



UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie “L. Spallanzani”

Laurea Magistrale in Biologia Sperimentale e Applicata

Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense
Unità di Medicina Legale e Scienze Forensi “A. Fornari”

Ottimizzazione di una metodica di estrazione del DNA da tessuti autoptici inclusi in paraffina (FFPE) in ambito forense, utilizzando un approccio di design of experiments (DoE)

Relatore:

Prof. Carlo Previderè

Correlatore:

Dott.ssa Pierangela Grignani

Tesi Sperimentale di
Riccardo Ciapparelli

Anno Accademico 2025/2026

Sommario

1 INTRODUZIONE	1
1.1 La Genetica forense.....	2
1.2 Variabilità genetica e polimorfismi del DNA	4
1.2.1 I polimorfismi di sequenza: <i>SNPs</i>	7
1.3 Il profilo genetico forense	16
1.4 La degradazione del DNA	22
1.4.1 Il DNA come molecola chimicamente instabile: depurinazione, deaminazione e rottura del legame fosfodiesterico.....	22
1.4.3 La degradazione enzimatica.....	25
1.5 Tessuti inclusi in paraffina (<i>FFPE</i>).....	27
1.5.1 L'effetto dei processi di fissazione e inclusione sulla molecola di DNA....	28
2 SCOPO DEL LAVORO	34
3 MATERIALI E METODI.....	36
3.1 Descrizione dei campioni	37
3.2 Progettazione degli esperimenti (<i>Design of experiments, DoE</i>) per lo studio del metodo di estrazione di DNA su tessuti <i>FFPE</i>	38
3.2.1 Obiettivi sperimentali e prove studiate	38
3.2.2 Fattori di processo e loro dominio sperimentale.....	39
3.2.3 Screening dei fattori con tessuto cardiaco	40
3.2.4 Estensione delle prove di estrazione a più tessuti.....	42
3.3 Estrazione del DNA	43
3.3.1 Estrazione manuale da tessuti inclusi in paraffina	43
3.3.2 Estrazione automatizzata da tessuti inclusi in paraffina	45
3.4 Quantificazione del DNA.....	48
3.4.1 Quantificazione tramite <i>Real Time PCR</i>	49
3.5 Amplificazione tramite <i>PCR multiplex</i>	56
3.5.1 Amplificazione con kit <i>GlobalFiler™ PCR</i>	58
3.6 Elettroforesi capillare	60
3.6.1 Preparazione della miscela di reazione per l'elettroforesi capillare	64
4 RISULTATI E DISCUSSIONE	67

4.1 Quantificazione del DNA.....	68
4.1.1 Quantificazione del campione autoptico FFPE 1	68
4.1.2 Quantificazione del campione autoptico FFPE 2	72
4.1.3 Quantificazione del campione autoptico FFPE 3	77
4.2 Analisi “Effetto tessuto”.....	81
4.3 Miscela di estrazione finale ottimizzata.....	84
4.4 Confronto tra estrazione con la miscela ottimizzata ed estrazione automatizzata con kit.....	87
4.5 Valutazione della qualità dei profili genetici	92
5 CONCLUSIONI	105
6 BIBLIOGRAFIA E SITOGRAFIA.....	109

1 INTRODUZIONE

1.1 La Genetica forense

La genetica forense rappresenta una delle discipline più rivoluzionarie nell'ambito delle scienze applicate alla Giustizia, costituendo oggi uno strumento imprescindibile per l'identificazione personale e la risoluzione di controversie giudiziarie. Questa branca specialistica impiega le tecniche della biologia molecolare e le conoscenze della genetica per rispondere a quesiti di natura legale, trasformando tracce biologiche in prove scientifiche di elevato valore probatorio. Le moderne tecniche molecolari, permettono di determinare un profilo genetico oltre che da campioni biologici convenzionali come sangue, saliva, sperma, etc., anche a partire da tracce minime di materiale biologico quali sudore, impronte, tracce lavate, etc. Ciò consente il raffronto tra il profilo genetico di tracce biologiche repertate sulla scena di un crimine con quelle di eventuali sospettati e/o vittime, al fine di escludere o ricondurre in maniera precisa la provenienza ad uno specifico individuo. L'importanza del ruolo della genetica, tra tutte le scienze forensi che vengono utilizzate, si è notevolmente accresciuta negli anni, grazie ai progressi nella conoscenza dei marcatori genetici polimorfici e all'evoluzione degli strumenti utilizzati per analizzarli. La genetica forense si è rivelata nel tempo particolarmente utile per diverse finalità quali, ad esempio:

- Attribuire tracce biologiche prelevate nell'ambito di indagini forensi a soggetti di interesse per l'Autorità Giudiziaria;
- Stabilire la sussistenza o meno di relazioni parentali tra soggetti (indagini di paternità o di parentela, in generale);
- Predire la provenienza biogeografica di un soggetto nell'ambito di indagini investigative
- Fornire una valutazione probabilistica dell'aspetto esteriore (fenotipo) del soggetto a cui appartiene una certa traccia biologica (*DNA Phenotyping*), per finalità investigative.
- Identificare cadaveri la cui identità sia ignota (vittime di catastrofi naturali, disastri aerei, guerre, genocidi)
- Investigazioni su persone scomparse

Prima dell'introduzione delle moderne tecniche molecolari, le indagini forensi prevedevano la caratterizzazione di polimorfismi proteici umani. Le origini della genetica forense possono essere fatte risalire ai primi decenni del Novecento, quando la scoperta del sistema dei gruppi sanguigni AB0 da parte di *Karl Landsteiner* aprì la possibilità di utilizzare marcatori biologici ereditari per scopi identificativi e di esclusione di paternità. Il sistema AB0, insieme ad altri sistemi di gruppi sanguigni

successivamente identificati, rappresentò per decenni il principale strumento di analisi genetica forense, nonostante il suo limitato potere discriminante, dovuto al numero ridotto di fenotipi possibili. L'era dell'emogenetica forense, che si estese approssimativamente dalla scoperta dei gruppi sanguigni fino agli anni Ottanta del secolo scorso, si basava sull'analisi di marcatori proteici presenti nel sangue e in altri fluidi corporei. Questi marcatori includevano, oltre ai gruppi sanguigni, enzimi eritrocitari, proteine sieriche e antigeni del sistema *HLA* (*Human Leukocyte Antigen*). Tuttavia, l'approccio emogenetico presentava due limitazioni fondamentali: la rapida degradazione delle proteine in condizioni ambientali avverse e il relativamente basso potere di discriminazione che rendeva possibile solo l'esclusione di individui ma non la loro identificazione positiva con elevata certezza statistica. Un progresso significativo si verificò negli anni Settanta con l'introduzione di tecniche di elettroforesi enzimatica e immunologica più sofisticate, che permisero l'analisi di un numero maggiore di marcatori proteici in contemporanea. Tuttavia, fu solo con l'avvento delle tecnologie di biologia molecolare e la possibilità di analizzare direttamente il DNA che la genetica forense conobbe una vera rivoluzione. Una reale svolta nelle analisi di genetica forense a scopo identificativo si ebbe nel 1985 quando si iniziò a parlare di DNA *fingerprinting*. Questa è una tecnica che fu sviluppata dal genetista britannico *Alec Jeffreys* il quale dimostrò che specifiche regioni del DNA contenevano sequenze ripetitive altamente variabili tra gli individui; grazie alla grande variabilità di queste sequenze era possibile attribuire caratteristiche genetiche univoche ai soggetti così da ottenere una impronta digitale genetica perfettamente identificativa (*DNA fingerprint*). Negli studi di *Jeffreys* vennero analizzati i cosiddetti polimorfismi VNTRs, (*Variable Number of Tandem Repeats*), ovvero sequenze ripetute in tandem in numero variabile tra individui diversi, usando la tecnica del *Southern blotting* (Southern E. M., 1975). Tale approccio ha fornito i primi strumenti genetico-molecolari per discriminare gli individui all'interno della popolazione, permettendo la sostituzione dei vecchi metodi basati su proteine nelle analisi forensi (Jeffreys AJ, 2005). La prima applicazione giudiziaria di questa tecnologia avvenne nel 1986, quando *Jeffreys* collaborò con il *Forensic Science Service* britannico per risolvere un caso di duplice omicidio nel *Leicestershire*. L'analisi del DNA non solo permise di scagionare un sospetto inizialmente accusato, ma condusse all'identificazione del vero colpevole attraverso uno screening di massa della popolazione maschile locale, rappresentando il primo caso al mondo risolto mediante analisi del DNA. Questo evento segnò l'inizio dell'era della genetica forense moderna e dimostrò il potenziale rivoluzionario dell'analisi del DNA nelle indagini giudiziarie. Nonostante il suo impatto rivoluzionario, la tecnica originale di *Jeffreys* presentava alcune limitazioni pratiche significative. Il metodo richiedeva quantità relativamente elevate di DNA di buona qualità, tempi di analisi lunghi procedure tecnicamente complesse e costose, e risultava poco adatto all'analisi

di campioni degradati o in tracce. Queste limitazioni stimolarono la ricerca di marcatori genetici alternativi e di tecnologie analitiche più efficienti.

1.2 Variabilità genetica e polimorfismi del DNA

Il progetto per il sequenziamento dell'intero genoma umano, Progetto Genoma Umano (*HGP*), iniziato nel 1990 e terminato nel 2001, ha portato alla luce numerose novità riguardo l'organizzazione e complessità del nostro patrimonio genetico. Dall'analisi del Progetto Genoma Umano, inoltre, si è potuto capire che, prendendo due individui scelti in maniera casuale nella popolazione generale, questi condividono circa il 99.7% del materiale genetico (Butler, 2011). La porzione di DNA di interesse per l'identificazione personale risiede, quindi, nello 0.3% dell'intero genoma umano; è, infatti, in tale quota del genoma che sono presenti i marcatori d'elezione per lo studio della variabilità la cui caratterizzazione consente di rendere ogni individuo unico all'interno della popolazione mondiale (Butler, 2011). Il genoma umano contiene tra i 20000 e i 25000 geni codificanti per proteine, che corrispondono a circa il 5% del DNA totale (Lander et al., 2001; *International Human Genome Sequencing Consortium*, 2004). Le sequenze ripetute, invece, rappresentano circa il 50% del genoma: definite inizialmente come *junk* DNA sono poi state rivalutate in quanto ricche di informazioni per gli studi evolutivisti e utili strumenti per studi di mutazione e selezione. Anche se la funzione delle sequenze ripetute non è ad oggi conosciuta in termini esaustivi, se ne è ipotizzata l'associazione con numerose patologie mentre è certo, invece, il ruolo nell'espressione di alcuni geni (Fan, Chu, 2007). Queste sequenze ripetute si presentano in tutte le dimensioni e sono tipicamente designate dalla lunghezza dell'unità di ripetizione centrale (*core repeat*) e dal numero di unità di ripetizione contigue o dalla lunghezza complessiva della regione ripetuta. Le unità di ripetizione lunghe possono contenere da diverse centinaia a diverse migliaia di basi nella ripetizione centrale. Queste regioni sono chiamate DNA satellite e possono essere trovate intorno al centromero cromosomico con le sequenze α -satellite, ma il DNA satellite si può trovare anche nelle regioni pericentromeriche, in prossimità del centromero cromosomico, e nelle regioni subtelomeriche dove sono presenti delle sequenze *satellite-like* (Aldrup-MacDonald M.E, Sullivan B.A., 2018). Il termine satellite è nato dal fatto che frequentemente una o più "bande satelliti" minori venivano viste nei primi esperimenti che coinvolgevano la centrifugazione in gradiente di densità di equilibrio (Britten & Kohne, 1968).

Le sequenze ripetute possono essere distinte in base alla tipologia di ripetizione (Lander et al., 2001) come segue:

- Derivate da trasposoni;
- Copie di geni inattivati (pseudogeni);
- Segmenti duplicati di 10-300kb da altre parti del genoma;
- Centromeri, telomeri, i bracci corti dei cromosomi acrocentrici;
- Ripetizioni di sequenze semplici di poche basi.

La variabilità rilevante per la genetica forense è meglio descritta dal concetto di polimorfismo: una specifica regione genomica viene definita polimorfica se nella popolazione sono presenti almeno due forme alleliche, di cui quella più rara tra gli individui presenta una frequenza maggiore o uguale all'1%; in caso di frequenze inferiori a questa soglia, non si parla più di polimorfismo, ma di variante allelica. Per convenzione, i polimorfismi vengono suddivisi in due categorie sulla base del differente meccanismo molecolare che dà loro origine: i polimorfismi di sequenza a singola base e le ripetizioni di sequenze semplici.

Il primo tipo di polimorfismi sono le variazioni a singola base. Una variazione di una singola base tra individui in uno specifico punto del genoma è spesso definita polimorfismo a singolo nucleotide (*Single Nucleotide Polymorphism, SNP*). Gli *SNP* sono delle singole posizioni nucleotidiche in corrispondenza delle quali un evento di mutazione o un evento di inserzione/delezione ha alterato la sequenza delle basi e quindi il codice genetico in quella specifica posizione del genoma. Questo tipo di polimorfismi sono molto abbondanti nel genoma umano e, proprio per questa ragione, vengono utilizzati negli studi di linkage per tracciare le malattie genetiche (Anthony Brookes, 1999). Ogni individuo possiede milioni di *SNP* e questa elevata abbondanza implica che essi possono essere impiegati anche per distinguere un individuo da un altro. Sono state sviluppate numerose tecnologie per automatizzare il processo di analisi degli *SNP*. Milioni di *SNP* possono essere analizzati simultaneamente mediante tecnologie di sequenziamento di nuova generazione (*Next-Generation Sequencing, NGS*) (Morozova, Marra, 2008), offerte da aziende come Illumina (San Diego, California).

L'altra tipologia di elemento variabile presente nel DNA sono le ripetizioni di sequenze semplici caratterizzate da una ripetizione centrale di media lunghezza, sono definite *VNTR (Variable Number of Tandem Repeat)* o minisatelliti e, questi minisatelliti sono caratterizzati da una sequenza ripetuta (9-100 bp) che si ripete in tandem. A partire dagli anni '80, l'impiego dei minisatelliti ha rivoluzionato

l'analisi del profilo genetico. Oltre all'elevata variabilità interindividuale, questi marcatori presentano una trasmissione ereditaria di tipo mendeliano, rendendo la profilazione genetica (*DNA fingerprinting*) uno strumento altamente efficace anche per l'attribuzione di relazioni parentali. Come illustrato nella Figura 1, l'analisi autoradiografica dimostra l'elevato potere discriminativo del sistema, consentendo sia l'identificazione univoca di individui già tipizzati, sia l'evidenziazione di legami di consanguineità e l'individuazione di eventuali campioni duplicati (Jeffreys A. J., 2005).

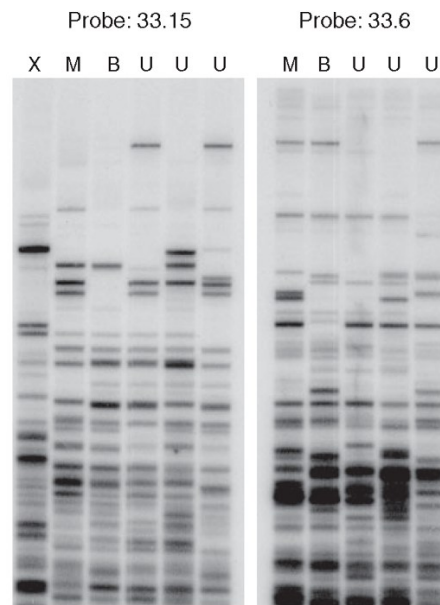


Figura 1 | Autoradiografia che mostra il profilo fingerprinting ottenuto dall'ibridazione di due differenti sonde (33.15 e 33.6) del DNA di una madre (M) e dei suoi 4 figli (B, U, U, U). Ogni figlio condivide una banda (allele) con la madre. X è un soggetto non correlato (Jeffreys A. J., 2005).

Tuttavia, la loro lunghezza risulta eccessiva nel caso di analisi di DNA degradato, come spesso accade in ambito forense (Gill, Jeffreys and Werrett, 1985). Infatti, le dimensioni degli alleli caratterizzati inizialmente da *Jeffreys et al.* andavano da 4 a 20 kb (Jeffreys, Wilson and Thein, 1985), incompatibili con DNA anche solo moderatamente degradato. Inoltre, la tecnica utilizzata per la caratterizzazione genetica (*Southern blotting*) prevedeva una discreta quantità iniziale di materiale biologico per consentire l'ottenimento di un profilo (Weber, 1990).

Un'altra categoria di ripetizioni di sequenze semplici sono i microsatelliti o *STR* (*Short Tandem Repeats*) che hanno delle sequenze ripetute più piccole in termini di numero di nucleotidi rispetto ai minisatelliti, da 2 a 6 bp (Tautz, 1993). Il numero di ripetizioni in tandem di questa sequenza identifica, per convenzione, l'allele d'interesse. Il marcatore minisatellite più comunemente usato negli anni '90 era il D1S80, che si trova sul cromosoma 1 e ha un'unità di ripetizione di 16 bp e contiene alleli che coprono l'intervallo da 14 a 41 unità di ripetizione (Kasai et al. 1990, Butler 2010).

La Tabella 1 mette a confronto i marcatori *SNP* e i marcatori *Short Tandem Repeat (STR)*, evidenziandone somiglianze e differenze.

Characteristics	Short Tandem Repeats (STRs)	Single Nucleotide Polymorphisms (SNPs)
Occurrence in human genome	≈1 in every 15kb	≈1 in every 1 kb
General informativeness	High	Low; only 20% to 30% as informative as STRs
Mutation rate	≈1 in 1000	≈1 in 100 000 000
Marker type	Di-, tri-, tetra-, penta-nucleotide repeat markers with many alleles	Mostly bi-allelic markers with six possibilities: A/G, C/T, A/T, C/G, T/G, A/C
Number of alleles per marker	Usually 5 to 20	Typically 2 (some tri-allelic SNPs exist)
Detection methods	Gel/capillary electrophoresis	Sequence analysis; microchip hybridization
Multiplex capability	> 10 markers with multiple fluorescent dyes	Difficult to amplify more than 50 SNPs well (detection of 1000s with microchips)
Amplicon size	≈75 to 400bp	Can be less than 100bp
Ability to predict ethnicity (biogeographical ancestry)	Limited	Some SNPs associated with ethnicity
Phenotypic information	No	Possible to predict some hair colors, etc.
Major advantages for forensic application	Many alleles enabling higher success rates for detecting and deciphering mixtures	PCR products can be made small potentially enabling higher success rates with degraded DNA samples; low mutation rate may aid kinship analysis; phenotype prediction
Limitations for forensic application	Data interpretation must account for artifacts such as dye blobs, stutter, spikes, etc.	No widely established core loci; large multiplexing assays required; mixture resolution issues/interpretation; multiple typing platforms make universal SNP selection difficult; population substructure due to low mutation rate

Tabella 1 | Comparazione tra le caratteristiche dei marcatori *STRs* e *SNPs* (Butler, 2011).

1.2.1 I polimorfismi di sequenza: *SNPs*

I *Single Nucleotide Polymorphism* – *SNPs* – sono polimorfismi che originano da un evento mutazionale che provoca la sostituzione di una singola base nucleotidica nella sequenza di DNA (Figura 2). Per la modalità di generazione, si tratta di polimorfismi completamente diversi dagli *STRs*, in quanto derivanti da una errata incorporazione di basi durante la replicazione del materiale genetico o generati da modificazioni chimiche e danni fisici a cui risulta esposta la doppia elica del DNA (Tagliabracci, 2009). Gli *SNP* sono abbondanti nel genoma umano e, come tali, vengono utilizzati negli studi di associazione (*linkage*) per tracciare le malattie genetiche (Brookes 1999). La peculiarità di questi polimorfismi risiede soprattutto nello scarso potere informativo, in quanto l'evento di mutazione determina, nella maggior parte dei casi, l'origine di 2 sole forme alleliche che, combinandosi tra loro, portano alla formazione di 3 possibili genotipi.

