.

Capitolo 7

BIBLIOGRAFIA

1. Ceppellini R., Curtoni E.S., Mattiuz P., Miggiano V., Scudeller G., Serra A., Tosi R.M. (Eds.), *Genetics of leukocyte antigens: A family study of segregation and linkage*, *Histocompatibility Testing*, Munksgaard, Copenhagen, 1967, pp. 149–187;
2. *International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome*, *Nature*. 2004 Oct; 431(7011) 931-45;
3. Tagliabracci A., Alessandrini F., Mazzarini L., Onofri V., Onori N., Turchi C., *Introduzione alla genetica forense. Indagini di identificazione personale e di paternità*, Ancona, Springer Biomed. 2010;
4. <https://www.fbi.gov/services/laboratory/biometric-analysis/codis#CODIS-Overview>;
5. Budowle B., Moretti T.R., Baumstark A.L., Defenbaugh D.A., Keys K.M., *Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians,*

- Jamaicans, and Trinidadians, Journal of Forensic Sciences.* 1999 Nov; 44(6) 1277-86;
6. Butler J.M., *Advanced Topics in Forensic DNA Typing: Methodology*, Gaithersburg (USA), Academic Press, 2011;
7. Barnholtz-Sloan J.S., Chakraborty R., Sellers T.A., Schwartz A.G., *Examining population stratification via individual ancestry estimates versus self-reported race. Cancer Epidemiology, Biomarkers & Prevention.* 2005 Jun;14(6) 1545-51;
8. Walsh S., Lindenbergh A., Zuniga S.B., Sijen T., de Knijff P., Kayser M., Ballantyne K.N., *Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence, Forensic Science International: Genetics.* 2011 Nov; 5(5) 464-71;
9. <https://hirisplex.erasmusmc.nl/>;
10. Oldoni F., Kidd K.K., Podini D., *Microhaplotypes in forensic genetics, Forensic Science International: Genetics.* 38 (2019) 54-69;
11. Kidd K.K., Pakstis A.J., Speed W.C., Lagacé R., Chang J., Wootton S., Haigh E., Kidd J.R., *Current sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics, Forensic Science International: Genetics.* 12 (2014) 215–224;

12. Kidd K.K., Pakstis A.J., Speed W.C., Lagacé R., Chang J., Wootton S., Ihuegbu N., *Microhaplotype loci are a powerful new type of forensic marker, Forensic Science International: Genetics - Supplement Series.* 4 (2013) e123-e124;
13. Daly M.J., Rioux J.D., Schaffner S.F., Hudson T.J., Lander E.S., *High-resolution haplotype structure in the human genome, Nature Genetics.* 29 (2001) 229–232;
14. Patil N., Berno A.J., Hinds D.A., Barrett W.A., Doshi J.M., Hacker C.R., Kautzer C.R., Lee D.H., Marjoribanks C., McDonough D.P., Nguyen B.T., Norris M.C., Sheehan J.B., Shen N., Stern D., Stokowski R.P., Thomas D.J., Trulson M.O., Vyas K.R., Frazer K.A., Fodor S.P., Cox D.R., *Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21, Science.* 294 (2001) 1719–1723;
15. Stephens J.C., Schneider J.A., Tanguay D.A., Choi J., Acharya T., Stanley S.E., Jiang R., Messer C.J., Chew A., Han J.H., Duan J., Carr J.L., Lee M.S., Koshy B., Kumar A.M., Zhang G., Newell W.R., Windemuth A., Xu C., Kalbfleisch T.S., Shaner S.L., Arnold K., Schulz V., Drysdale C.M., Nandabalan K., Judson R.S., Rúaño G., Vovis G.F., *Haplotype variation and linkage disequilibrium in 313 human genes, Science.* 293 (2001) 489–493;