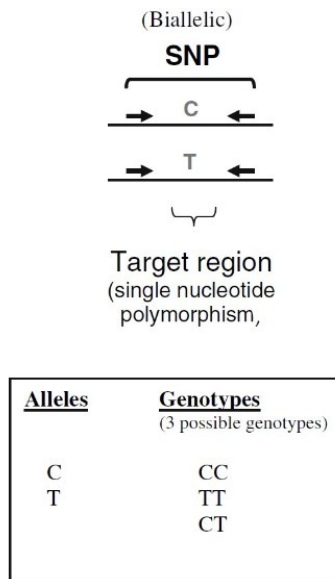


Figura 2 | Rappresentazione grafica semplificata di un *locus* con polimorfismo *SNP* con i due alleli possibili e le loro combinazioni per formare i 3 genotipi possibili (Butler, 2007).

Ad esempio, se gli alleli per un *locus SNP* sono A e B, i tre genotipi possibili dati dalla combinazione degli alleli sono AA, BB o AB. Il numero limitato di alleli rende ancora più complicata l'interpretazione di campioni misti, dato che gli alleli con una maggiore frequenza risultano comunemente condivisi nella popolazione generale (Butler, 2011). Nonostante lo scarso potere informativo, gli *SNP* sono stati presi in considerazione nelle analisi di DNA forense come potenziali marcatori genetici per diverse ragioni. In primo luogo, e cosa più importante, i prodotti della reazione a catena della polimerasi (*PCR*) ottenuti dagli *SNP* possono avere dimensioni inferiori a 100 bp e ciò significa che questi marcatori sono in grado di recuperare informazioni da campioni di DNA degradato meglio degli *STR*, i cui ampliconi possono raggiungere dimensioni comprese tra 300 bp e 400 bp. In secondo luogo, possono essere inseriti in sistemi *multiplex* in grado di caratterizzare contestualmente centinaia o migliaia di marcatori, poiché alcuni metodi di rilevazione, come ad esempio l'ibridazione su *microarray* superano i limiti legati alla possibilità di amplificare solo un numero limitato di marcatori *STR* a causa del numero di fluorocromi utilizzabile e alla limitazione dei sistemi di *multiplex PCR* e separazione elettroforetica. In terzo luogo, il processamento dei campioni e l'analisi dei dati possono essere automatizzati in modo più completo, poiché non è necessaria una separazione basata sulle dimensioni. In quarto luogo, non vi sono artefatti di polimerizzazione tipici dei marcatori *STR* (picchi stutter, artefatti di cui si parlerà successivamente) che possono complicare l'interpretazione dei profili. Questo dovrebbe contribuire a semplificare l'interpretazione della determinazione degli alleli. Infine, questi marcatori polimorfici, per contro, sono piuttosto stabili in quanto presentano un tasso di mutazione di circa 10^{-9} - 10^{-11} per gamete,

per generazione (Tagliabracci, 2009), di gran lunga inferiore rispetto al tasso di mutazione caratteristico degli *STRs* di circa 2×10^{-3} /gamete/generazione; ciò li rende idonei ed estremamente informativi in contesti analitici che si discostano dall'identificazione personale ma che risultano, in tutti i casi, fondamentali per la previsione dell'ascendenza biogeografica e dei caratteri fenotipici (Butler, 2007). Sono state sviluppate numerose tecnologie per miniaturizzare e automatizzare la procedura di analisi degli *SNP*. Oggi è possibile elaborare simultaneamente milioni di *SNP* utilizzando le tecnologie di sequenziamento di nuova generazione (*NGS*, *Next-Generation Sequencing*) (Ku et al. 2010). In relazione a quest'ultima applicazione in ambito forense, Walsh e collaboratori (Walsh et al., 2012) hanno messo a punto un pannello *SNPs* che consente di effettuare una previsione circa il colore degli occhi di un soggetto a partire dal DNA estratto da una traccia. Per una più completa ricostruzione delle caratteristiche fenotipiche di un soggetto, è stato fondato il consorzio *Visible Attributes through Genomics – VISAGE* – il quale ha messo a punto un pannello contenente 524 *SNPs* in grado di fornire informazioni fenotipiche aggiuntive quali la presenza/assenza di lentiggini, il colore delle sopracciglia, la predisposizione alla calvizie, la tipologia di capello (liscio o riccio), l'ascendenza biogeografica e l'età (Xavier et al., 2022). Il progetto *VISAGE* è attivo e in corso ancora oggi e ha permesso di raggiungere un elevato potere informativo sulla predizione dell'età, provenienza geografica e caratteristiche fenotipiche andando ad aumentare il numero di marcatori *SNP* e unendo pannelli di *SNPs* del DNA nucleare autosomi e specifici del cromosoma Y con un pannello di *SNPs* del DNA mitocondriale (mtDNA) (Amory et al., 2026). Le principali categorie di *SNPs* oggetto di studio e di applicazioni in ambito forense sono le seguenti (Butler et al., 2008; Butler, 2011).

- *LISNPs – Lineage Informative SNPs* – pannello di *SNPs* in *linkage disequilibrium* che porta alla formazione di un blocco alplotipico, cioè un marcatore multi-allelico, in grado di restituire informazioni relative a possibili relazioni parentali in quanto caratterizzati da trasmissione uniparentale (*SNPs* presenti nel DNA mitocondriale e *SNPs* localizzati sul cromosoma Y) (Ge et al., 2010);
- *AISNPs – Ancestry Informative SNPs* – pannello di *SNPs* autosomici che fornisce indicazioni circa l'ascendenza biogeografica di un individuo; a causa di un ridotto tasso di mutazione, gli *SNPs* vengono tramandati inalterati tra generazioni successive, sedimentandosi in maniera specifica in ciascuna popolazione (Phillips et al., 2014);

- *PISNPs – Phenotype Informative SNPs* – pannello di *SNPs* che restituisce una predizione circa le caratteristiche fenotipiche di un soggetto ignoto, quali colore dei capelli, della pelle e degli occhi e altre caratteristiche fenotipiche (Walsh et al., 2012).

1.2.2 Polimorfismi di lunghezza *STR*

I genomi eucariotici sono ricchi di sequenze di DNA ripetute (Ellegren 2004). Queste sequenze possono avere dimensioni molto diverse e vengono generalmente classificate in base:

- alla lunghezza dell'unità di ripetizione fondamentale (*core repeat unit*);
- al numero di unità ripetute contigue;
- oppure alla lunghezza complessiva della regione ripetuta.

Le regioni di DNA caratterizzate da polimorfismi basati sulla variazione nel numero di unità ripetute di lunghezza tra 9 e 100 bp sono chiamate “*ministelliti*”. Questi sono i polimorfismi del DNA caratterizzati inizialmente da Jeffreys a definire il DNA “*DNA fingerprinting*”. Che presentano un elevato numero di alleli di peso molecolare tra 0.4 e 10 kb, la cui caratterizzazione con la metodica Southern blotting è complessa. Polimorfismi con unità ripetute lunghe da 2 a 7 bp sono chiamati “*microsatelliti*”, o *SSRs (Simple Sequence Repeats)* oppure, più comunemente, *STRs (Short Tandem Repeats)*. La caratterizzazione di tali marcatori avviene in PCR e il peso molecolare degli alleli varia tra 80 a 500 bp, compatibile con DNA degradato. La nomenclatura degli alleli dei marcatori STR si basa sul numero dell'unità di ripetizione che li caratterizza. Quindi gli alleli sono identificati direttamente dal numero di volte che una certa sequenza si ripete in tandem; per esempio, se una sequenza “TCTA” sul cromosoma materno di un certo individuo è ripetuta 8 volte e su quello paterno è ripetuta 6 volte, allora questo individuo per questo *STR* con sequenza “TCTA” è eterozigote e i suoi alleli sono 8 e 6 (Tagliabracci, 2009) (Figura 3).

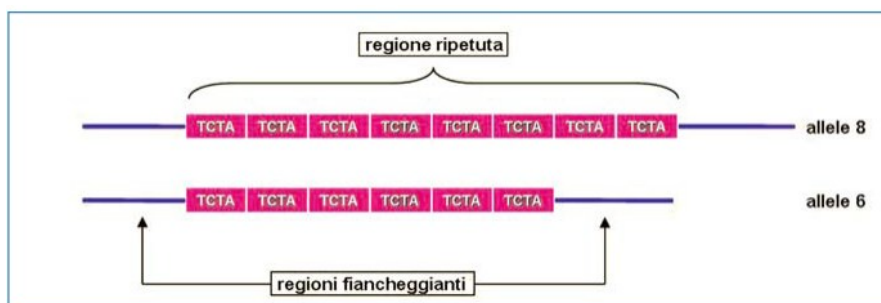


Figura 3| Struttura di un *STR*. I due alleli differiscono nella lunghezza della regione ripetuta, costituita da 8 ripetizioni del tetranucleotide TCTA nell'allele in alto e da 6 ripetizioni in quello in basso. La regione ripetuta è delimitata dalle regioni fiancheggiatrici (linea blu) identiche in entrambi gli alleli (Tagliabracci, 2009).

Le sequenze ripetute di tipo *STR* prendono il nome dalla lunghezza dell'unità di ripetizione e ogni *repeat* può contenere un diverso numero di basi. Le ripetizioni dinucleotidiche presentano due nucleotidi ripetuti consecutivamente più e più volte. Le trinucleotidiche hanno tre nucleotidi nell'unità di ripetizione, le tetranucleotidiche ne hanno quattro, le pentanucleotidiche cinque e le esanucleotidiche sei nucleotidi nella ripetizione di base (Figura 4). Per le ragioni che verranno discusse in seguito, le ripetizioni tetranucleotidiche sono diventate i marcatori *STR* di elezione per l'applicazione in ambito identificativo forense.

- **Dinucleotide** (CA)(CA)(CA)(CA)
- **Trinucleotide** (GCC)(GCC)(GCC)
- **Tetranucleotide** (AATG)(AATG)(AATG)
- **Pentanucleotide** (AGAAA)(AGAAA)
- **Esanucleotide** (AGTACA)(AGTACA)

Figura 4 | *STR* con lunghezza diversa dell'unità di ripetizione (modificato da Ahlquist, 2011).

Gli *STR* sono spesso divisi in diverse categorie in base alla sequenza dell'unità ripetuta:

- Ripetizioni semplici: contengono unità di lunghezza e sequenza identiche;
- Ripetizioni composte: comprendono due o più ripetizioni semplici adiacenti;
- Ripetizioni complesse: possono contenere diversi blocchi di ripetizione con unità di lunghezza variabile, oltre a sequenze interposte variabili (Urquhart et al. 1994) (Tabella 2).

Non tutti gli alleli di un *locus STR* contengono unità di ripetizione complete. Persino le ripetizioni semplici possono contenere alleli *non-consensus* che si collocano tra gli alleli con unità di ripetizione intere. Le microvarianti sono alleli che contengono unità di ripetizione incomplete. Forse l'esempio più comune di microvariante è l'allele 9.3 al *locus* TH01, che contiene nove ripetizioni tetranucleotidiche e una ripetizione incompleta di tre nucleotidi (Puers et al. 1993) (Butler, 2011) (Tabella 2).

Categoria di <i>STR</i>	Esempio di sequenza	13 <i>CODIS Loci</i>
Ripetizioni semplici	(GATA)(GATA)(GATA)	TPOX, CSF1PO, D5S818, D13S317, D16S539
Ripetizioni composte	(GATA)(GATA)(GACA)	VWA, FGA, D3S1358, D8S1179
Ripetizioni complesse	(GATA)(GACA)(GA)(CATA)	D21S11
Ripetizioni semplici con alleli non- consensus (TH01 9.3)	(GATA)(GAT-)(GATA)	TH01, D18S51, D7S820

Tabella 2 | Diverse categorie di marcatori *STR* con i relativi esempi di struttura delle unità ripetute e *loci* del *CODIS* (modificato da Urquhart et al., 1994).

Esistono anche ripetizioni ipervariabili complesse con numerosi alleli *non-consensus* che differiscono sia per dimensioni che per sequenza, e sono quindi difficili da genotipizzare in modo riproducibile (Urquhart et al. 1993, Gill et al. 1994). Quest'ultima categoria di marcatori *STR* non è comunemente utilizzata nella tipizzazione del DNA forense a causa delle difficoltà nella nomenclatura degli alleli e della variabilità di misurazione tra i laboratori, sebbene diversi kit commerciali includano oggi il *locus STR* ipervariabile complesso SE33, talvolta chiamato ACTBP2 (Urquhart et al. 1993) (Figura 5).

Allele no.	Repeat segment														Length (bp)
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
	(AAAG) ₂ AA	(AAAG) ₃ AG	(AAAG) ₃ AG	(AAAG) ₁₁ (AA)	1 ⁻ ----	(AAAG) ₁₅ GAAGG	(AAAG) ₂ -----	(AG) ₁ CAAG							291
1	(AAAG) ₂ AA	(AAAG) ₃ AG	(AAAG) ₃ AG	(AAAG) ₁₁ -----						GAAG	-(AAAG) ₂ -----	(AG) ₁ CAAG			228
2	(AAAG) ₂ AA	(AAAG) ₂ AG	(AAAG) ₃ AG	(AAAG) ₁₁ ----	(AG) ₁ (AAAG) ₁₉ GAAGG	(AAAG) ₂ -----	(AG) ₁ CAAG								303
3	(AAAG) ₂ AA	(AAAG) ₃ AG	(AAAG) ₃ AG	(AAAG) ₂₀ -----						G	(AAAG) ₃ -----	(AG) ₁ CAAG			265
4	(AAAG) ₂ AA	(AAAG) ₃ AG	(AAAG) ₃ AG	(AAAG) ₁₂ (AA)	1 ⁻ ----	(AAAG) ₁₈ GAAGG	(AAAG) ₂ -----	(AG) ₁ CAAG							307
5	(AAAG) ₂ -----	AG	(AAAG) ₃ AG	(AAAG) ₁₈ -----										(AG) ₁ CAAG	230

Figura 5 | Sequenze delle regioni variabili degli alleli nel *locus STR* SE33 (ACTBP2) (Urquhart et al. 1993).

Ai fini dell'identificazione umana, è fondamentale disporre di marcatori del DNA che siano altamente polimorfici, che mostrino cioè la massima variazione possibile tra individui diversi, e che possano essere combinati per ottimizzare la capacità di discriminare tra i campioni provenienti da soggetti diversi. Inoltre, è necessario considerare marcatori compatibili con la caratterizzazione di DNA gravemente degradato e frammentato, quale quello tipico di reperti forensi. A tale scopo l'impiego di *loci STR* nella tipizzazione di del DNA forense, rispetto ai minisatelliti *VNTR* presenta diversi vantaggi.

- Un range delle dimensioni in bp degli alleli limitato che permette il *multiplexing*, cioè l'analisi simultanea di più *loci* in un'unica reazione di amplificazione (*PCR*);
- La capacità di generare prodotti di *PCR* in un range di peso molecolare compatibile con quello di DNA molto degradati. Le dimensioni ridotte degli alleli *STR* (da circa 100 bp a 400 bp) rispetto agli alleli dei minisatelliti *VNTR* (da circa 90 bp a 1000 bp) rendono i marcatori *STR* candidati migliori per l'uso in applicazioni forensi dove il DNA degradato è condizione frequente (Butler, 2011). L'amplificazione *PCR* di campioni di DNA degradato può essere realizzata con successo più facilmente a partire da frammenti di DNA di dimensioni minori che quindi hanno più probabilità di essere amplificati mantenendo integra tutta la dimensione degli alleli *STR*. Se invece venissero utilizzati i marcatori *VNTR* per applicazioni forensi, in molti casi non si riuscirebbe a caratterizzare per intero tutto il marcatore, per assenza di DNA in quel range di peso molecolare.

Pertanto, sia per ragioni molecolari che tecnologiche, la caratterizzazione degli *STR* risulta più favorevole rispetto a quella dei minisatelliti (Tagliabracci, 2009) (Butler, 2011).

Nel caso di tipizzazione di DNA molto degradato è, inoltre, possibile pianificare un approccio utilizzando ampliconi *STR* di dimensioni ridotte, definiti *miniSTR*. Questi non sono altro che i normali marcatori *STR* che vengono utilizzati per l'identificazione personale, ma con la sola differenza che l'amplicone di *PCR* di un *miniSTR* è di dimensioni ridotte rispetto a quello di un classico amplicone *STR*. Questo è semplicemente ottenuto individuando sequenze dei *primers* in posizioni più prossime alla regione polimorfica composta dalle sequenze ripetute in tandem (Figura 6) (Coble & Butler, 2005) (Hill, 2008). In questo modo si ottiene un prodotto di *PCR* di dimensioni ridotte in termini di bp, in grado di fornire genotipi per le applicazioni forensi.

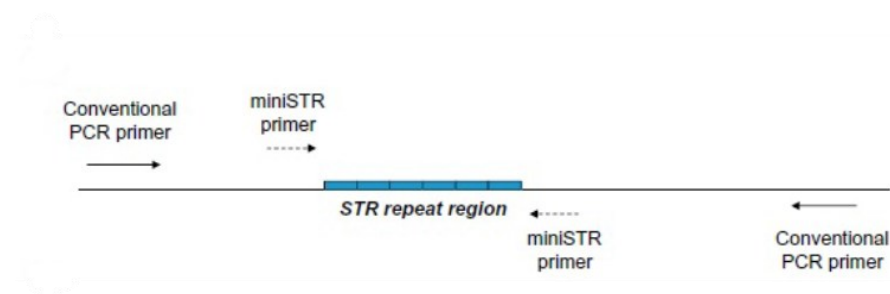


Figura 6 | Rappresentazione grafica della differenza tra un amplicone classico *STR* e un amplicone *miniSTR* (Butler Jhon .M, 2007).

I polimorfismi *STR* di elezione per l'analisi genetico forense sono quelli a base tetramerica (Bacher et al. 1999). Tra i vantaggi dell'utilizzo di tali marcatori nella tipizzazione di DNA in ambito forense, ricordiamo i due principali, qui di seguito riportati:

- Ridotta generazione di prodotti *stutter*; questi sono degli ineliminabili prodotti di PCR artefattuali che si generano in ragione della struttura a unità ripetute di questi polimorfismi *STR*. I prodotti *stutter* sono ampliconi che hanno tipicamente una unità di ripetizione in meno ($n-1$) o una unità di ripetizione in più ($n+1$) rispetto all'allele reale e si generano durante l'amplificazione dell'allele caratteristico del campione (Figura 7).

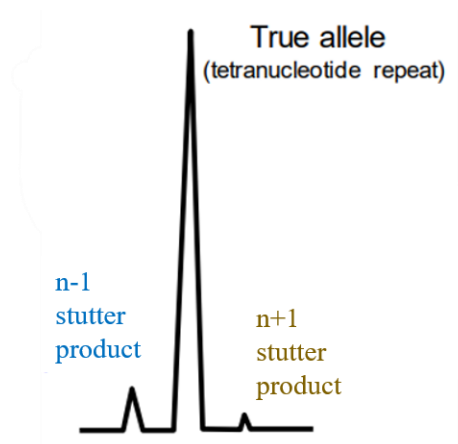


Figura 7 | Esempio di un picco elettroforetico di un allele *STR* a base tetramerica e le sue *stutter* correlate ($n-1$) e ($n+1$). (modificato da strbase.nist.gov).

Questo fenomeno di formazione dei picchi *stutter* è causato dallo scivolamento della DNA polimerasi che "slitta" sul template (*strand slippage*), producendo un prodotto accessorio che differisce dall'allele reale di esattamente una unità ripetitiva (Figura 8).

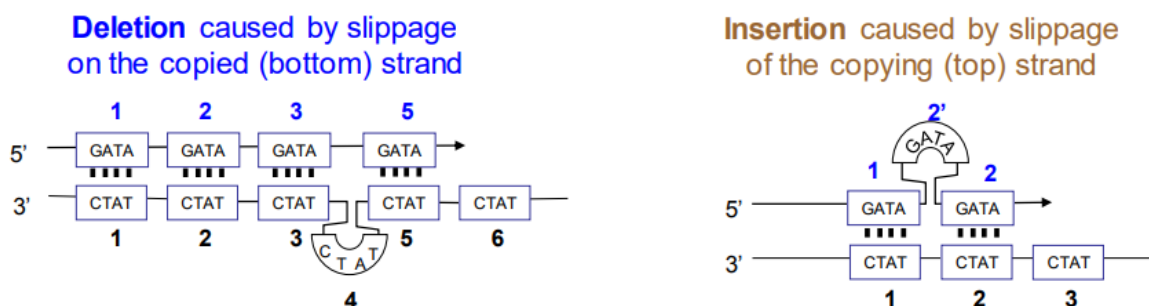


Figura 8 | Rappresentazione grafica dello *slippage* della DNA Polimerasi che causa delezione o inserzione di una ripetizione in più durante il processo di amplificazione di un marcatore *STR*. Questa delezione o inserzione di una unità ripetuta si manifesta, nel tracciato elettroforetico, come picco *stutter* (modificata da strbase.nist.gov).