- 16.Kidd K.K., *Proposed nomenclature for microhaplotypes*, *Human Genomics*. 10 (2016)16;
- 17.<https://www.genenames.org/> HUGO Gene Nomenclature Committee;
- 18.Rajeevan H., Soundararajan U., Kidd J.R., Pakstis A.J., Kidd K.K., *ALFRED: an allele frequency resource for research and teaching*, *Nucleic Acids Research*. 2012 Jan; 40(Database issue): D1010–D1015;
- 19.<http://alfred.med.yale.edu>;
- 20.Butler J.M., *Fundamentals of Forensic DNA Typing*, Gaithersburg (USA), Accademic Press, 2009;
- 21.Kidd K.K., Speed W.C., Pakstis A.J., Furtado M.R., Fang R., Madbouly A., Maiers M., Middha M., Friedlaender F.R., Kidd J.R., *Progress toward an efficient panel of SNPs for ancestry inference*, *Forensic Science International: Genetics*. 10 (2014) 23–32;
22. Kayser M., *Forensic DNA Phenotyping: predicting human appearance from crime scene material for investigative purposes*, *Forensic Science International: Genetics*. 18 (2015) 33–48;
- 23.Kidd K.K., Speed W.C., *Criteria for selecting microhaplotypes: mixture detection and deconvolution*, *Investigative Genetics*. 6 (2015) 1;
- 24.Turchi C., Melchionda F., Pesaresi M., Tagliabracci A., *Evaluation of a microhaplotypes panel for forensic genetics using massive parallel*

- sequencing technology, Forensic Science International: Genetics.* 41 (2019) 120-127;
25. Rosenberg N.A., Li L.M., Ward R., Pritchard J.K., *Informativeness of genetic markers for inference of ancestry, The American Journal of Human Genetics.* 2003 Dec; 73(6)1402-22;
26. Bulbul O., Pakstis A.J., Soundararajan U., Gurkan C., Brissenden J.E., Roscoe J.M., Evsanaa B., Togtokh A., Paschou P., Grigorenko E.L., Gurwitz D., Wootton S., Lagace R., Chang J., Speed W.C., Kidd K.K., *Ancestry inference of 96 population samples using microhaplotypes, International Journal of Legal Medicine.* 132 (2017) 703–711;
27. Bennett L., Oldoni F., Long K., Cisana S., Madella K., Wootton S., Chang J., Hasegawa R., Lagacé R., Kidd K.K., Podini D., *Mixture deconvolution by massively parallel sequencing of microhaplotypes, International Journal of Legal Medicine.* 2019 May; 133(3) 719-729;
28. Bose N., Carlberg K., Sensabaugh G., Erlich H., Calloway C., *Target capture enrichment of nuclear SNP markers for massively parallel sequencing of degraded and mixed samples, Forensic Science International: Genetics.* 34 (2018) 186–196;

29. Weir B.S., Anderson A.D., Hepler A.B., *Genetic relatedness analysis: modern data and new challenges*, *Nature Reviews Genetics*. 7 (2006) 771–780;
30. Kidd K.K., Speed W.C., Pakstis A.J., Podini D.S., Lagacé R., Chang J., Wootton S., Haigh E., Soundararajan U., *Evaluating 130 microhaplotypes across a global set of 83 populations*, *Forensic Science International: Genetics*. 29 (2017) 29–37;
31. Oldoni F., Hart R., Long K., Maddela K., Cisana S., Schanfield M., Wootton S., Chang J., Lagace R., Hasegawa R., Kidd K., Podini D., *Microhaplotypes for ancestry prediction*, *Forensic Science International: Genetics - Supplement Series*. 6 (2017) 513–515;
32. Chen P., Yin C., Li Z., Pu Y., Yu Y., Zhao P., Chen D., Liang W., Zhang L., Chen F., *Evaluation of the Microhaplotypes panel for DNA mixture analyses*, *Forensic Science International: Genetics*. 35 (2018) 149–155;
33. Voskoboinik L., Motro U., Darvasi A., *Facilitating complex DNA mixture interpretation by sequencing highly polymorphic haplotypes*, *Forensic Science International: Genetics*. 35 (2018) 136–140;
34. Phillips C., *Forensic genetic analysis of bio-geographical ancestry*, *Forensic Science International: Genetics*. 18 (2015) 49–65;

35. Bulbul O., Cherni L., Khodjet-el-khil H., Rajeevan H., Kidd K.K., *Evaluating a subset of ancestry informative SNPs for discriminating among Southwest Asian and circum-Mediterranean populations, Forensic Science International: Genetics.* 23 (2016) 153–158;
36. Kosoy R., Nassir R., Tian C., White P.A., Butler L.M., Silva G., Kittles R., Alarcon-Riquelme M.E., Gregersen P.K., Belmont J.W., De La Vega F.M., Seldin M.F., *Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America, Human Mutation.* 30 (2009) 69–78;
37. Gettings K.B., Lai R., Johnson J.L., Peck M.A., Hart J.A., Gordish-Dressman H., Schanfield M.S., Podini D.S., *A 50-SNP assay for biogeographic ancestry and phenotype prediction in the U.S. population, Forensic Science International: Genetics.* 8 (2014) 101–108;
38. Li C.X., Pakstis A.J., Jiang L., Wei Y.L., Sun Q.F., Wu H., Bulbul O., Wang P., Kang L.L., Kidd J.R., Kidd K.K., *A panel of 74 AISNPs: improved ancestry inference within Eastern Asia, Forensic Science International: Genetics.* 23 (2016) 101–110;
39. Thiede C., Florek M., Bornhäuser M., Ritter M., Mohr B., Brendel C., Ehninger G., Neubauer A., *Rapid quantification of mixed chimerism*

- using multiplex amplification of short tandem repeat markers and fluorescence detection, Bone Marrow Transplantation. 23 (1999) 1055;*
40. Schichman S.A., Suess P., Vertino A.M., Gray P.S., *Comparison of short tandem repeat and variable number tandem repeat genetic markers for quantitative determination of allogeneic bone marrow transplant engraftment, Bone Marrow Transplantation. 29 (2002) 243;*
41. Thiede C., Bornhäuser M., Oelschlägel U., Brendel C., Leo R., Daxberger H., Mohr B., Florek M., Kroschinsky F., Geissler G., Naumann R., Ritter M., G. Prange- Krex, T. Lion, A. Neubauer, G. Ehninger, *Sequential monitoring of chimerism and detection of minimal residual disease after allogeneic blood stem cell transplantation (BSCT) using multiplex PCR amplification of short tandem repeat-markers, Leukemia Nature. 15 (2001) 293;*
42. Ning Q., Yifan X., Haiyan L., Hao L., Shaobin L., Erwen H., Jun G., Fang C., Yanwei S., Xueling O., *Non invasive prenatal paternity testing using targeted massively parallel sequencing, Transfusion. 58 (2018) 1792–1799;*

43. Alaeddini R., Walsh S.J., Abbas A., *Forensic implications of genetic analyses from degraded DNA - a review*, *Forensic Science International: Genetics*. 2010 Apr; 4(3) 148-57;
44. Antheunisse J., *Decomposition of nucleic acids and some of their degradation products by microorganisms*, *Antonie Van Leeuwenhoek*. 1972, 38(3) 311-27;
45. Paabo S., Poinar H., Serre D., Jaenicke-Despres V., Hebler J., Rohland N., Kuch M., Krause J., Vigilant L., Hofreiter M., *Genetic analyses from ancient DNA*, *Annual Review of Genetics*. 38(2004) 645–679;
46. Bessetti J., *An introduction to PCR inhibitors*, *Profiles in DNA*. 2007, 10(1) 9–10;
47. Wilson I.G., *Inhibition and facilitation of nucleic acid amplification*. *Applied and Environmental Microbiology*. 63(1997) 3741–3751;
48. Akane A1, Matsubara K, Nakamura H, Takahashi S, Kimura K., *Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of polymerase chain reaction*, *Journal of Forensic Sciences*. 39(1994) 362–372;
49. Rådström P., Knutsson R., Wolffs P., Lövenklev M., Löfström C., *Pre-PCR processing: strategies to generate PCR-compatible samples*. *Molecular Biotechnology*. 2004 Feb; 26(2) 133-46;

50. <https://www.genome.gov> / National Human Genome Research Institute
51. Metzker M. L., *Sequencing technologies — the next generation. Nature Reviews Genetics*, 11(2010), 31–46;
52. Progetto “NGS Core Facility” – Azienda Ospedaliera, Ospedale Infantile Regina Margherita – S. Anna di Torino – Dipartimento di diagnostica e servizi
53. <https://www.ampliseq.com>
54. QIAamp® DNA Micro Handbook, Third Edition, December 2014, *QIAGEN®*;
55. QIAamp® DNA Investigator Handbook, June 2012, *QIAGEN®*;
56. Fattorini P., Previderè C., Sorçaburu-Cigliero S., Marrubini G., Alù M., Barbaro A.M., Carnevali E., Carracedo A., Casarino L., Consoloni L., Corato S., Domenici R., Fabbri M., Giardina E., Grignani P., Lonero Baldassarra S., Moratti M., Nicolin V., Pelotti S., Piccinini A., Pitacco P., Plizza L., Resta N., Ricci U., Robino C., Salvaderi L., Scarnicci F., Schneider P.M., Seidita G., Trizzino L., Turchi C., Turrina S., Vatta P., Vecchiotti C., Verzeletti A., De Stefano F., *The molecular characterization of a depurinated trial DNA sample can be a model to understand the reliability of the results in forensic genetics, Electrophoresis*. 35 (2014) 3134–3144;