Questo prodotto anomalo appare nell'elettroferogramma come un picco più piccolo, posizionato a sinistra del picco principale (Figura 7) (Walsh et al. 1996). L'altezza del prodotto di *stutter* varia a seconda del *locus STR* e persino della lunghezza dell'allele all'interno del *locus*, ma è solitamente inferiore al 10-12% dell'altezza dell'allele caratteristico del campione. Più piccola è la lunghezza dell'unità ripetuta, maggiore è l'altezza del picco *stutter*. Per questo la scelta dei marcatori è caduta su quelli a base tetranucleotidica, in relazione ai quali l'impatto della *stutter* è limitato. L'ambiguità di tale artefatto è legata alla possibile difficoltà nell'interpretazione dei campioni che contengono una mistura di DNA poiché dei picchi *stutter* possono essere confusi con picchi che rappresentano invece alleli reali e viceversa;

- Migliore risoluzione nell'elettroforesi capillare; la separazione dei frammenti avviene in base alla dimensione in bp. Alleli contigui di un *locus STR* devono essere distinguibili come picchi separati. La separazione elettroforetica dei marcatori a base tetrametrica è migliore rispetto a quella dei marcatori a base di o trinucleotidica (Butler, 2010, 2011).

I criteri di selezione per i *loci STR* a base tetrametrica candidati nelle applicazioni di identificazione umana includono le seguenti caratteristiche (Gill et al. 1996, Carracedo e Lareu 1998):

- Elevato potere discriminante, con un'eterozigosi osservata > 70%;
- Posizioni su cromosomi distinti (o, se sullo stesso cromosoma, ampiamente distanziate) per garantire l'indipendenza nella segregazione dei marcatori;
- Robustezza e riproducibilità dei risultati quando analizzati in *multiplex* con altri marcatori;
- Minore impatto dell'artefatto di PCR "*stutter*";
- tasso di mutazione relativamente basso (2×10^{-3} /gamete/generazione);
- Lunghezza prevista degli alleli compresa nell'intervallo tra 90 bp e 500 bp, con dimensioni minori più adatte all'analisi di campioni di DNA degradati.

Al fine di sfruttare la regola del prodotto (*product rule*) e poter combinare le informazioni genetiche di più *loci*, i marcatori *STR* utilizzati nella tipizzazione del DNA forense hanno localizzazioni su cromosomi diversi o, comunque, se sullo stesso cromosoma ampiamente distanziati, così da evitare qualsiasi problema di associazione (*linkage*) tra i marcatori.

1.3 Il profilo genetico forense

Il profilo genetico a fini identificativi forensi è l'insieme dei genotipi che caratterizza ogni singolo soggetto in relazione ai *loci* del DNA definiti nell'ambito di uno standard internazionale. Viene espresso con un codice nel quale vengono indicati i marcatori *STR* analizzati e i corrispondenti genotipi, sulla base del numero di ripetizioni in tandem che compongono gli alleli.

Le prime pubblicazioni che dimostrarono il potenziale dei marcatori *STR* per la tipizzazione del DNA umano, identificando *loci* che sarebbero poi diventati parte di uno standard forense, risalgono agli inizi degli anni 90'. Tra questi, quelli inizialmente caratterizzati e sviluppati nel laboratorio del Dr. *Thomas Caskey* presso il *Baylor College of Medicine* (Edwards, Hammond, Caskey et al. 1991). Il contributo fondamentale di tale ricercatore fu dimostrare che i tetranucleotidi erano superiori ai trinucleotidi per la tipizzazione umana, proprio per i motivi discussi in precedenza. Questo lavoro aprì concretamente la strada all'applicazione forense degli *STR*, che fino ad allora era dominata dai minisatelliti *VNTR* di *Jeffreys*. Solo due anni dopo, presso il *Forensic Science Service (FSS)* in Inghilterra, venne introdotto l'uso dei fluorofori per la marcatura dei primer e la rilevazione degli *STR*, rendendo possibile la futura automazione e il *multiplexing* in PCR (Kimpton et al. 1993). L'anno successivo fu sviluppato il primo vero sistema *multiplex* di marcatori *STR* a scopo identificativo forense, il cosiddetto "*quadruplex*" (*multiplex* a quattro *loci*) creato dal *Forensic Science Service*, che comprendeva i quattro *loci* TH01, FES/FPS, vWA e F13A1 (Lygo et al., 1994) (Kimpton et al. 1994). Questo cosiddetto "*multiplex* di prima generazione" aveva una *match probability*, cioè la probabilità di trovare due soggetti a caso nella popolazione con gli stessi genotipi, di circa 1 su 10.000 soggetti. Il FSS sviluppò successivamente una *multiplex* di seconda generazione (*SGM, Second-Generation Multiplex*) "*six-plex*" composta da sei *STR* polimorfici e da un marcatore per l'identificazione del sesso (marcatori TH01, vWA, FGA, D8S1179, D18S51 e D21S11 e Amelogenina) in grado di fornire una *match probability* di circa 1 su 50 milioni (Sparkes et al. 1996). Successivamente alla reazione di *multiplex PCR*, i prodotti di amplificazione venivano separati su gel di poliacrilammide denaturante e sequenziatori automatici. Il sistema si dimostrò robusto e riproducibile, specifico per il DNA umano, e adatto anche a materiale datato e degradato. A partire dagli anni successivi alla *Second generation multiplex*, alcune aziende commerciali iniziarono a produrre kit commerciali. In questo modo si iniziò a standardizzare, in maniera globale, il metodo di analisi di marcatori *STR* a scopo di identificazione personale in ambito forense. Infatti, grazie alla standardizzazione della metodica e alla selezione dei marcatori *STR* identificativi, nel 1995 il Regno Unito istituì lo *UK National DNA Database*, il primo sistema di archiviazione di profili genetici a livello nazionale, progettato per contribuire

all'identificazione di individui attraverso la comparazione di tali profili genetici con quelli rinvenuti sulle scene del crimine (Amankwaa & McCartney, 2019). Negli Stati Uniti, il Laboratorio dell'*FBI* sponsorizzò uno sforzo scientifico-forense a livello comunitario per stabilire i *loci STR* fondamentali da includere all'interno della banca dati nazionale del DNA, nota come *CODIS* (*COmbined DNA Index System*). Questo progetto coinvolse diversi laboratori che valutarono 13 *STR* fondamentali per la costruzione della banca dati nazionale statunitense del DNA (marcatori CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51 e D21S11 in Figura 9) (Budowle et al. 1998). Quando vennero testati tutti i 13 *loci* fondamentali del *CODIS*, la probabilità media di compatibilità casuale (*random match probability*) si dimostrò essere inferiore a una su mille miliardi ($1/10^{12}$) tra individui non imparentati all'interno della popolazione (Chakraborty et al. 1999).

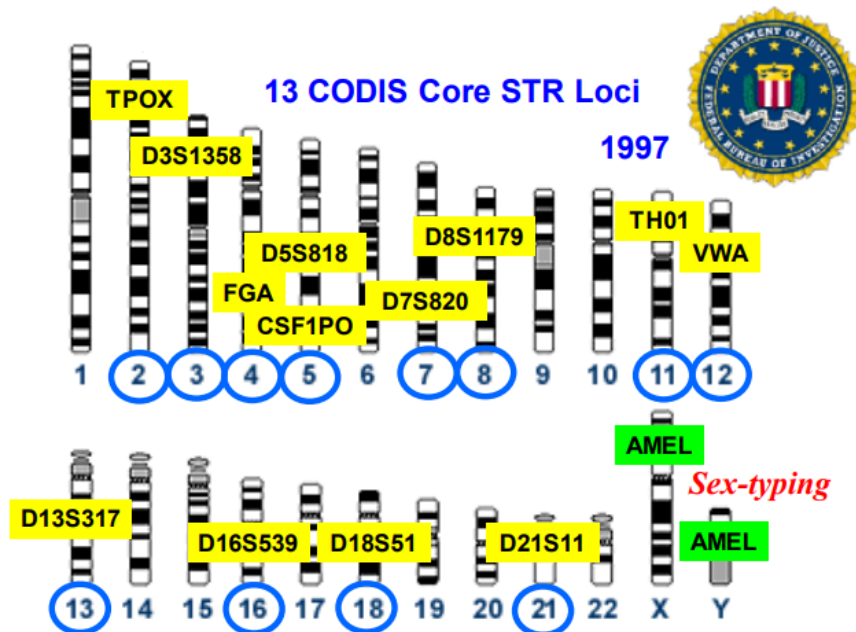


Figura 9 | Schema della posizione sui cromosomi dei marcatori *STR* del *CODIS* (strbase.nist.gov).

In Europa, simultaneamente all'istituzione del *CODIS*, sulla base del lavoro iniziale svolto dal *Forensic Science Service* (*FSS*), venne selezionato un set standard europeo di *loci STR*, denominato *ESS* (*European Standard Set*) (Schneider, 2009). L'*ESS* originale includeva i *loci STR* FGA, TH01, vWA, D3S1358, D8S1179, D18S51 e D21S11. Sei di tali sette *loci STR* sono quelli del sistema *SGM* con l'aggiunta del *locus* D3S1358. Una volta aggiunto anche il *locus* D16S539, che fa parte dei *loci* fondamentali del *CODIS*, si ottenevano otto *loci STR* sovrapponibili tra lo standard di marcatori europeo e statunitense. L'*ESS* ha subito un'implementazione negli anni, partendo da 7 *loci STR* nel

1999 fino a 12 *loci* nel 2009 (Gill et al., 2006; Schneider, 2009). Tuttavia, man mano che le banche dati del DNA crescevano numericamente, si rese necessario l'implementazione di un numero maggiore di *loci* per aumentare la capacità di discriminazione tra individui all'interno della popolazione al fine di evitare corrispondenze casuali tra due individui (*adventitious matches*). Mentre l'ESS europeo, e di conseguenza quello dell'Interpol (ISSOL), era fermo a dodici marcatori minimi da analizzare dopo la revisione del 2009, negli USA nel 2010 l'FBI istituì un CODIS Core Loci Working Group per valutare l'espansione del pannello, con l'obiettivo di minimizzare i match avventizi, migliorare la compatibilità internazionale e il potere discriminante per i casi di persone scomparse. A partire dal 2017, il set di marcatori STR del CODIS è stato implementato a venti *loci*, i 13 originali del 1997 con l'aggiunta di 7 nuovi marcatori (marcatori D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433 e D22S1045).

In Italia, prima dell'istituzione della Banca Dati Nazionale del DNA, non esisteva un set di marcatori standardizzato. I laboratori forensi italiani operavano adottando di fatto i kit commerciali europei disponibili, basati sull'ESS originale di 7 *loci* definito dall'ENFSI negli anni '90 e l'SGM (6 *loci*, 1995) che per molti anni costituì lo standard operativo *de facto* nei laboratori forensi europei, inclusi quelli italiani. Con la Legge 30 giugno 2009, n. 85, l'Italia ha aderito al Trattato di Prüm del 27 maggio 2005, volto a rafforzare la cooperazione di polizia in materia di lotta al terrorismo, alla criminalità transfrontaliera e all'immigrazione clandestina. Il Trattato disciplinava tra le parti l'impegno a creare archivi nazionali profili genetici e la possibilità di scambio di informazioni a livello internazionale. Con questa legge l'Italia si allineava formalmente all'ESS espanso a 12 *loci* che l'ENFSI stava adottando nello stesso anno (2009), aggiungendo ai 7 *loci* originali i cinque nuovi marcatori: D10S1248, D12S391, D22S1045, D1S1656 e D2S441, adottati con raccomandazione del Consiglio UE nel 2009. Successivamente sono stati pubblicati i regolamenti attuativi della Legge 85/2009 (D.P.R. 7 aprile 2016, n. 87) che hanno definito le regole per il funzionamento e l'alimentazione della Banca Dati Nazionale del DNA (BDND). Attualmente lo standard dei marcatori della BDND è a 21-24 *loci*, questo poiché i nuovi kit moderni più utilizzati permettono *multiplex* di questo numero di marcatori (Marafioti e Luparia, 2010). La standardizzazione dei marcatori ha consentito alle aziende di biologia molecolare di sviluppare kit commerciali *ad hoc* per l'applicazione forense. Tali kit consentono di effettuare reazioni di PCR *multiplex*, in cui tutti i marcatori selezionati sono amplificati contemporaneamente. Così facendo, è possibile massimizzare il risultato anche in caso di una bassa quantità di DNA (DNA Low Copy Number) o in presenza di DNA degradato. I kit sviluppati consentono inoltre di ottenere profili completi anche in presenza di inibitori condizione che in passato impediva totalmente o parzialmente l'amplificazione dei *loci* di interesse.

La specie-specificità dei primers elimina gli amplificati di DNA non umano, con solo rare eccezioni dovute a primati evolutivamente vicini all'uomo (Tucker et al., 2011).

Per ottenere un profilo genetico interpretabile è necessario sottoporre il campione in analisi ad amplificazione mediante *PCR*. I frammenti amplificati, corrispondenti agli alleli di ogni singolo marcatore, sono separati in elettroforesi capillare. Al fine di consentire l'attribuzione certa di un allele ad uno specifico marcatore, la separazione di questi avviene sulla base di specifici range di peso molecolare e di fluorocromi associati ai primer per l'amplificazione in *PCR* (Figura 10).

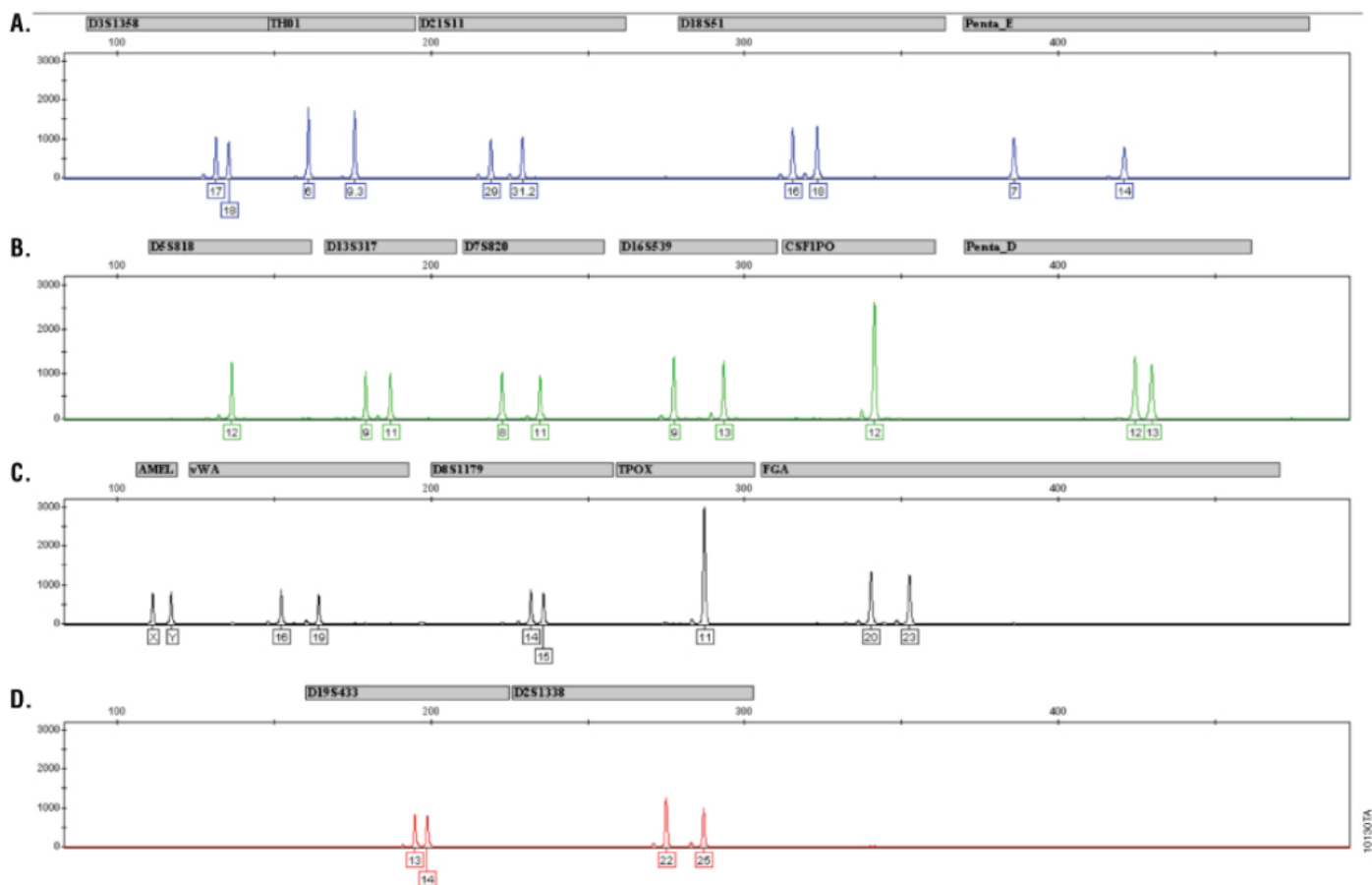


Figura 10 | Rappresentazione di un elettroferogramma del DNA di controllo 2800M (Promega) ottenuto tramite amplificazione con il kit *Promega PowerPlex® 18D System* (ita.promega.com). La scala sottostante ai riquadri grigi con i nomi dei marcatori *STR* (da circa 100 a 400 bp) è quella dei pesi molecolari dei frammenti amplificati. I marcatori sono identificati da specifici fluorocromi (blu, verde, nero e rosso). I picchi corrispondono, quindi, agli alleli amplificati per quel dato marcatore ad originare un genotipo omozigote (un solo allele) o eterozigote (due alleli). Ogni allele è identificato dal numero di ripetizioni, riportate nelle etichette sottostanti. Sull'asse verticale sono identificati gli *RFU* (*relative fluorescence unit*), ovvero la fluorescenza emessa da ciascun allele, direttamente proporzionale alla quantità di prodotto amplificato.

Nel caso di un campione biologico riconducibile ad un singolo soggetto si originerà un profilo genetico (tracciato elettroforetico) caratterizzato da uno o due picchi per ciascun *locus*, a seconda che il soggetto sia omozigote o eterozigote, come quello mostrato in Figura 10.

L'esame del DNA è un accertamento comparativo. Una volta acquisito il profilo genetico a partire da una traccia, sarà necessario compararlo con un soggetto di riferimento come, ad esempio, l'indagato, la vittima, o soggetti contenuti all'interno della Banca Dati Nazionale del DNA. I risultati della comparazione possono essere sostanzialmente due: nel caso in cui i genotipi ottenuti dalla traccia siano diversi da quelli del soggetto in analisi, si può giungere ad un giudizio perentorio di esclusione; nel caso in cui si registri una concordanza di genotipi traccia-indagato, si giungerà ad una valutazione di compatibilità rispetto all'ipotesi che tale soggetto abbia originato la traccia in questione, supportata necessariamente da una valutazione statistica. In questi casi si procede quindi al calcolo biostatistico che prevede di calcolare la frequenza nella popolazione della combinazione dei genotipi che compongono il profilo della traccia. Questa frequenza è denominata *Random Match Probability (RMP)* e indica la probabilità di trovare un altro soggetto casuale nella popolazione generale che condivide casualmente le stesse caratteristiche genetiche dell'individuo compatibile geneticamente con la traccia. Più è bassa questa frequenza, minori saranno le probabilità di rinvenire per caso nella popolazione generale un soggetto compatibile. La *RMP* viene calcolata moltiplicando fra loro le frequenze genotipiche dei marcatori previsti nel kit utilizzato. Nel caso di analisi con i 13 marcatori del vecchio *CODIS*, la *RMP* media è di circa 10^{-12} , mentre quella media di un'analisi di un pannello di 24 marcatori è di circa 10^{-26} .

Un secondo parametro statistico, strettamente correlato alla *RMP*, nel caso dell'analisi di un profilo a singolo contributore, è la *Likelihood ratio (LR)* in cui il profilo genetico viene analizzato mettendo a confronto due ipotesi alternative: al numeratore si avrà l'ipotesi dell'accusa, quindi che la traccia sia stata originata dal soggetto di interesse, mentre al denominatore l'ipotesi della difesa, ovvero che il profilo della traccia sia riconducibile ad un altro individuo presente nella popolazione, diverso dal, ma geneticamente compatibile. Il numeratore sarà, dunque, sempre 1, mentre il denominatore sarà pari alla *RMP*. Se il valore di *LR* è maggiore dell'unità, l'evidenza favorisce l'ipotesi dell'accusa, se è minore dell'unità favorisce l'ipotesi della difesa; se $LR = 1$, i dati sono perfettamente neutri rispetto alle due ipotesi. Nella seguente Tabella 3 vengono riportati i predicati verbali associati ai valori numerici di *LR* ottenuti a favore dell'accusa, come proposto dall'. (*European Network of Forensic Science Institutes*) nel documento del marzo 2015 "*Guideline for evaluative reporting in forensic science*". È del tutto ovvio che a valori di *LR* con esponente negativo corrispondono le medesime modulazioni a favore della difesa (Tagliabracci, 2009).

Values* of likelihood ratio	Verbal equivalent (two options of phrasing are suggested)
1	The forensic findings do not support one proposition over the other. The forensic findings provide no assistance in addressing the issue.
2 - 10	The forensic findings provide weak support** for the first proposition relative to the alternative. The forensic findings are slightly more probable given one proposition relative to the other.
10 - 100	...provide moderate support for the first proposition rather than the alternative ...are more probable given...proposition...than proposition...
100 - 1000	...provide moderately strong support for the first proposition rather than the alternative ...are appreciably more probable given... proposition...than proposition...
1000 - 10,000	...provide strong support for the first proposition rather than the alternative ...are much more probable given... proposition...than proposition...
10,000 - 1,000,000	...provide very strong support for the first proposition rather than the alternative ...are far more probable given... proposition...than proposition...
1,000,000 and above	...provide extremely strong support for the first proposition rather than the alternative ...are exceedingly more probable given... proposition...than proposition...

Tabella 3 | Associazione tra predicati verbali e *LR* in termini di valenza statistica (*ENFSI Guideline for Evaluative Reporting in Forensic Science*).

1.4 La degradazione del DNA

1.4.1 Il DNA come molecola chimicamente instabile: depurinazione, deaminazione e rottura del legame fosfodiesterico

Il DNA è comunemente percepito come una molecola stabile, custode fedele dell'informazione genetica nel corso delle generazioni. In realtà, a livello chimico, il DNA è una molecola intrinsecamente labile, soggetta a un continuo processo di decadimento spontaneo che avviene anche in condizioni fisiologiche, senza alcun intervento di agenti esterni. Già a partire dagli anni Settanta, i lavori pionieristici di *Lindahl* e collaboratori dimostrarono che il DNA nativo va incontro, ad ogni istante della vita cellulare in condizioni fisiologiche, a migliaia di eventi di danno spontaneo per ogni cellula, questi principali danni alla molecola di DNA sono la depurinazione (Figura 11), deaminazione (Figura 12) e rottura del legame fosfodiesterico.

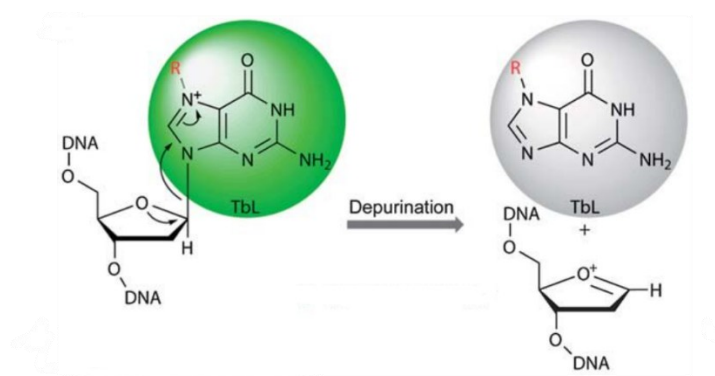


Figura 11| Rappresentazione grafica della reazione di depurinazione di un nucleotide della molecole di DNA (Modificato da: Wang et al., 2013).

La depurinazione (Figura 11) consiste nell'idrolisi del legame N-glicosidico che unisce una base purinica (adenina o guanina) al deossiribosio, con conseguente formazione di un sito abasico (sito AP, apurinico). Questo evento rappresenta una delle forme più frequenti di danno spontaneo al DNA. Gli studi sulla cinetica di questo processo, stimano che, a pH e temperatura fisiologici, un numero considerevole di basi puriniche venga perso da ciascuna cellula di mammifero ogni giorno (Lindahl T. et al., 1972). Studi successivi hanno mostrato che la depurinazione può avvenire anche in maniera sito-specifica e auto-catalizzata in particolari contesti di sequenza, un fenomeno che ha implicazioni sia per la stabilità genomica sia per l'interpretazione dei pattern di danno osservati nel DNA antico o degradato (Amosova et al., 2006). Il sito abasico che ne risulta non è di per sé una rottura del filamento, ma rappresenta una lesione instabile: se non riparato dal sistema di *Base Excision Repair*

(*BER*), può evolvere in una rottura del filamento per idrolisi del legame fosfodiesterico adiacente, oppure può causare l'inserimento preferenziale di un'adenina durante la replicazione (regola "*A-rule*"), generando mutazioni. La deaminazione idrolitica (Figura 12) consiste nella perdita del gruppo amminico esociclico delle basi, con conseguente conversione della citosina in uracile, della 5-metilcitosina in timina, dell'adenina in ipoxantina e della guanina in xantina. Questo tipo di danno è particolarmente insidioso perché altera direttamente le proprietà di appaiamento delle basi: l'uracile, ad esempio, si appaia con l'adenina anziché con la guanina, generando transizioni C→T se la lesione non viene corretta prima della replicazione (Lindahl, 1993). Il gruppo di *Lindahl* caratterizzò già negli anni Settanta la cinetica di deaminazione delle citosine nel DNA nativo a doppio filamento, mostrando come questo processo sia fortemente dipendente dalla temperatura e dal pH, e come sia accelerato nel DNA a singolo filamento rispetto a quello *duplex*, poiché l'appaiamento delle basi nella doppia elica offre una parziale protezione sterica (Lindahl, 1974).

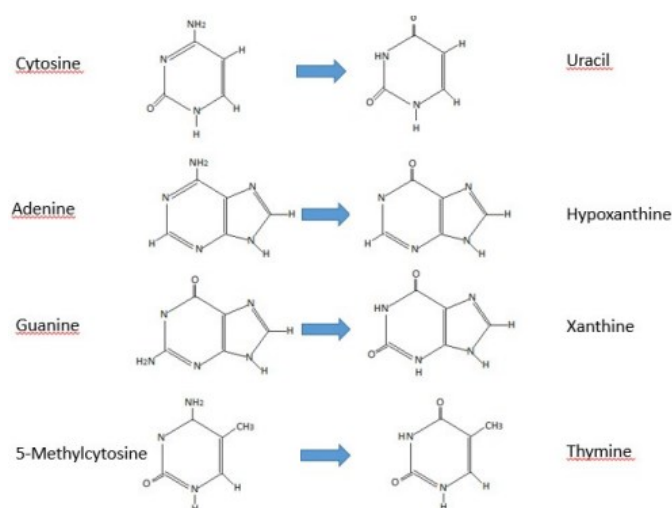


Figura 12 | Reazione di deaminazione di Citosina, Adenina, Guanina e 5-metilcitosina (Yoca, 2021).

Oltre al danno sulle basi, l'acqua può attaccare direttamente i legami fosfodiesterici dello scheletro zuccherino-fosfato, causando rotture dirette del filamento. Questo processo è notevolmente accelerato in presenza di siti abasici, poiché l'assenza della base destabilizza il legame glicosidico adiacente rendendolo più suscettibile alla scissione, sia per via chimica sia per l'azione di endonucleasi AP-specifiche (Lindahl et al., 1972). Questi tre processi idrolitici di depurinazione, deaminazione e rottura del *backbone*, agiscono in maniera sinergica e cumulativa nel tempo, rendendo la degradazione idrolitica una delle cause più importanti di frammentazione del DNA in campioni conservati per periodi prolungati, anche in assenza di stress ambientali particolarmente severi.

Questa instabilità chimica intrinseca non è un difetto evolutivo, ma una conseguenza inevitabile della struttura chimica dei nucleotidi: i legami N-glicosidici che uniscono le basi azotate allo scheletro zuccherino-fosfato, così come i gruppi amminici ciclici delle basi stesse, sono termodinamicamente instabili in ambiente acquoso. A questa fragilità si sommano numerosi fattori esogeni tra i quali radiazioni, agenti chimici, calore, attività microbica ed enzimatica che accelerano il decadimento molecolare, in particolare dopo la morte dell'organismo o in campioni conservati in condizioni non ottimali. Comprendere i meccanismi di degradazione del DNA è un tema centrale nelle analisi in genetica forense e in tutte le discipline che si basano sull'amplificazione e sul sequenziamento di acidi nucleici, poiché la qualità del template condiziona direttamente l'affidabilità del dato analitico ottenuto e la sua interpretazione finale (Lindahl T. et al., 1972).

1.4.2 Danno ossidativo

Accanto ai processi idrolitici, il danno ossidativo rappresenta una seconda, importantissima via di degradazione del DNA. Le specie reattive dell'ossigeno (*ROS*), generate sia da processi metabolici endogeni come la respirazione mitocondriale o i processi infiammatori sia da agenti esogeni come radiazioni, xenobiotici, calore, attaccano sia le basi azotate sia lo zucchero pentoso deossiribosio. Tra le basi, la guanina è la più suscettibile all'ossidazione, a causa del suo basso potenziale *redox*. Il prodotto di ossidazione più studiato è la 8-oxo-7,8-diidroguanina (8-oxoG), oggi universalmente riconosciuta come il biomarcatore di riferimento del danno ossidativo al DNA, sia in ambito clinico sia in ambito tossicologico e forense (Figura 13). È stato dimostrato che anche il semplice riscaldamento di soluzioni acquose può indurre la formazione di specie reattive dell'ossigeno e, di conseguenza, danno ossidativo al DNA: un dato che collega direttamente lo stress termico alla componente ossidativa della degradazione, mostrando come le diverse vie di danno raramente agiscano in maniera del tutto indipendente. Se non riparata dai sistemi di *Base Excision Repair*, la 8-oxoG può appaiarsi erroneamente con l'adenina durante la replicazione, generando transversioni G→T, una delle firme mutazionali più comuni osservate nei tumori umani e nei campioni sottoposti a stress ossidativo cronico. Il danno ossidativo, inoltre, può colpire direttamente lo zucchero desossiribosio, generando rotture dirette del filamento senza il passaggio intermedio per un sito abasico (Bruskov V.I et al., 2002).

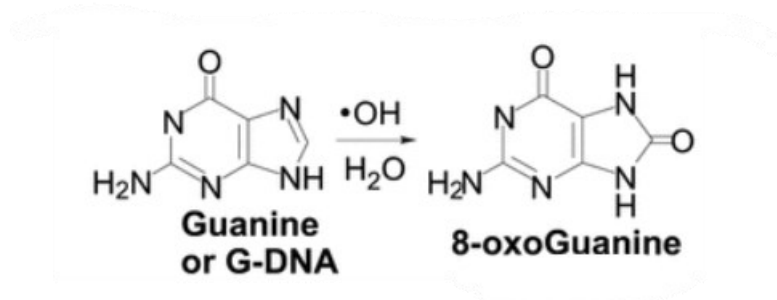


Figura 13| Reazione di ossidazione della Guanina in 8-oxo-7,8-diidroguanina la quale si appaia preferenzialmente con l'adenina causando una mutazione (Modificato da: White B., 2003).

1.4.3 La degradazione enzimatica

Diversamente dai processi chimici spontanei descritti nei capitoli precedenti, la degradazione enzimatica è mediata da proteine specializzate come le nucleasi che sono enzimi capaci di idrolizzare i legami fosfodiesterici con elevata efficienza e, spesso, con specificità di sequenza o di struttura. Le nucleasi endogene rientrano tra i primi agenti ad avviare il processo di frammentazione del DNA nel periodo *post-mortem* (Hofreiter et al. 2001). Quando avviene il processo di morte cellulare per apoptosi, entrano in azione degli enzimi lisosomiali che digeriscono le proteine istoniche legate alla cromatina: ciò facilita la successiva degradazione del DNA da parte delle endonucleasi, tra cui le DNasi I e II, riducendo quindi il peso molecolare dell'acido nucleico. La digestione delle proteine della cromatina da parte delle proteasi lisosomiali facilita, a sua volta, il processo di digestione casuale del DNA da parte delle endonucleasi. Ciò produce le tipiche scie di DNA (*smear*) all'elettroforesi su gel, in particolare nei casi medico-legali *post-mortem*. Diversi fattori influenzano la velocità della frammentazione (e di conseguenza la rilevabilità del DNA residuo nei tessuti), tra cui il livello di espressione degli enzimi, la disponibilità di cationi e di soluzione per le reazioni, la temperatura e i livelli di pH (Anzai *et al.* 1995, Alaeddini, Walsh, and Abbas 2010). Si distinguono due grandi categorie funzionali: le esonucleasi, che degradano il DNA progressivamente a partire dalle estremità libere, e le endonucleasi, che tagliano internamente al filamento in siti specifici o in corrispondenza di lesioni strutturali. In base alla periodicità del ripiegamento del DNA nella struttura della cromatina (J. Filipiak et al., 1990), a condizione che la struttura istonica legata ai nucleosomi rimanga intatta, il DNA può andare incontro a una serie di fasi successive di frammentazione in relazione alle quali i prodotti della frammentazione possono essere di grandi dimensioni, quindi ad alto peso molecolare, (*HMW*) o frammenti di piccole dimensioni a basso peso molecolare (*LMW*). Le nucleasi intervengono fisiologicamente in numerosi processi cellulari: la replicazione e la riparazione del DNA, la

ricombinazione, la maturazione degli acidi nucleici e, in particolare, l'apoptosi, durante la quale endonucleasi specifiche frammentano il genoma in maniera ordinata come parte del programma di morte cellulare. In condizioni favorevoli, come le basse temperature, o laddove esistano misure protettive naturali dipendenti dal tessuto, gli enzimi nucleasici possono inattivarsi prima di frammentare completamente i lunghi filamenti di DNA in frammenti più corti e meno informativi. I frammenti *HMW* vengono prodotti in presenza di Mg^{2+} come unico catione bivalente (P. Huang et al., 1997), mentre i frammenti oligonucleosomiali vengono generati quando nel tampone di digestione sono presenti ioni divalenti come il Mg^{2+} e il Ca^{2+} (X.M. Sun, Cohen G. M., 1994). È stato riportato che la DNasi Ca^{2+}/Mg^{2+} -dipendente è una caratteristica principale dell'apoptosi (Nakamura M. et al., 1981). Milza, fegato, reni e timo esprimono alti livelli di questo enzima, mentre polmoni, cervello, cuore e testicoli mostrano una scarsa attività. Nei tipi cellulari privi di attività calcio-dipendente, avviene una frammentazione del DNA differente attraverso l'azione di altri gruppi di endonucleasi (Barry M.A., Eastman A., 1993). Inoltre la rottura delle membrane cellulari durante le fasi tardive della morte cellulare porta al rilascio di fluidi ricchi di nutrienti, che favoriscono la crescita di microrganismi ambientali e, di conseguenza, un'ulteriore degradazione delle macromolecole, in particolare di acidi nucleici (Hofreiter et al. 2001). Più del 70% dei microrganismi del suolo contiene enzimi nucleasici. I microrganismi predominanti nel terreno sono in grado di decomporre gli acidi nucleici e la maggior parte dei loro prodotti di degradazione. Nel contesto *post-mortem* e in ambito forense, l'attività nucleasica endogena e, soprattutto, quella di origine microbica costituisce una delle cause principali di degradazione rapida del DNA nei tessuti biologici non adeguatamente conservati, agendo in sinergia con i processi idrolitici e ossidativi già descritti.

1.4.4 Degradazione da agenti chimici e fisici

Esistono moltissimi agenti chimici e fisici che possono entrare in contatto con le cellule di tessuti biologici e ne causano dei danni. In particolare, molti agenti chimici vanno a reagire con la molecola di DNA e ne causano una modificazione strutturale alterandone la morfologia e di conseguenza causando danni. Questi agenti chimici possono essere di diversa natura e vanno a interagire con la molecola di DNA in diversi modi; i principali sono gli agenti alchilanti, gli agenti intercalanti e gli agenti *cross-linkanti*. Gli agenti alchilanti trasferiscono gruppi metilici o etilici alle basi azotate, alterandone le proprietà di appaiamento e la struttura elettronica (Kondo N. et al., 2010). Gli agenti intercalanti si inseriscono tra le coppie di basi, distorcendo la geometria dell'elica e interferendo con i processi di replicazione e la trascrizione, fondamentali per la vita di una cellula (Palchaudhuri,

Hergenrother, 2007). Gli agenti *cross-linkanti*, come alcuni farmaci chemioterapici, creano legami covalenti tra i due filamenti o tra basi dello stesso filamento, impedendo la separazione dei filamenti necessaria per la replicazione (Deans A.J., West S.C., 2011) (Rajski S. R., Williams R.M, 1998). In tutti questi casi si ha una modificazione della morfologia della molecola di DNA che subirà diverse conseguenze, tutte dannose. Tra i principali agenti fisici che vanno a danneggiare la molecola di DNA ci sono le radiazioni, quelle più studiate in questo ambito ci sono le radiazioni ultraviolette (UV) e le radiazioni ionizzanti a più alta energia. Le radiazioni ultraviolette inducono la formazione di dimeri di pirimidine adiacenti, in particolare dimeri ciclobutano-pirimidinici (CPD) tra timine contigue, che distorcono localmente la doppia elica e bloccano la progressione delle polimerasi se non rimossi dal sistema di *Nucleotide Excision Repair* (NER) (Rastogi R.P.et al., 2010). Le radiazioni ionizzanti come i raggi X e i raggi gamma, invece, possono indurre rotture dirette a singolo o doppio filamento, oltre a generare specie reattive dell'ossigeno per radiolisi dell'acqua, sommando quindi un effetto diretto e uno indiretto sul danno al DNA (Saini et al., 2025). Inoltre anche il calore può andare a influenzare i processi di danneggiamento alla molecola desossiribonucleica. Il calore accelera pressoché tutte le reazioni chimiche fin qui descritte, agendo come un fattore trasversale piuttosto che come una via di danno a sé stante. L'innalzamento della temperatura aumenta la velocità di depurinazione e di deaminazione, favorisce la formazione di specie reattive dell'ossigeno, come dimostrato da Bruskov e collaboratori, e può indurre direttamente la denaturazione della doppia elica, esponendo il DNA a singolo filamento a un rischio di danno ancora maggiore rispetto alla forma *duplex* (Bruskov V.I et al., 2002).