57. Quantifiler® Trio DNA Quantification Kit, User Guide, *Applied Biosystem*, Rev.E;
58. E-Gel™ Agarose Gels with SYBR Safe™, *Invitrogen, ThermoFisher Scientific* Rev. A.0;
59. Agilent 2200 TapeStation System, User Manual, *Agilent Technologies*, Rev. C;
60. Precision ID Panels with the Ion PGM™ System, Application Guide, *Applied Biosystem*, Rev. B.0;
61. Ion PGM™ Hi-Q™ View OT2 Kit, User Guide, *Ion Torrent, ThermoFisher Scientific*, Rev. C.0;
62. Ion PGM™ Hi-Q™ View Sequencing Kit, User Guide, *Ion Torrent, ThermoFisher Scientific*, Rev. C.0;
63. <https://igv.org/> Integrative Genomics Viewer;
64. Bruijns B., Tiggelaar R., Gardeniers H., *Massively parallel sequencing techniques for forensics: A review. Electrophoresis*. 2018 (39) 2642-2654;
65. Brown, T.A., Brown, K.A., *Ancient DNA: using molecular biology to explore the past. Bioessays*, 16 (1994),719-726.

Capitolo 8

RINGRAZIAMENTI

Mi sento di dover ringraziare diverse persone che hanno permesso la realizzazione di questo elaborato.

In primo luogo, voglio ringraziare la mia relattrice, la dottoressa Chiara Turchi. Grazie a lei sono entrata nel mondo della medicina legale e in particolare della genetica forense. La sua disponibilità e il suo esempio sono stati fondamentali per la stesura di tutto il lavoro.

Desidero inoltre ringraziare la mia correlatrice, la dottoressa Filomena Melchionda. La sua infinita pazienza e i suoi indispensabili consigli mi hanno aiutato nelle attività di laboratorio, e non solo.

Ringrazio il professor Adriano Tagliabracci per aver permesso lo svolgimento della mia tesi nel laboratorio di genetica forense del Dipartimento di Scienze Biomediche e Sanità Pubblica, così attrezzato e proiettato verso il futuro. Grazie a Valerio Onofri per essere stato un riferimento professionale. Grazie a Federica Alessandrini per le sue preziose indicazioni.

Grazie a Beatrice. Senza il suo supporto psicologico, al suo aiuto nelle attività di laboratorio e al suo incoraggiamento nei momenti di sconforto non avrei raggiunto questo risultato.

Grazie a Lucia. Amica che mi ha accompagnato in tutto il percorso universitario. Negli alti e bassi che abbiamo vissuto, siamo state sempre vicine. Grazie a lei e a Beatrice ho capito realmente cosa significa lavorare in team e sostenersi a vicenda.

Grazie a Lorenzo. Grazie alla sua vicinanza in ogni momento. All'aiuto fisico che mi ha dato nelle serate passare a ripetere gli argomenti d'esame e all'aiuto psicologico fornito fin dall'inizio. Unico in grado di calmarmi quando l'ansia prendeva il sopravvento, sei stato il tassello indispensabile per la realizzazione di questo enorme puzzle!

Grazie alla mia famiglia per il sostegno dimostratomi. Soprattutto ringrazio mia madre per esserci sempre stata, in tutti i momenti di gioia e in quelli di delusione.

Ringrazio la mia amica Giada. Grazie ai suoi consigli e alla sua vicinanza non è mancato mai un appoggio durante tutto il mio percorso di studi.

Infine ringrazio tutti i miei amici per avermi sempre incoraggiata e sostenuta nelle mie scelte. Non scorderò mai i vostri visi soddisfatti per i miei traguardi!