1.5 Tessuti inclusi in paraffina (FFPE)

I tessuti fissati in formalina e inclusi in paraffina (FFPE), comunemente noti come tessuti paraffinati, rappresentano la più vasta collezione di materiale biologico umano disponibile, conservata per decenni negli archivi dei reparti di Anatomia Patologica e negli istituti di Medicina Legale. I campioni biologici inclusi in paraffina risultano essere ampiamente utilizzati nella diagnostica anatomo-patologica e medico-legale anche per studi di diagnostica molecolare e per studi retrospettivi di associazione *GWAS*, rivelandosi particolarmente utili negli studi clinici sui tumori (Gaffney *et al.* 2018). Questi campioni vengono trattati tramite fissazione per prevenire l'autolisi e la putrefazione, seguita dall'inclusione in paraffina per consentire il taglio di sezioni sottili destinate all'esame istologico. In ambito genetico forense, l'analisi del DNA estratto da tali tessuti è uno strumento molto utile come campione biologico di partenza per poi effettuare un'analisi di identificazione personale

in procedimenti penali o civili. Un ruolo di rilievo è ricoperto dai tessuti autoptici, che costituiscono una risorsa fondamentale per la medicina legale, la neuropatologia e l'epidemiologia molecolare, offrendo la possibilità di analizzare campioni che non sarebbero disponibili da altre fonti. Tuttavia, l'utilizzo di tessuti provenienti da autopsie presenta complessità tecniche superiori rispetto ai comuni campioni biotici: fattori come l'intervallo *post-mortem*, il tipo di fissativo e i tempi di fissazione prolungati (che spesso superano le 24-48 ore standard per i campioni di interesse clinico-diagnostico) influenzano drasticamente la qualità delle macromolecole. In particolare, nei tessuti paraffinati autoptici o clinici d'archivio si riscontra spesso una degradazione estesa del DNA, con frammenti che possono ridursi in termini di peso molecolare, limitando l'informatività dei profili genetici. Nonostante queste alterazioni, i blocchetti istologici di paraffina conservano il materiale genetico in modo tale da renderlo comunque analizzabile per scopi forensi e diagnostici, anche a distanza di molti anni (Stanta G., 2011).

1.5.1 L'effetto dei processi di fissazione e inclusione sulla molecola di DNA

Nel panorama della patologia forense e della medicina legale, dunque, i tessuti fissati in formalina ed inclusi in paraffina rappresentano una risorsa di inestimabile valore. Questi campioni, conservati spesso per decenni a temperatura ambiente, custodiscono informazioni genetiche cruciali per l'identificazione personale, per esempio in analisi retrospettive di *cold case*, e la determinazione di mutazioni patologiche correlate a morti improvvise o patologie oncologiche occulte (Panchal et al., 2020). Tuttavia, l'isolamento di acidi nucleici da questa tipologia di matrice costituisce una delle sfide più complesse per il genetista forense. La quantità e la qualità del DNA estratto sono intrinsecamente limitate dai processi chimico-fisici di fissazione e conservazione, rendendo necessaria la ricerca e la validazione di nuovi metodi di estrazione che ottimizzino la resa senza esacerbare il danno molecolare (Bauer, Pelisek, 2021). La formalina (soluzione acquosa di formaldeide solitamente al 10 %) agisce come fissativo ideale per la conservazione morfologica dei tessuti, ma si rivela un agente chimico altamente aggressivo nei confronti degli acidi nucleici. Dal punto di vista chimico, la formaldeide (CH_2O) reagisce elettrofilicamente con i gruppi amminici primari delle basi azotate, in particolare con due di queste che sono specialmente l'adenina e la citosina, portando alla formazione di basi di *Shiff* quindi di derivati instabili. Successivamente, questi intermedi reagiscono con altri gruppi amminici vicini, generando ponti metilenici stabili inter-molecolari DNA-DNA e DNA-proteina; questo processo di formazione di legami stabili è detto *crosslink* (Kennedy-darling, Smith, 2014) (Hoque et al., 2025). Oltre alla formazione di legami crociati, che intrappolano fisicamente l'acido

nucleico impedendone la denaturazione e l'aggancio da parte delle polimerasi, la formalina induce una severa frammentazione dell'acido nucleico. Un altro fattore critico consiste nell'evitare tempi di fissazione in formalina eccessivamente prolungati (>24–48 ore), i quali aumentano marcatamente il danno a carico del DNA. Anche l'impiego della formalina tamponata, in alternativa alla formalina non tamponata, esercita un impatto significativo sulla qualità del DNA. La formalina convenzionale presenta un carattere acido ($\text{pH} < 4$), che porta a un'intensa degradazione del DNA e a tassi più elevati di mutazioni, tra cui la deaminazione della citosina in uracile. La formalina tamponata (comunemente tamponata con fosfato, $\text{pH} \sim 7$) stabilizza l'ambiente di conservazione, limitando l'idrolisi e la frammentazione dell'acido nucleico (Lisman D. et al., 2025) (Kun Lu et al., 2010). Durante lo stoccaggio prolungato dei blocchetti, la formalina subisce un processo spontaneo di ossidazione che porta alla formazione di acido formico. L'abbassamento del pH all'interno della matrice tissutale innesca una reazione di idrolisi acida: questo fenomeno causa delle reazioni chimiche che vanno a intaccare la struttura dell'acido nucleico, in particolare avviene la depurinazione del DNA che, come detto in precedenza, causa una rottura del legame glicosidico tra la base purinica e il desossiribosio innescando una possibile successiva scissione idrolitica dello scheletro fosfodiesterico (Kazmi et al., 2023). Le condizioni acide, infatti, promuovono la formazione di forti legami crociati DNA-proteina, ostacolando la successiva estrazione e incrementando l'insorgenza di mutazioni spontanee del DNA (come le transizioni $\text{C} > \text{T}$), che possono interferire con le analisi genetiche (Koshiba M et al., 1993). Se la formalina è la causa principale del danno chimico, il processo di inclusione in paraffina introduce ulteriori variabili critiche. La paraffina è una miscela di idrocarburi solidi che agisce da matrice di supporto idrofobica. Per consentire la completa infiltrazione della paraffina nel tessuto, il campione deve subire una disidratazione progressiva in alcol e un passaggio in solventi chiarificanti come lo Xilene e suoi analoghi. Questo ambiente idrofobico isola il DNA, ma la fase di inclusione richiede temperature elevate, tipicamente comprese tra i 56°C e i 60°C . Questo stress termico accelera i processi di degradazione cinetica del DNA già parzialmente depurato o *cross-linkato* (Bauer, Pelisek, 2021). Inoltre, la paraffina residua che non viene rimossa efficacemente durante le fasi iniziali di estrazione agisce come un potente inibitore chimico della *PCR*, bloccando stericamente il sito attivo della Taq polimerasi. Il risultato finale è un DNA fortemente frammentato, il cui peso molecolare medio decresce progressivamente in funzione del tempo di fissazione in formalina (Soldati G. et al., 2026). Questo è molto rilevante nelle analisi genetico forensi poiché spesso queste avvengono su campioni biologici inclusi in paraffina i quali contengono frammenti di tessuti prelevati in sede autoptica. I campioni che vengono prelevati in autopsia, per poi essere inclusi al fine di fare le analisi istologiche, provengono da organi che non sono più vitali e più o meno interessati da

processi di autolisi cellulare e degradazione post-mortale. Questo quadro biologico rende l'isolamento e l'analisi del DNA significativamente più complessi e problematici qualora i tessuti inclusi in paraffina provengano da campionamenti autoptici piuttosto che da biopsie cliniche o chirurgiche. Nel caso di quest'ultime, il tessuto viene prelevato *in vivo* e processato in tempi molto brevi.

Il campione autoptico sconta inevitabilmente gli effetti del tempo intercorso tra il decesso e il prelievo (Post-Mortem Interval) e del tempo trascorso nel fissativo poiché in questo caso i tessuti non vengono processati in tempi brevi ma possono rimanere in formaline per molte ore (anche giorni e settimane) prima di subire la successiva disidratazione e inclusione in paraffina. Di conseguenza, nei tessuti provenienti da autopsia, la frammentazione chimica causata dalla formalina e dallo stress termico della paraffina non fa che sovrapporsi a una degradazione biologica spontanea, intrinseca negli organi già massivamente avviata, riducendo drasticamente la quota di DNA integro e amplificabile rispetto a un corrispettivo biptico.

Il materiale genetico che si estrae da tessuti fissati in formalina e inclusi in paraffina è degradato e frammentato per i motivi esplicitati in precedenza. La ridotta lunghezza dei frammenti di DNA limita ulteriormente l'efficacia dei metodi analitici, in particolare di quelli che si basano su ampliconi lunghi, come l'analisi dei marcatori *STR*, comunemente applicata nella genetica forense. La qualità del DNA nei tessuti *FFPE* è inoltre fortemente influenzata dal tempo e dalle condizioni di conservazione; i campioni che hanno subito un periodo di fissazione molto prolungato o campioni stoccati per molti anni mostrano spesso una degradazione addizionale, il che complica ulteriormente l'analisi (Zhang J. Et al., 2025). I tracciati elettroforetici ottenuti da questi campioni evidenziano profili genotipici parziali o ampiamente parziali. Tale esito è direttamente imputabile alla frammentazione e alla modificazione della struttura primaria del DNA, fenomeni intrinseci ai tessuti *FFPE* che riducono drasticamente il peso molecolare del materiale estratto. Di conseguenza, i loci con dimensioni molecolari più elevate risultano spesso non amplificabili. In Figura 14 sono rappresentate tre tipologie di tracciati elettroforetici: uno non degradato, uno parziale ottenuto da un DNA degradato e un tracciato ampiamente parziale ottenuto dalla tipizzazione di un DNA molto degradato.

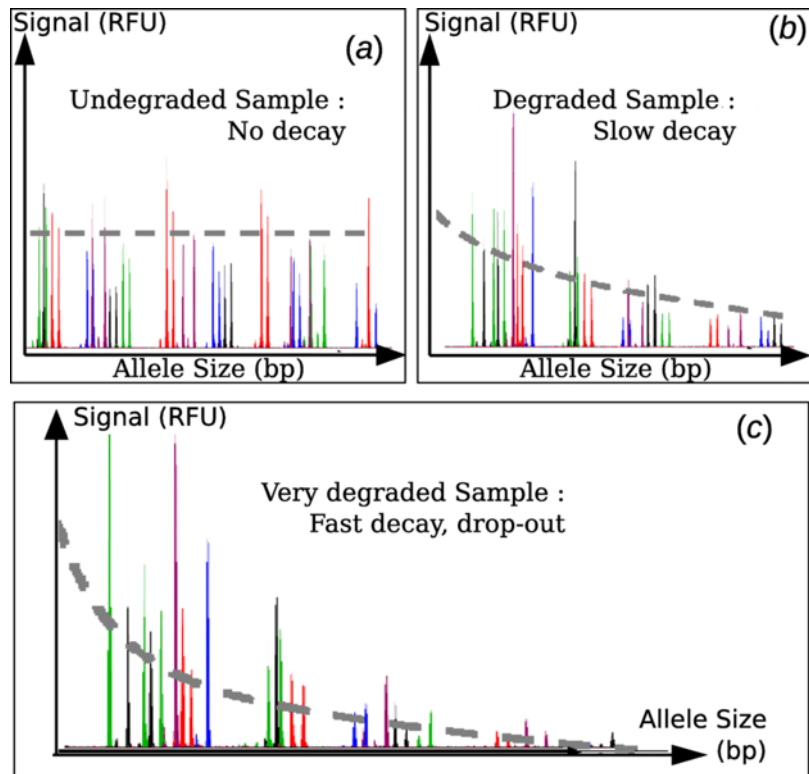


Figura 14 | Rappresentazione di tre tracciati elettroforetici. a) Tracciato elettroforetico di buona qualità proveniente da un campione non degradato. I picchi hanno altezza molto simili b) Tracciato elettroforetico che presenta un lieve decadimento di segnale (*RFU*) all'aumentare del peso molecolare degli alleli (verso destra), questo è tipico di un DNA degradato. c) Tracciato elettroforetico che presenta sia un rapido decadimento di segnale (*RFU*) all'aumentare del peso molecolare sia dei *Drop-out*, questo proviene da un campione di DNA molto degradato (Karkar S. et al., 2019) .

I profili genetici provenienti da DNA degradato presentano, infatti, problematiche connesse ad un aumentato della presenza di artefatti di amplificazione in grado di generare ambiguità nell'interpretazione dei risultati al fine identificativo. Alcuni dei tracciati elettroforetici ottenuti dall'analisi di DNA di bassa qualità, come quello degli *FFPE*, possono presentare, infatti, risultati tali da originare dubbi nell'attribuzione del profilo ad uno specifico soggetto. Ci sono cinque principali tipologie di fenomeni artefattuali che si possono generare durante il processo di amplificazione (Figura 15).

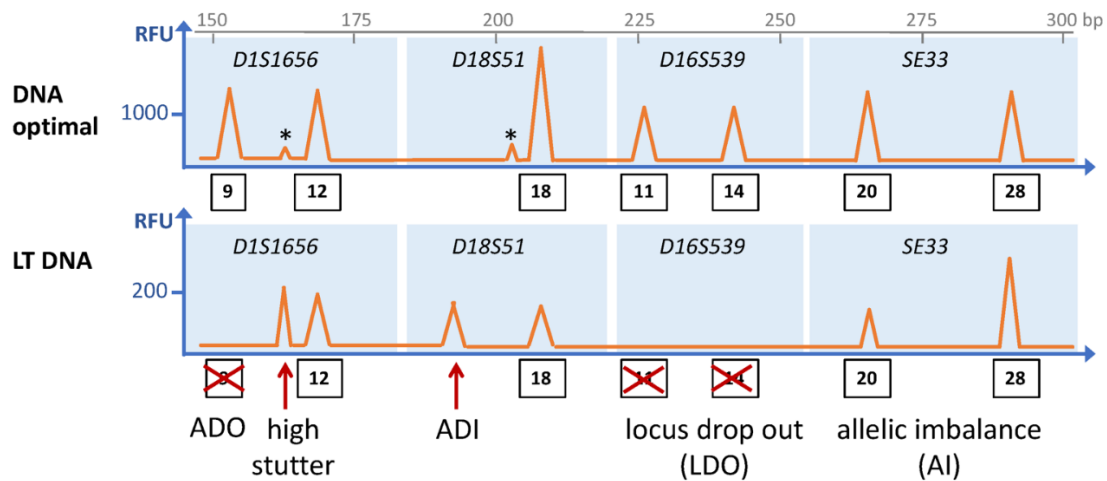


Figura 15 | Rappresentazione di *loci* con genotipo buona qualità con bilanciamento tra i due alleli quando eterozigote e la presenza di picchi stutter di dimensioni ridotte (inferiore al 10% secondo le linee guida) (parte superiore dell'immagine). Nella parte inferiore si possono vedere le diverse situazioni artefattuali: nel locus D1S1656 è presente un *Drop-out* allelico (ADO) e una banda *stutter* di dimensioni superiori alla soglia (*high stutter*). Nel locus D18S51 è rappresentato un *Drop-in* allelico (ADI). Nel locus D16S539 è presente un *Drop-out* dell'intero *locus* (LDO) e infine nel locus SE33 è presente uno sbilanciamento allelico (AI) (Jäger R., 2022).

- *Drop-out* allelico: caratterizzato dalla mancata amplificazione di un allele di un genotipo eterozigote la cui conseguenza è di scambiare un genotipo eterozigote (due alleli) per uno omozigote (un allele). La frequenza con cui avviene il fenomeno del *Drop out* è legata a molteplici fattori, quali il maggior peso molecolare dell'allele in questione e la possibile degradazione del campione di DNA (Figura 15) (Gill P. et al., 2000);
- Incremento dell'altezza del picco *stutter*: l'amplificazione dei marcatori utilizzati per la definizione del profilo genetico produce sistematicamente un artefatto di amplificazione definito "*stutter band*", nelle immediate vicinanze di un allele. In condizioni di DNA di bassa qualità, tale artefatto, normalmente presente in maniera moderata, può subire un incremento quantitativo nell'amplificazione, creando ambiguità sulla finale interpretazione del profilo (Figura 15) (Gill P. et al., 2000);
- *Drop-in* allelico: descritto dalla presenza di uno o più alleli aggiuntivi rispetto ad un genotipo principale (nella maggior parte dei casi). La conseguenza è di non riuscire ad identificare chiaramente il numero di contributori alla traccia e il relativo genotipo. Tali alleli extranumerari possono originarsi a seguito di contaminazione ovvero possono essere artefatti di amplificazione. Questo fenomeno artefattuale è più comune per i DNA in bassa quantità, il cosiddetto *LCN-DNA* (*Low copy number DNA*), rispetto al DNA degradato ma in quantità elevate come quello proveniente dagli *FFPE* (Figura 15) (Gill P., Buckleton J., 20120) ;

- *Locus Drop-out*: caratterizzato dalla mancata amplificazione di entrambi gli alleli di un determinato marcatore STR, la cui conseguenza è l'assenza totale di segnale elettroforetico per quel dato locus. Rappresenta la forma più estrema di fallimento dell'amplificazione e si verifica tipicamente in campioni caratterizzati da severa frammentazione del DNA o da bassissimo quantitativo iniziale. Poiché la perdita di segnale colpisce prevalentemente i loci aventi un peso molecolare maggiore, il fenomeno del locus drop-out comporta una perdita netta di informazioni genetiche, riducendo complessivamente il potere discriminativo del profilo ottenuto (Gill P. et al., 2000);
- Sbilanciamento allelico: in caso di amplificazione in condizioni ottimali, i picchi dei due alleli di un eterozigote presentano un'altezza simile; in condizioni di DNA degradato, vi può essere l'amplificazione preferenziale di uno dei due alleli dell'eterozigote con un rapporto di altezze tra i due picchi sbilanciato. In questo caso, vi possono essere dubbi circa l'identificazione del numero dei contributori, in special modo in una traccia commista (Figura 15) (Budowle B. et al., 2009).

La possibilità di riscontrare gli artefatti sopra descritti viene fortemente incrementata in presenza di DNA degradato. Nel caso in cui si verificano questi artefatti di amplificazione, nel profilo possono realizzarsi condizioni tali da insinuare il dubbio circa l'attribuibilità del profilo ottenuto dalla traccia all'indagato o, al contrario, da fare giungere a “falsi” giudizi di esclusione nelle comparazioni dei profili. Caratteristica peculiare di tali artefatti è quella di non presentarsi in modo sistematico nei singoli esperimenti ma casuale, seguendo criteri probabilistici definiti “stocastici”. In condizioni di DNA di qualità molto bassa, non vi è modo di prevederne la comparsa così come non può essere escluso -a priori- che ne sia avvenuto uno o più di uno.

2 SCOPO DEL LAVORO

I tessuti fissati in formalina ed inclusi in paraffina (*FFPE*), sebbene rappresentino una matrice biologica notoriamente complessa per le indagini molecolari, costituiscono frequentemente gli unici campioni biologici disponibili ai fini dell'identificazione personale in casi giudiziari. Se da un lato l'inclusione in paraffina assicura la conservazione morfologica dei tessuti nel tempo, dall'altro la combinazione dei processi di fissazione chimica e inclusione induce gravi alterazioni a carico della struttura del DNA. Tra le principali criticità, si annoverano un'estesa frammentazione meccanica, la formazione di legami crociati (*cross-linking*) covalenti intra- e inter-molecolari (DNA-DNA e DNA-proteina) e la presenza di inibitori della reazione di amplificazione in catena della polimerasi (*PCR*). Tali lesioni si traducono frequentemente in basse rese estrattive e nella generazione di profili genotipici (*STR*) incompleti o inficiati da artefatti analitici.

In questo contesto, il presente lavoro di tesi si propone in primo luogo di ottimizzare una miscela di estrazione del DNA già descritta in letteratura ed in uso presso il laboratorio di Genetica forense dell'Università di Pavia. L'obiettivo primario di questa prima fase sarà quello di perfezionare la ricetta originale, eliminando i componenti non essenziali, non in grado di fornire un contributo significativo all'incremento della resa totale di DNA dal campione.

Una volta ottimizzata la composizione della miscela di estrazione, verrà allestito uno studio comparativo condotto su campioni autoptici provenienti da differenti tipologie di tessuto, al fine di valutare l'efficienza di estrazione del DNA rispetto a un sistema automatizzato di riferimento commerciale, il *Maxwell® RSC DNA FFPE Kit*, specificamente progettato per il recupero di DNA da matrici paraffinate. Il confronto sarà focalizzato sulla determinazione della resa quantitativa totale di DNA estratto e sulla sua qualità, stimata attraverso l'indice di degradazione (*Degradation Index*).

Infine, la ricerca ha come obiettivo una valutazione della qualità dei profili genotipici (*STR*) ottenuti su un set rappresentativo di campioni estratti con entrambi i metodi. I tracciati elettroforetici verranno analizzati in dettaglio per comprendere l'impatto dei due approcci estrattivi sulla qualità finale della tipizzazione, prendendo in esame parametri quali la pendenza della curva di degradazione del DNA (*Degradation Slope*), il numero di marcatori completamente tipizzati e l'incidenza degli artefatti analitici correlati alla degradazione intrinseca delle matrici paraffinate.

3 MATERIALI E METODI

3.1 Descrizione dei campioni

Il presente elaborato di tesi ha preso in considerazione campioni biologici umani prelevati in sede autoptica da tre soggetti, due di sesso femminile e uno maschile. I prelievi relativi al soggetto di sesso maschile (FFPE 1) e a uno dei soggetti di sesso femminile (FFPE 3), entrambi deceduti in età compresa tra i diciotto e i vent'anni, sono stati eseguiti presso il settore dell'Unità di Medicina Legale e Scienze Forensi dell'Università degli Studi di Pavia. Il campione appartenente al secondo soggetto di sesso femminile (FFPE 2), deceduto all'età di 77 anni, è stato invece campionato durante l'esame autoptico condotto presso l'UCO di Medicina Legale dell'Università di Trieste. Le biocassette contenenti i tessuti paraffinati sono state preparate da personale medico, al fine di effettuare le necessarie valutazioni istopatologiche per la definizione della causa di morte. Per la prima fase degli esperimenti sono stati analizzati i tessuti cardiaci di tutti e tre i soggetti, mentre nelle fasi successive di studio per la comparazione dei tre tessuti, sono stati selezionati, oltre a quello cardiaco, anche tessuti encefalico e polmonare relativi al medesimo soggetto di sesso femminile sottoposto ad autopsia presso il settore della Medicina Legale di Pavia (FFPE 3).

Per quanto concerne i criteri di selezione dei campioni provenienti dal settore di Pavia (FFPE 1 e 3), i tessuti inclusi in paraffina utilizzati per questo studio sono provenienti da soggetti sottoposti ad autopsia nei 4-5 giorni successivi al decesso e i cui corpi non presentavano evidenti fenomeni cadaverici trasformativi. Per quanto riguarda il campione di Trieste (FFPE 2), l'autopsia è stata eseguita a 40 ore dal decesso, quindi ad un tempo *post-mortem* significativamente ridotto, rispetto a quello degli altri due campioni.

Il tempo di fissazione in formalina è stato tra 1 e 2 settimane per i campioni forensi (FFPE 1 e 3) e di 30 ore per il campione di interesse anatomico-patologico (FFPE 2). Anche in questo caso, il tempo di fissazione di tale campione era significativamente inferiore rispetto a quello degli altri due campioni. A partire dai tessuti inclusi nei blocchetti di paraffina, per tutte le analisi condotte, sono state tagliate sezioni istologiche di identico spessore (10 μm) utilizzate per l'estrazione del DNA impiegando i protocolli selezionati.

Per la prima fase dello studio, quella in cui è stata effettuata l'analisi solo sul tessuto cardiaco al fine di individuare la formulazione migliore della miscela per l'estrazione del DNA, sono stati utilizzati i seguenti campioni (Tabella 4):

Campione	Tessuto	Tempo di fissazione
FFPE 1	Cardiaco	n. settimane >1 e < 2
FFPE 2	Cardiaco	30 h
FFPE 3	Cardiaco	n. settimane >1 e < 2

Tabella 4| Campioni biologici utilizzati nella prima fase dello studio per andare ad individuare la formulazione migliore per la miscela per l'estrazione di DNA. Il tempo di permanenza in formalina, prima dell'inclusione in paraffina, è indicato nell'ultima colonna.

Per la seconda fase dello studio, invece, si è studiata l'eventuale presenza di un effetto tessuto cioè se tessuti differenti da quello cardiaco (i tessuti encefalico e polmonare) producessero risultati diversi in termini di quantità e qualità del DNA, a concentrazioni dei reagenti diverse da quelle ottimali stabilite per il tessuto cardiaco. A tale scopo, sono stati utilizzati i seguenti tessuti del soggetto FFPE 3 (Tabella 5):

Campione	Tessuto	Tempo di fissazione
FFPE 3	Cardiaco	n. settimane >1 e < 2
FFPE 3	Encefalico	n. settimane >1 e < 2
FFPE 3	Polmonare	n. settimane >1 e < 2

Tabella 5| Campioni biologici utilizzati nella seconda fase dello studio.

3.2 Progettazione degli esperimenti (*Design of experiments, DoE*) per lo studio del metodo di estrazione di DNA su tessuti *FFPE*

3.2.1 Obiettivi sperimentali e prove studiate

Al fine di identificare la composizione ottimale della soluzione di estrazione del DNA da tessuti inclusi in paraffina (*FFPE*), è stato adottato un approccio di *Design of Experiments (DoE)* in fase di screening a partire da una formula estrattiva (Higuchi R., 1989), con l'obiettivo di individuare, tra i componenti chimici e le condizioni di estrazione della miscela in uso presso il laboratorio di Genetica Forense per l'estrazione del DNA da tessuti inclusi in paraffina (KCl, Tris-HCl, MgCl₂, Gelatina, NP40, Tween20, Proteinasi K e tempo di denaturazione).

L'approccio di screening consente di individuare i parametri critici con un numero contenuto di prove sperimentali. Lo studio valuta simultaneamente l'impatto dei fattori su due variabili di risposta quantitative: la prima risposta è la misura della quantità totale di DNA estratto; la seconda risposta è l'indice di degradazione del DNA. Lo scopo dello studio è stato quello di massimizzare la quantità totale di DNA estratto per garantire elevate rese analitiche e, allo stesso tempo, minimizzare l'indice di degradazione del DNA per la caratterizzazione della qualità del campione.

3.2.2 Fattori di processo e loro dominio sperimentale

Il dominio o spazio sperimentale comprende otto fattori di processo derivati dalla formulazione attualmente in uso del laboratorio di Genetica Forense dalla pubblicazione di *Russel G. Higuchi* (Higuchi R., 1989) definita soluzione 13 "originale" nel seguito. I fattori studiati comprendono sette componenti chimici della miscela e l'ottavo fattore è una condizione fisica di processo. I fattori chimici analizzati sono KCl, Tris-HCl (pH 8.3), MgCl₂, Gelatina, NP40, Tween20 e Proteinasi K. Il fattore fisico di processo è il tempo di denaturazione termica (*Denaturation Time*, DT). Tutti i fattori sono stati esaminati a due livelli, codificati come livello minimo (-1) e livello massimo (+1). Gli intervalli di variazione sono stati scelti di una ampiezza tale che consentisse di rilevare in modo netto un effetto reale sulle risposte. La formulazione della soluzione 13, basata sul protocollo di *Higuchi*, è stata integrata nel piano come dato di riferimento di partenza, per il confronto di prestazioni delle diverse condizioni di estrazione studiate. Per tale ragione, la ricetta della soluzione 13 non corrisponde con il centro dello spazio fattoriale esplorato (Tabella 6).

#	Fattori (Unità di misura)	Abbreviazione	Ricetta Originale	Livello minimo (-1)	Livello massimo (+1)
1	KCl (mM)	KCl	50	10	100
2	Tris-HCl pH 8.3 (mM)	Tris	10	5	20
3	MgCl ₂ (mM)	Mg	2.5	0	3
4	Gelatina (mg/mL)	Gel	0.1	0	0.5
5	NP40 (%)	NP40	0.45	0	2
6	Tween20 (%)	Tw20	0.45	0	2
7	Proteinase K in H ₂ O (mg/mL)	PK	1	0.2	2
8	<i>Denaturation Time</i> (min)	DT	10	5	10

Tabella 6| Fattori di processo della soluzione di estrazione, ricetta originale (soluzione 13) e livelli definiti per il DoE.

3.2.3 Screening dei fattori con tessuto cardiaco

Per questa fase di *screening* è stato impiegato un disegno di *Plackett-Burman* a 12 corse, scelto per la sua efficienza nell'esplorare un numero elevato di fattori, in questo caso otto che sono i componenti della soluzione di estrazione, con un numero di esperimenti pianificati di numero contenuto. Il *Plackett-Burman* è un disegno fattoriale frazionato di risoluzione III ideato per valutare nel modo più efficiente in assoluto gli effetti dei fattori sulle risposte (Plackett, Burman, 1946). Il disegno adottato è stato analizzato assumendo un modello lineare privo di termini di interazione, coerentemente con il disegno di *Plackett-Burman* usato per costruire questo piano sperimentale. Ciascun fattore è stato testato a due soli livelli, codificati come -1 (minimo) e +1 (massimo), su intervalli di concentrazione e tempo elevati, scelti in modo da massimizzare la probabilità di rilevare un effetto reale qualora presente. Sono state inoltre impiegate delle variabili fittizie: e1, e2, e3 che non sono associate ad alcun fattore chimico reale (qui omesse per semplificare la lettura delle tabelle che illustrano i piani sperimentali). Dunque il loro effetto stimato, teoricamente nullo, fornisce una stima surrogata della varianza d'errore sperimentale, necessaria per valutare la significatività degli effetti principali in assenza di repliche. L'effetto stimato delle variabili fittizie misura la varianza dell'errore sperimentale residuo in assenza di

repliche delle prove sperimentali. Tale stima permette di valutare la significatività statistica degli effetti confrontando direttamente la magnitudo dei coefficienti dei fattori di processo con quelli dei fattori fittizi. Solo i coefficienti di fattori di processo più grandi numericamente dei coefficienti dei fattori fittizi sono considerati “significativi” o, per meglio dire, “importanti” ai fini dello studio. La sperimentazione iniziale è stata condotta su tessuto cardiaco includendo la soluzione 13 come riferimento (Tabella 7).

Exp#	KCl (mM)	Tris (mM)	Mg (mM)	Gel (mg/mL)	NP40 (%)	Tw20 (%)	PK (mg/mL)	DT (min)
1	100	20	0	0.5	2	2	0.2	5
2	10	20	3	0	2	2	2	5
3	100	5	3	0.5	0	2	2	10
4	10	20	0	0.5	2	0	2	10
5	10	5	3	0	2	2	0.2	10
6	10	5	0	0.5	0	2	2	5
7	100	5	0	0	2	0	2	10
8	100	20	0	0	0	2	0.2	10
9	100	20	3	0	0	0	2	5
10	10	20	3	0.5	0	0	0.2	10
11	100	5	3	0.5	2	0	0.2	5
12	10	5	0	0	0	0	0.2	5
13 (originale)	50	10	2.5	0.1	0.45	0.45	1	10

Tabella 7| Piano sperimentale di Plackett-Burman per lo screening dei fattori condotto su tessuto cardiaco. Il piano comprende 12 esperimenti con formulazioni di cui non è noto a priori il rendimento e dalla formulazione di riferimento, soluzione 13 (Exp# 13).

3.2.4 Estensione delle prove di estrazione a più tessuti

Nella fase successiva a quella delle analisi iniziali solo sui tessuti cardiaci seguendo il *design* sperimentale riportato in Tabella 7, ne è stato allestito uno nuovo. La seconda fase di prove ha esteso la valutazione ad altri tessuti per verificare la reale efficacia e validità generale della miscela di estrazione individuata come migliore su tessuto cardiaco. A tale scopo sono stati sottoposti ad estrazione tessuto cardiaco, encefalico e polmonare e si sono confrontati i risultati ottenuti in termini di DNA totale estratto e indice di degradazione. Questo secondo *design* sperimentale si basa sullo stesso disegno del primo, quello di *Plackett-Burman* a 12 esperimenti. Questo secondo modello sperimentale denominato “effetto tessuto” è stato ideato al fine di verificare se la miscela risultata migliore nella prima fase, utilizzando solo tessuti cardiaci, risultasse tale anche applicata ad altri due tessuti, quello encefalico e polmonare o se da questa analisi su altri tessuti portasse alla modifica dei componenti della miscela di estrazione finale (Tabella 7). L'obiettivo è stato quindi quello di trovare una miscela di estrazione “ottimale” cercando la miscela che performa meglio "in media" su tutti e tre i tessuti provenienti da organi diversi. In questo modello è stato quindi aggiunta la variabile qualitativa "tessuto", tenendo come livello di riferimento quello del tessuto cardiaco analizzato singolarmente in precedenza nella prima fase. Le 12 prove previste sono state ripartite tra i tre organi usando il criterio di D-ottimalità (Leardi, 2018): 3 prove su tessuto cardiaco (C), 4 su tessuto encefalico (E) e 5 su tessuto polmonare (P). Inoltre, in questo *design* “effetto tessuto” sono state aggiunte tre estrazioni aggiuntive alle dodici del modello, per le prove di riferimento utilizzando la miscela originaria (13 Tabella 6), una per ciascun tessuto nelle medesime condizioni.

Exp#	KCl (mM)	Tris (mM)	Mg (mM)	Gel (mg/mL)	NP40 (%)	Tw20 (%)	PK (mg/mL)	DT (min)	Tessuto
1	10	5	0	0.5	0	2	2	5	C
2	100	5	0	0	2	0	2	10	C
3	10	20	3	0.5	0	0	0.2	10	C
4	100	20	0	0.5	2	2	0.2	5	E
5	100	5	3	0.5	0	2	2	10	E
6	10	5	3	0	2	2	0.2	10	E
7	100	20	3	0	0	0	2	5	E
8	10	20	3	0	2	2	2	5	P
9	10	20	0	0.5	2	0	2	10	P
10	100	20	0	0	0	2	0.2	10	P
11	100	5	3	0.5	2	0	0.2	5	P
12	10	5	0	0	0	0	0.2	5	P
13C	50	10	2.5	0.1	0.45	0.45	1	10	C
13E	50	10	2.5	0.1	0.45	0.45	1	10	E
13P	50	10	2.5	0.1	0.45	0.45	1	10	P

Tabella 8| Piano sperimentale di Plackett-Burman per studiare l'effetto del tessuto sulla resa di estrazione. Il piano è costituito da 12 esperimenti, distribuiti per valutare l'efficienza di estrazione del DNA da tre tessuti biologici (Exp# 1-12). A queste prove sono state aggiunti 3 esperimenti condotti con la formulazione originale (soluzione 13) su ciascun tessuto considerato (Exp# 13-15).

3.3 Estrazione del DNA

3.3.1 Estrazione manuale da tessuti inclusi in paraffina

Per l'isolamento del materiale genetico a partire da campioni di tessuto fissati in formalina e inclusi in paraffina (*FFPE*), nello specifico sezioni di miocardio, encefalo e parenchima polmonare, è stato impiegato un protocollo di lisi chimico-enzimatica basato su un tampone di estrazione “originale” che era già in uso in laboratorio e presente in letteratura (Higuchi R., 1989). Questo metodo consiste in una procedura di estrazione grezza, in quanto il DNA non viene purificato ma, semplicemente, liberato nella soluzione. La componente proteica è degradata enzimaticamente dalla Proteinasi K e denaturata termicamente. Dopo centrifugazione, la paraffina solidifica formando un una sorta di tappo sulla soluzione acquosa di lisato grezzo. Una volta bucato il tappo, il lisato è trasferito in una provetta. Il vantaggio di questa procedura è che è una “*single-tube extraction*”, con ciò significando che il

campione rimane sempre nella medesima provetta fino al suo trasferimento finale, minimizzando i trasferimenti tra provette e garantendo così l'originalità del campione. Per procedere con l'estrazione di DNA dai campioni di tessuto inclusi in paraffina, per ciascun campione è stata tagliata una sezione di 10 µm al microtomo dal blocchetto contenente il frammento autoptico e inserite in una provetta *Eppendorf* da 1.5 mL. Ai campioni sono stati in seguito aggiunti 150 µL della seguente soluzione i cui reagenti sono alla concentrazione finale seguente:

- 50 mM KCl;
- 10 mM Tris-HCl (pH 8.3);
- 2.5 mM MgCl₂;
- 0.1 mg/mL Gelatina;
- 0.45 % NP-40;
- 0.45 % Tween-20;
- 1 mg/mL Proteinasi K (in H₂O).

La soluzione è quindi costituita da una base salina comprendente cloruro di potassio (KCl) e per il mantenimento del potere tamponante è presente il Tris-HCl (pH8.3). E' presente anche il magnesio cloruro (MgCl₂) che, al tempo della messa a punto della miscela originale serviva come cofattore della *Taq* Polimerasi durante la reazione di *PCR*; ciò poiché un'aliquota dell'estratto grezzo veniva direttamente inserita nella miscela di amplificazione per l'esecuzione della *PCR*. La gelatina alla soluzione è stata addizionata con funzione stabilizzante e protettiva per gli acidi nucleici. L'NP40 e il Tween 20 sono due detergenti non ionici che, in sinergia, contribuiscono alla disgregazione delle membrane cellulari e la solubilizzazione dei complessi lipo-proteici. Infine, la degradazione strutturale del tessuto e delle proteine istoniche è stata garantita dall'aggiunta della Proteinasi K. Dopo aver aggiunto tale enzima, i campioni sono stati messi in lisi in un termostato a 56° C per 4 ore. A seguito del periodo di incubazione necessario per la digestione enzimatica, il protocollo ha previsto uno step di denaturazione termica finale a 95° C per 10 minuti per inattivare l'enzima proteolitico. I campioni sono stati in seguito centrifugati a 13000 *rpm* per 5 minuti e trasferiti in una nuova provetta *Eppendorf*, prestando attenzione a non trasferire eventuali impurità presenti sul fondo della provetta e a non trasferire paraffina. Il lisato così ottenuto è stato successivamente destinato alla fase di quantificazione in *Real Time PCR* per valutare la resa finale di estrazione. Uno schema dell'intero procedimento è rappresentato in Figura 16.

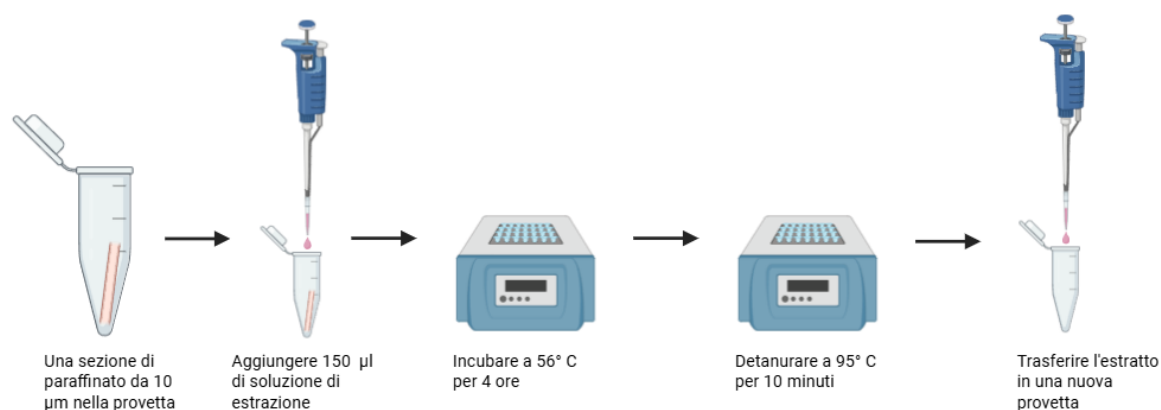


Figura 16 | Schema del procedimento di estrazione del DNA manuale da una sezione di tessuto incluso in paraffina (© Riccardo Ciapparelli, BioRender.com).

La formulazione sopra riportata corrisponde a quella “originale” presente nella pubblicazione di Higuchi del 1989 ed è stata riportata al fine di esplicitare il ruolo di ogni componente. Tuttavia, nel corso dello studio e sulla base del disegno degli esperimenti, sono state utilizzate diverse formulazioni riportate nelle precedenti tabelle 7 e 8, testate con le medesime modalità della miscela originale. Un controllo negativo di estrazione, finalizzato ad escludere contaminazioni durante il processo di estrazione, è stato allestito per ogni set di esperimenti.

3.3.2 Estrazione automatizzata da tessuti inclusi in paraffina

L'estrazione di DNA da tessuto cardiaco, encefalico e polmonare è stata effettuata anche con un processo estrattivo automatizzato. Per svolgere questo protocollo estrattivo è stato utilizzato un kit commerciale *Maxwell® RSC DNA FFPE kit*. Il kit viene utilizzato con gli strumenti *Maxwell® RSC* per offrire un metodo semplice ed efficiente per la purificazione automatizzata del DNA genomico da campioni di tessuto fissati in formalina e inclusi in paraffina (*FFPE*). Il kit non richiede l'uso di solventi organici pericolosi, come lo xilene, garantendo un protocollo di estrazione del DNA da *FFPE* più sicuro rispetto ad altri metodi comuni. Questo kit prevede l'utilizzo di cartucce pre confezionate *Maxwell® FFPE Cartridges*. Una cartuccia è composta da otto pozzetti, all'interno dei singoli pozzetti di ciascuna cartuccia sono presenti reagenti liquidi pre-dosati e particelle paramagnetiche che si legano agli acidi nucleici e li catturano. Lo strumento *Maxwell® RSC* impiegato può estrarre da 1 a 16 campioni, in una singola corsa (Figura 17). Gli strumenti *Maxwell® RSC* vengono forniti

con procedure di purificazione pre-programmate e sono progettati per l'uso con cartucce di reagenti già processati, massimizzando la semplicità e la praticità. Il DNA purificato è pronto per l'uso diretto in una varietà di applicazioni a valle basate sull'amplificazione.



Figura 17| Rappresentazione dello strumento di estrazione automatizzata *Maxwell® RSC 16* utilizzando per svolgere le estrazioni (ita.promega.com).

La fase di estrazione del DNA, come sopra riportato, avviene in maniera automatizzata nello strumento utilizzando le cartucce contenute nel kit. Prima però di caricare i campioni nelle cartucce e quindi sullo strumento, le sezioni di paraffinato contenenti il tessuto devono subire una lisi al fine di digerire tutta la componente tissutale. Per ciascun campione è stata tagliata una sezione di 10 μm al microtomo dal blocchetto contenente il frammento autoptico e inserite in una provetta *Eppendorf* da 1.5 mL. A questo punto si procede con la fase di estrazione del DNA dal campione seguendo il protocollo del kit *Maxwell® RSC DNA FFPE kit* riportato nel manuale tecnico.

- Aggiungere 300 μl di olio minerale (contenuto nel kit) nella provetta contenente la sezione di paraffinato;
- Inserire la provetta in un termostato a 80° C per 2 minuti;
- Successivamente lasciare la provetta a temperatura ambiente durante il tempo della preparazione della soluzione di lisi;
- Preparare la soluzione di lisi contenente: 224 μl di *Lysis Buffer*, 25 μl di Proteinasi K e 1 μl di *Blue Dye* per ciascun campione in studio (tutti contenuti nel kit). Preparare la soluzione di lisi per 1 o 2 campioni in più rispetto a quelli di cui si vuole procedere con l'estrazione;
- Aggiungere 250 μl di soluzione di lisi per ciascun campione e vortexare 5 secondi;
- Centrifugare a 10 000 g per 20 secondi, se presente un pellet nella fase acquosa (di colore azzurro) mixare gentilmente con una pipetta per disgregarlo;

- Trasferire le provette in termostato a 56° C per 30 minuti;
- Trasferire le provette in un termostato a 80° C per 4 ore;
- Rimuovere le provette e lasciarle a temperatura ambiente per 5 minuti;
- Aggiungere 10 µl di RNasi alla fase acquosa di ciascun campione e mischiare pipettando;
- Incubare a temperatura ambiente per 5 minuti;
- Centrifugare alla massima velocità per 5 minuti.

Posizionare le cartucce da utilizzare nei vassoi di alloggiamento, orientando il pozzetto n. 1 dalla parte opposta rispetto ai tubi di eluizione. Premere sulla cartuccia per incastrarla in posizione. Rimuovere delicatamente la pellicola sigillante, assicurandosi che tutta la plastica venga via dalla parte superiore della cartuccia. Posizionare un tubo di eluizione vuoto nell'apposito alloggiamento per ciascuna cartuccia presente nei vassoi. Aggiungere 50 µl di acqua priva di nucleasi (*Nuclease-Free Water*) (fornita all'interno del kit) sul fondo di ciascun tubo di eluizione.

Successivamente tutta la fase acquosa, di colore azzurro grazie alla presenza del colorante *Blue Dye*, deve essere trasferita all'interno del primo pozzetto della cartuccia. Si procede poi ad inserire un plunger nel pozzetto n. 8 di ciascuna cartuccia (Figura 18).

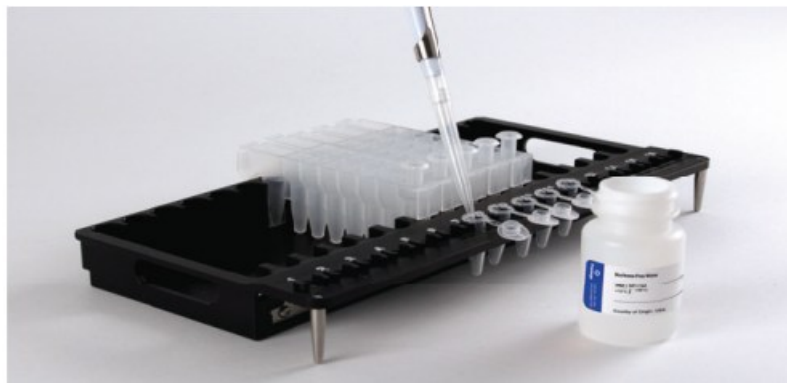


Figura 18 | Immagine rappresentativa del vassoio del *Maxwell® RSC 16* con le cartucce posizionate, i plungers inseriti e i tubi di eluizione disposti correttamente (*Maxwell® RSC DNA FFPE kit* Technical manual).

A questo punto si inserisce il vassoio con le cartucce, i plungers e i tubi di eluizione all'interno del *Maxwell® RSC 16* e si procede con l'impostare il kit utilizzato tramite lo scan del QR code presente sulla scatola del kit al fine di far partire il protocollo di estrazione automatico. Questo procedimento è stato eseguito anche per un controllo negativo di estrazione, in quanto è importante accertarsi che non ci siano state contaminazioni durante il processo di estrazione.

3.4 Quantificazione del DNA

Quando le prove biologiche di una scena del crimine vengono processate per isolarne il DNA, tutte le fonti di DNA vengono estratte e risultano presenti nei campioni da esaminare. Pertanto, nel DNA totale recuperato dal campione, oltre al DNA umano di interesse, potrebbe esserci anche DNA non umano, come materiale batterico, fungino, vegetale o animale. Per garantire che il DNA recuperato da un'estrazione sia effettivamente umano e non provenga da altre fonti come i batteri, gli Standard di Assicurazione Qualità dell'*FBI* (Standard 9.4, QAS 2009) richiedono una quantificazione del DNA specificamente umana. La quantità e la qualità del DNA in un campione possono essere valutate in modo affidabile solo dopo che quest'ultimo è stato isolato. Lo scopo principale della quantificazione del DNA nei casi giudiziari forensi è determinare la quantità di materiale genetico presente nel campione per procedere nel migliore dei modi per effettuare le analisi successive (Nicklas & Buel, 2003). In particolare questo passaggio è fondamentale per avere una appropriata quantità di DNA template da includere nell'amplificazione, tramite *PCR*, dei *loci STR*, al fine di evitare dati fuori scala e i relativi artefatti. La determinazione della quantità di DNA in un campione è essenziale per la maggior parte dei saggi basati su *PCR*, poiché solo un intervallo ristretto di concentrazione offre i risultati migliori con la tipizzazione *multiplex* dei marcatori. Con gli attuali kit commerciali per *STR*, l'input ottimale di DNA umano è tipicamente compreso tra 0.5 ng e 1.0 ng. L'impiego di una quantità di DNA *input* errata può compromettere il risultato delle analisi: un eccesso di DNA stampo può portare all'ottenimento di una quantità di prodotti amplificati troppo elevata con artefatti ed effetti stocastici che potrebbero compromettere la corretta interpretazione dei profili, viceversa, una ridotta quantità di DNA *input* può produrre una quantità insufficiente di amplificati che può portare all'ottenimento di profili parziali con la perdita di alleli (Figura 19) (Butler, 2010). Se al momento della quantificazione, la concentrazione di DNA è superiore o inferiore al *range* ottimale da inserire per la reazione di amplificazione, è necessario rispettivamente diluire il campione (se troppo concentrato) o concentrarlo (se troppo poco concentrato) rimuovendo fluido in eccesso. La diluizione può essere effettuata semplicemente aggiungendo acqua all'estratto per diminuirne la concentrazione finale, mentre per concentrare il campione e quindi rimuovere liquido in eccesso può essere usato un dispositivo che sfrutta l'ultrafiltrazione per centrifugazione. Un esempio di filtro centrifugo può essere il *Microcon® DNA Fast Flow*.

È un filtro centrifugo a membrana con un *cut-off* di peso molecolare (*MWCO*) definito, tipicamente 30 kDa per le applicazioni forensi sul DNA. La membrana trattiene le molecole grandi (DNA genomico) e lascia passare quelle piccole (sali, inibitori, primer, piccoli frammenti).

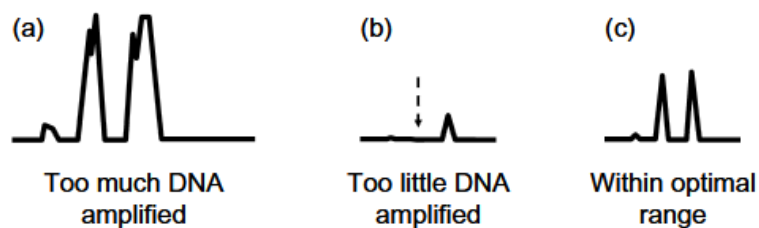


Figura 19 | Illustrazione dei risultati della tipizzazione *STR* in un singolo locus eterozigote per un campione a singolo contributore con: (a) una quantità eccessiva di DNA templatato, che mostra picchi fuori scala e sdoppiati ; (b) una quantità insufficiente di DNA templatato, in cui la freccia indica la perdita di un allele (*allele dropout*) dovuta a effetti stocastici; oppure (c) la giusta quantità, tale per cui i due picchi allelici risultano bilanciati e all'interno della scala (strbase.nist.gov).

3.4.1 Quantificazione tramite *Real Time PCR*

Lo scopo principale dell'esecuzione di un test di quantificazione del DNA è determinare la quantità di DNA "amplificabile". Una reazione di amplificazione *PCR* può fallire a causa della presenza di inibitori co-estratti, di DNA altamente degradato, di una quantità di DNA insufficiente o di una combinazione di tutti questi fattori. Di conseguenza, un test in grado di riflettere accuratamente sia la qualità sia la quantità del DNA templatato presente in un campione estratto è di grande aiuto per decidere come procedere. I saggi di *PCR* "in tempo reale" offrono una simile valutazione sia della resa sia della purezza del DNA ai fini dell'amplificazione. Oggi sono disponibili strumenti e saggi in grado di monitorare il processo di *PCR* durante il suo svolgimento, consentendo una raccolta dati "in tempo reale". La *PCR* quantitativa in tempo reale (*qPCR*) descrittiva per la prima volta da Higuchi e collaboratori presso la *Cetus Corporation* nei primi anni '90 (Higuchi et al. 1992, 1993). La *Real Time PCR* è una tecnica di biologia molecolare che si basa sulla reazione a catena della polimerasi ma che, al contrario della tecnica classica, monitora l'amplificazione della molecola *target* durante la fase esponenziale della reazione e non alla fine; infatti, durante la reazione viene rilevata la fluorescenza emessa da determinati costrutti molecolari (sonde), indice della produzione di ampliconi a ogni ciclo di *PCR*. Questa tecnica sfrutta l'attività 5'- 3' esonucleasica della *Taq* polimerasi per clivare una sonda oligonucleotidica fluorogena di ibridazione doppiamente marcata (*TaqMan*®) durante la fase di estensione della reazione.

La sonda *TaqMan*® è marcata a un'estremità con un fluoroforo che ha la funzione di *reporter*, mentre all'altra estremità da un secondo fluoroforo che ha la funzione di *quencher*; quando la sonda è integra, si ha un trasferimento di energia fluorescente, in quanto la fluorescenza emessa dal *reporter* viene assorbita dal *quencher*, risultando nell'annullamento del segnale fluorescente emesso dalla sonda

(Heid et al. 1996). Inoltre, la sua estremità 3' non può essere estesa, in modo tale che non possa fungere da *primer* (Lee, 1993) (Arya et al. 2005). L'azione esonucleasica della polimerasi durante la fase di estensione della *PCR* degrada la sonda, allontanando così i due marcatori e ponendo fine all'azione di repressione della fluorescenza del *quencher* sul *reporter*: ciò risulta in un picco nello spettro di fluorescenza emessa dal *reporter* che viene in seguito rilevato per la quantificazione della molecola (Heid et al. 1996), come rappresentato in Figura 20.

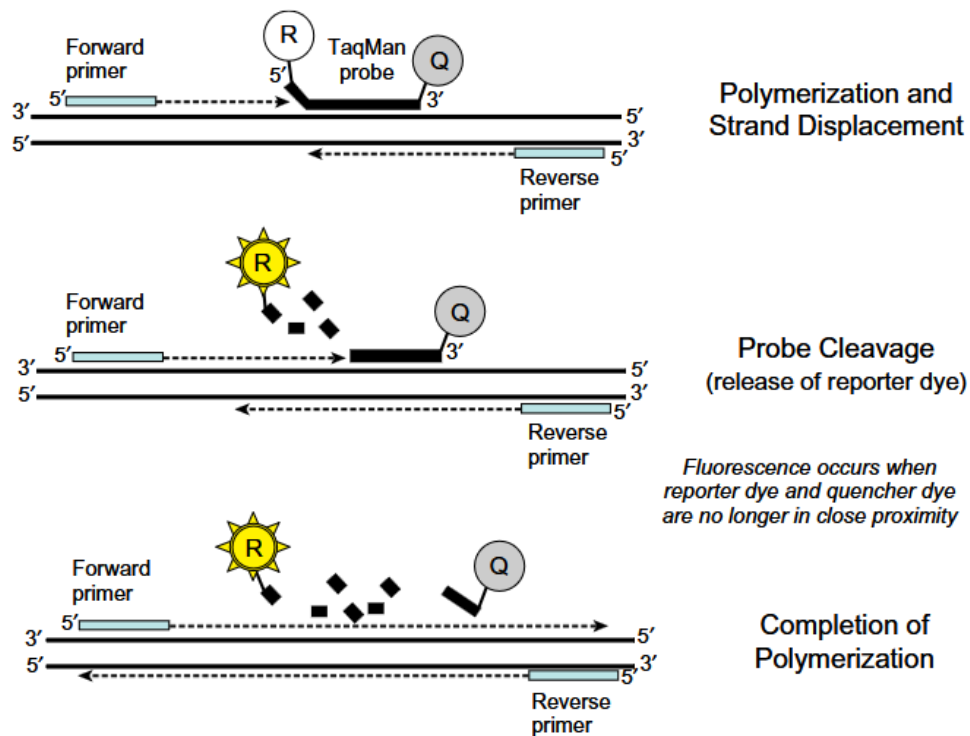


Figura 20 | Schema illustrativo della funzione di una sonda *TaqMan*® con R = *reporter* e Q = *quencher* (Butler, 2011).

All'interno della *master mix* della *PCR* si trova un marcatore passivo di referenza, la cui intensità di fluorescenza emessa rimane pressoché invariata durante la reazione ed è quindi utilizzato come standard interno per normalizzare le variazioni nella fluorescenza del marcatore *reporter* secondo la seguente formula: $\Delta Rn = (Rn^+) - (Rn^-)$ dove Rn^+ (Rn : *report* normalizzato) indica l'intensità di emissione del *reporter*/l'intensità di emissione del marcatore passivo in qualsiasi momento della reazione, mentre Rn^- (Rn : *report* normalizzato) indica l'intensità di emissione del *reporter*/l'intensità di emissione del marcatore passivo misurate prima che inizi la reazione di amplificazione. Questo calcolo viene effettuato a ogni ciclo di *PCR* e i risultati vengono plottati in un grafico di amplificazione su scala logaritmica: il valore medio per ciclo di ΔRn viene rappresentato

graficamente sull'asse y, mentre sull'asse x vengono indicati i cicli della reazione (Heid et al. 1996). Nel processo della *qPCR* si possono individuare tre fasi distinte (Figura 21):

- Fase esponenziale dove avviene la vera e propria reazione di amplificazione ed è rappresentata con la parte di curva che va dalla rilevazione del segnale e la fase lineare. L'aumento del segnale fluorescente emesso dal *reporter* della sonda è direttamente proporzionale ai prodotti di *PCR* che si formano, più prodotto di amplificazione è presente nella reazione più sarà intenso il segnale di fluorescenza. L'amplificazione esponenziale è descritta nella seguente relazione: $X_n = X_0(1 + E_x)^n$ dove X_n indica il prodotto di *PCR* al ciclo n , X_0 indica il numero di copie iniziali, E_x indica l'efficienza di amplificazione. Con un'efficienza teorica assoluta ad ogni ciclo di *Real Time PCR* la quantità di molecola prodotta viene duplicata. Il ripetersi di numerosi cicli di amplificazione provoca l'aumento della quantità di DNA consentendo, per ciascun campione, il raggiungimento di un ciclo, definito Ciclo Soglia o *Threshold Cycle (CT)*, a partire dal quale l'amplificazione esponenziale viene rivelata dallo strumento. Grazie al valore del *CT* che si ottiene è quindi possibile stimare la quantità di DNA presente nel campione. Ovviamente, maggiore è la quantità di DNA minore sarà il *CT* ovvero minore sarà il numero di cicli di amplificazione necessari per consentire alla fluorescenza emessa dai *reporter* di superare il valore soglia. Infatti, la tipologia di relazione che si deduce tra concentrazione di DNA e valore soglia è inversamente proporzionale, all'aumentare dell'una si osserva una riduzione dell'altra e viceversa (Figura 21).
- Fase lineare durante la quale la pendenza della curva diminuisce. Ciò indica che la concentrazione dei reagenti utilizzati nella reazione è drasticamente diminuita, essendo indirettamente proporzionale alla quantità di molecola target prodotta, diminuendo così l'efficienza della reazione.
- Fase di *plateau* in cui il segnale ΔRn e la quantità di amplificato rimangono costanti, coincide con la fine della reazione.

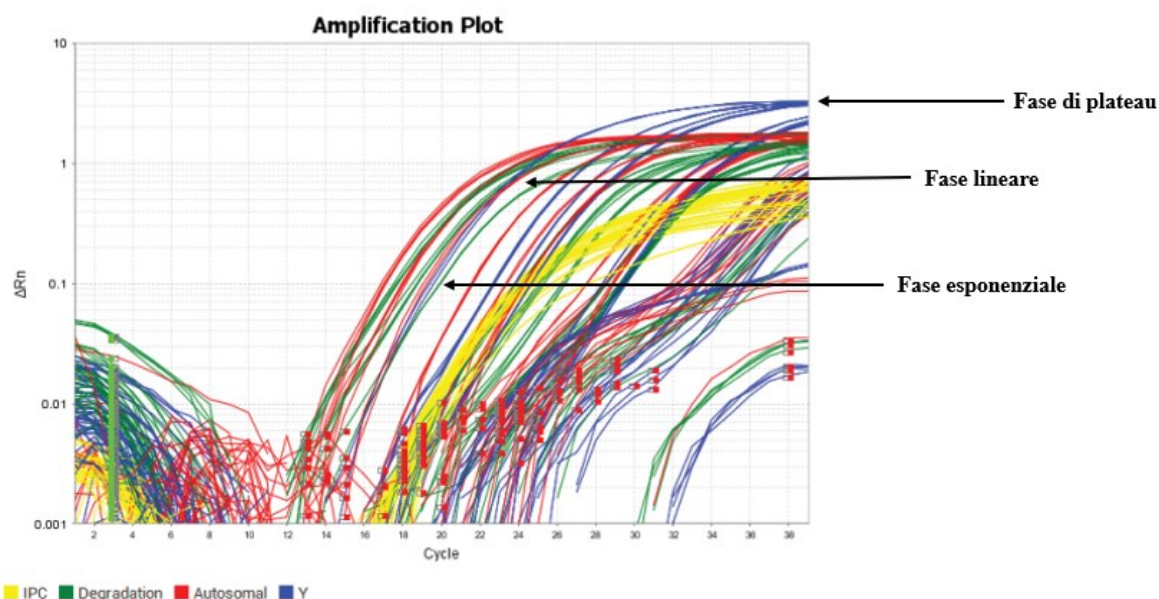


Figura 21| Grafico di amplificazione *Real Time PCR*, con ΔRn sull'asse y e il numero di cicli sull'asse x. Sono illustrate anche le diverse fasi del grafico: la fase esponenziale, la fase lineare e la fase di *plateau* (modificata da *PowerQuant® System, Promega*; ita.promega.com).

I campioni di DNA estratto sono stati quantificati mediante il kit *PowerQuant® System*, Promega utilizzando lo strumento *7500 fast Real-Time Applied Biosystems®*, seguendo il protocollo *PowerQuant® System* - Promega. Il sistema *Power Quant™* è una *multiplex qPCR* a cinque fluorocromi e quattro sequenze di DNA bersaglio che utilizza la tecnologia delle sonde *TaqMan*. Il kit contiene quattro diverse componenti. La prima è la *PowerQuant® 20X Primer/Probe/IPC Mix* che include appunto i *primers*, le sonde, il controllo interno positivo della *PCR (IPC)* e il marcatore passivo di riferimento.

- *Primers* e sonda per il *target* del DNA autosomico: la sonda marcata con fluorocromo *FAM* del sistema *PowerQuant®* rileva un target multi copia di DNA autosomico umano. I *primer* vengono utilizzati per amplificare un amplicone di 84 bp. I dati derivanti da questa reazione quantificano la quantità totale di DNA umano presente in un campione. Il *target* autosomico è corto, resistente agli inibitori e ha meno probabilità di risentire di eventi di degradazione rispetto al *target* di degradazione, che è più lungo;
- *Primers* e sonda per il *target* del cromosoma Y: la sonda marcata con fluorocromo *CAL Fluor® Gold 540* del sistema *PowerQuant®* rileva la presenza di DNA del cromosoma Y. I *primer* vengono utilizzati per amplificare due *loci* multi copia (due ampliconi, uno di 81 pb e l'altro di 136 pb). L'impiego di due *loci* multi copia riduce al minimo l'effetto che la variazione del numero di copie di un determinato marcatore può avere sul rapporto $[Auto]/[Y]$ e aumenta la sensibilità per il DNA maschile;

- *Primers* e sonda per il *target* di degradazione: la sonda marcata con fluorocromo *Quasar*® 670 del sistema *PowerQuant*® rileva la presenza di un amplicone più lungo di 294 bp derivato da una regione diversa dello stesso *locus* del *target* autosomico. A causa della sua maggiore lunghezza, l'amplicone di degradazione è più suscettibile alla degradazione e alla presenza di inibitori. Il rapporto tra le concentrazioni di DNA determinate con il *target* autosomico e quello di degradazione (rapporto [Auto]/[Deg]) può essere utilizzato per valutare il grado di degradazione;
- *Primers*, sonda e template per il Controllo Interno della *PCR* (*IPC*): la sonda marcata con fluorocromo *TMR* del sistema *PowerQuant*® rileva il Controllo Interno della *PCR*, un template di DNA ingegnerizzato *ex novo* che viene incluso in ogni reazione di amplificazione. I *primers* generano un prodotto amplificato di 435 pb. L'efficienza di amplificazione dell'*IPC* viene utilizzata per rilevare la presenza di inibitori nel campione. Questo è il *target* più lungo del sistema *PowerQuant*®, il che rende l'*IPC* più suscettibile agli inibitori rispetto agli altri *target* della *multiplex*;
- Colorante di riferimento passivo: il fluorocromo *CXR* del sistema *PowerQuant*® viene utilizzato come riferimento passivo. Il colorante *CXR* è incluso in ogni reazione di amplificazione e i dati provenienti dagli altri canali di fluorescenza vengono normalizzati rispetto a questo segnale (*PowerQuant*® *System Technical Manual*, Promega) (ita.promega.com);
- La seconda componente del kit è la soluzione *PowerQuant*® *2X Master Mix* che offre prestazioni paragonabili a quelle dei sistemi *STR* più recenti e utilizza la tecnologia *PCR hot-start*. L'allestimento della reazione può essere eseguito a temperatura ambiente ed è predisposto per l'automazione. L'amplificazione si completa in 1 ora circa;
- La terza e la quarta componente sono il *PowerQuant*® *Male gDNA standard* e il *PowerQuant*® *Dilution Buffer*. Il *PowerQuant*® *Male gDNA standard* è un DNA a concentrazione nota di 50 ng/μl e il *PowerQuant*® *Dilution Buffer* è la soluzione con cui vengono fatte le diluizioni per ottenere diverse concentrazioni finali. La quantificazione concentrazione di DNA del campione ignoto è resa possibile dalla valutazione dei cicli soglia (*CT*) e confrontando i campioni con una curva di taratura a quattro punti ottenuta con una serie di diluizioni a partire dal Il *PowerQuant*® *Male gDNA standard*. Prima di procedere con la quantificazione è necessario allestire le diluizioni utilizzando il DNA di riferimento ed il buffer *PowerQuant*® *Dilution Buffer* come di seguito riportato (Tabella 9) (*PowerQuant*® *System Technical Manual*, Promega) (ita.promega.com).

Concentrazione dello <i>standard</i> (ng/μl)	<i>PowerQuant® Male gDNA standard</i> (μl)	<i>PowerQuant® Dilution Buffer</i> (μl)
50 ng/μl	<i>PowerQuant® Male gDNA standard</i> non diluito	0 μl
2 ng/μl	4 μl di <i>PowerQuant® Male gDNA standard</i> non diluito	96 μl
0.08 ng/μl	4 μl della diluizione 2 ng/μl	96 μl
0.0032 ng/μl	4 μl della diluizione 0.08 2 ng/μl	96 μl

Tabella 9 | Concentrazioni di DNA utilizzati per costruire la curva di taratura a quattro punti con il kit *PowerQuant® System* Technical Manual, Promega (modificata da ita.promega.com).

La quantità di DNA presente negli standard a concentrazione nota viene stimata interpolando i *CT* corrispondenti, ottenuti sperimentalmente ed indicati sull'asse delle ordinate, con la retta di taratura costruita sulla base delle quantità riportate sull'asse delle ascisse (Figura 22).

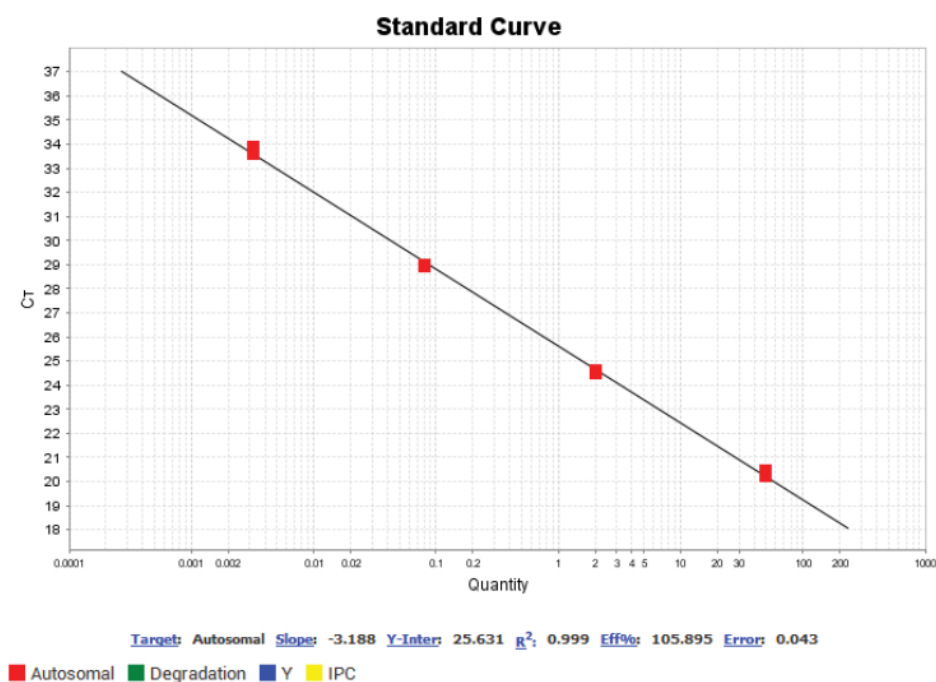


Figura 22 | Esempio di retta di taratura utilizzata per il calcolo grafico della quantità di DNA conoscendo il *CT* (ita.promega.com).

In merito alla pratica svolta in laboratorio, infatti, per ogni piastra di quantificazione, oltre ai campioni da quantificare, viene effettuata l'analisi in doppio dei 4 campioni *standard*, ottenuti a seguito di diluizioni scalari di DNA a concentrazione nota (Tabella 9). Per la preparazione della piastra si procede con la preparazione di una *mix* di amplificazione che verrà messa in ogni pozzetto della piastra di *PCR 96 Well*, le quantità di ogni componente della *mix* sono riportate di seguito (Tabella 10).

Componente	Volume per reazione (µl)
<i>PowerQuant® 2X Master Mix</i>	10 µl
<i>PowerQuant® 20X Primer/Probe/IPC Mix</i>	1 µl
<i>Water, Amplification Grade</i>	7 µl
Volume finale	18 µl

Tabella 10| Reagenti e volumi utilizzati per la preparazione della *mix* di amplificazione con il kit *Power Quant® System* (Promega) (modificata da ita.promega.com).

Le quantità di reagenti sono state moltiplicate per il numero complessivo di campioni da esaminare e sono stati aliquotati 18 µl per ogni pozzetto, successivamente sono stati aggiunti 2 µl di ogni DNA sia quelli ignoti da quantificare che gli *standard*. Prima di essere utilizzati, i reagenti del kit sono stati scongelati a temperatura ambiente, vortexati e centrifugati brevemente. Per ogni piastra è stato aggiunto un controllo negativo (*NTC*) ed il bianco estratto nella fase di estrazione del DNA. L'*NTC* è il bianco della reazione di quantificazione di *RT-PCR*, in cui sono presenti tutti i reagenti necessari all'amplificazione ad eccezione del DNA templatato estratto. Eseguendo questo controllo si può verificare che non ci siano contaminazioni da parte di DNA estraneo, garantendo così l'autenticità e la pulizia dei reagenti della metodica. Dopo l'aggiunta dei DNA per arrivare al volume totale per pozzetto di 20 µl, la piastra è stata chiusa con le pellicole adesive *MicroAmp™ Optical Adhesive Film* (Applied Biosystems™) e successivamente inserita nello strumento *7500 fast Real-Time Applied Biosystems®*, impostando le seguenti condizioni di amplificazione:

- *Holding stage* a 98° C per 2 minuti;
- *Cycling stage*: *Step 1* 98°C per 15 secondi e lo *Step 2* a 62°C per 35 secondi.

Per un numero di cicli pari almeno a 39 (*PowerQuant® System Technical Manual*, Promega) (ita.promega.com). Con l'utilizzo del *7500 Software v 2.3* sono stati inseriti il numero e la tipologia di campioni analizzati durante l'esperimento (traccia, *standard*, bianco, *NTC*) e i codici alfanumerici identificativi di ogni campione. Ad ogni posizione della piastra è stato poi assegnato il codice. I valori ottenuti per ogni sonda nella reazione di quantificazione sono stati inseriti nel software *PowerQuant Analysis 1.0.0.0.xlsm*. Una volta inseriti i dati nel software, quest'ultimo calcola in automatico la retta di taratura e la concentrazione dei vari campioni.

3.5 Amplificazione tramite *PCR multiplex*

I laboratori di scienze forensi e di tipizzazione del DNA hanno tratto enorme beneficio dalla scoperta di una tecnica nota come reazione a catena della polimerasi (*PCR*). Descritta per la prima volta nel 1985 da *Kary Mullis* e dai membri del gruppo di Genetica Umana della *Cetus Corporation*, la *PCR* ha rivoluzionato la biologia molecolare grazie alla capacità di produrre centinaia di milioni di copie di una specifica sequenza di DNA nel giro di poche ore (Mullis et al., 1986). L'impatto della *PCR* è stato tale che il suo inventore, *Kary Mullis*, ha ricevuto il Premio *Nobel* per la Chimica nel 1993, meno di dieci anni dopo la sua prima descrizione. Senza la capacità di duplicare le molecole di DNA, molti campioni forensi sarebbero impossibili da analizzare. Il DNA rinvenuto sulla scena del crimine è spesso limitato sia in termini di quantità che di qualità, e ottenere un campione più pulito e concentrato è normalmente fuori questione. La tecnologia di amplificazione del DNA tramite *PCR* si adatta perfettamente all'analisi dei campioni biologici forensi perché è sensibile, rapida e meno vincolata alla qualità del DNA rispetto ai metodi originari di polimorfismo della lunghezza dei frammenti di restrizione (*RFLP*) utilizzati in passato (Butler, 2011). La *PCR multiplex* è una tecnica utilizzata per co-amplificare più frammenti simultaneamente partendo da uno stesso DNA stampo; è un metodo di *parallel sample processing* e si possono amplificare oltre a 20 filamenti in un'unica reazione, evitando così lo spreco di materiale. La reazione è di tipo ciclico e viene ripetuta un numero di volte che varia in base alla quantità di *target* desiderata. La *PCR* replica *in vitro* un passaggio specifico della duplicazione cellulare: la sintesi di un segmento di DNA a doppio filamento (*dsDNA*) a partire da un filamento a singola elica (*ssDNA*). Partendo da un filamento di DNA denaturato, un nuovo filamento complementare viene sintetizzato aggiungendo nucleotidi nella sequenza corretta (Van Pelt-Verkuil et al., 2008). Questo processo è svolto da una DNA polimerasi termoresistente, che utilizza come innesco delle brevi sequenze di DNA (circa 19-25 basi) chiamate *primers*, complementari alle regioni che affiancano la sequenza bersaglio alle estremità 5' e 3' (Holland et al., 1991). La miscela di amplificazione deve contenere:

- L'enzima DNA Polimerasi, in particolare la *Taq* Polimerasi che deriva dall'enzima naturalmente presente nel batterio termofilo *Thermus Aquaticus*. La *Taq* Polimerasi è un tipo particolare di DNA Polimerasi che è termoresistente e riesce a lavorare in maniera efficiente anche a temperature elevate e questo è fondamentale poiché le varie fasi della reazione di amplificazione richiedono dei passaggi a temperature molto alte;
- Magnesio cloruro ($MgCl_2$) necessario in quanto cofattore dell'enzima;

- I *primers* che vanno ad appaiarsi in maniera complementare in un punto specifico del DNA e ne delimitano il frammento che verrà amplificato. I primers sono due uno che va ad appaiarsi sul filamento *forward* e l'altro che va sul filamento *reverse* fungendo da innesco per la *Taq* Polimerasi;
- I Deossinucleotidi trifosfato (*dNTPs*) che sono i nucleotidi che vengono utilizzati dalla *Taq* Polimerasi per sintetizzare il nuovo filamento;
- Un *buffer* che consiste in una soluzione tampone composta da Tris HCl a pH 8.3 e KCl. Il Tris HCl permette di mantenere il pH stabile anche nelle elevate temperature della reazione permettendo una condizione ottimale di lavoro per la *Taq* Polimerasi. Il KCl favorisce l'*annealing* dei *primers* poiché fornisce ioni K^+ che stabilizzano le interazioni elettrostatiche tra i primer e il DNA stampo, favorendo l'appaiamento.

I passaggi che consentono l'amplificazione sono scanditi da cambi di temperatura che favoriscono ciascun processo nell'ordine prestabilito. Operativamente le variazioni di temperatura, che devono essere precise e mantenute per un corretto periodo di tempo, vengono gestite grazie ad uno strumento chiamato termociclatore, sul quale si possono impostare i protocolli in uso che esso andrà a seguire.

Le fasi di una *PCR* sono:

- Denaturazione iniziale: avviene a circa 95°C e serve a separare le due eliche di DNA rendendole disponibili all'attacco della polimerasi;
- Denaturazione: questa seconda denaturazione verrà fatta all'inizio di ogni ciclo, infatti, da questo punto in avanti, i passaggi verranno ripetuti un numero variabile di volte, a seconda del protocollo in uso;
- *Annealing*: la temperatura viene abbassata tra i 40 e i 60°C per permettere ai *primers* di appaiarsi con le sequenze corrispondenti;
- Estensione: a questo punto si ha la replicazione vera e propria del DNA. Avviene alla temperatura ottimale per la *Taq* Polimerasi, ovvero 72°C, e consiste nell'allungamento dei *primers* sulla base del DNA stampo. Se si lavora con quantità minime di DNA saranno necessari più cicli;
- Estensione finale: quest'ultimo passaggio viene fatto una sola volta al termine del processo. Serve a consentire alla polimerasi di ultimare le estensioni.

Questo ciclo viene ripetuto generalmente per circa 30 volte, permettendo ai nuovi filamenti di fungere da stampo per i cicli successivi. Durante la PCR, l'amplificazione del DNA segue una cinetica suddivisa in tre fasi distinte, condizionate dalla disponibilità dei reagenti e dall'efficienza dell'enzima: Fase esponenziale: la reazione presenta la massima efficienza e reagenti in abbondanza; il DNA raddoppia ad ogni ciclo con crescita ideale pari a 2^n , dove n è il numero di cicli (Saiki et al. 1988). Fase lineare: la velocità rallenta poiché i reagenti diventano limitanti e i prodotti competono con i primer per l'annealing, traducendosi in una crescita non più esponenziale ma lineare. Fase di plateau: la quantità di prodotto si stabilizza al massimo raggiungibile a causa dell'esaurimento dei reagenti o della perdita di attività dell'enzima. La cinetica complessiva si schematizza riportando su un piano cartesiano il logaritmo della quantità di amplificato in ordinata e il numero di cicli termici in ascissa (Figura 21).

3.5.1 Amplificazione con kit *GlobalFiler™ PCR*

Il kit *GlobalFiler™ PCR Amplification* (ThermoFisher Scientific), validato per applicazioni forensi, prevede l'amplificazione di 23 marcatori *STR*, di cui 21 autosomici tra cui 14 *CODIS loci*: CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, e D21S11, 6 *loci* dall' *expanded European Standard Set of Loci (ESSL)*: D2S441, D22S1045, D10S1248, D1S1656, D12S391, D2S1338 e il *locus* con elevato tasso di discriminazione SE33. E' presente inoltre 1 *STR* del cromosoma Y che è il DYS391. Inoltre è presente anche un marcatore polimorfico *insertion/deletion* sul cromosoma Y: *Y indel*. A tutti questi si aggiunge l'Amelogenina per la determinazione del sesso (Tabella 12). Si tratta, dunque, di una *PCR multiplex*, ovvero una reazione singola di *PCR* che prevede l'amplificazione di più *loci* contemporaneamente. Ciò consente di ottenere risultati anche da quantità iniziali minime di DNA. I vari *loci* amplificati vengono differenziati grazie a fluorofori specifici, in questo caso sono cinque, a cui si aggiunge quello dello *standard* interno. Ciascun fluoroforo è associato a diversi *STR* che si distinguono in base al peso molecolare (Tabella 11,12).

Dye	Color	Label
<i>6-FAMTM</i>	<i>Blue</i>	<i>Samples, allelic ladder and controls</i>
<i>VICTM</i>	<i>Green</i>	
<i>NEDTM</i>	<i>Yellow</i>	
<i>TAZTM</i>	<i>Red</i>	
<i>SIDTM</i>	<i>Purple</i>	
<i>LIZTM</i>	<i>Orange</i>	<i>GeneScanTM 600 LIZTM size standard v2.0</i>

Tabella 11 | Elenco dei fluorofori presenti all'interno del kit *GlobalFilerTM* con i rispettivi colori e gli elementi a cui sono legati (*GlobalFilerTM PCR Amplification Kits User Guide*; documents.thermofisher.com).

Locus designation	Chromosome location	Dye label
D3S1358	3p21.31	<i>6-FAMTM</i>
vWA	12p13.31	
D16S539	16q24.1	
CSF1PO	5q33.3–34	
TPOX	2p23–2per	
Y indel	Yq11.221	<i>VICTM</i>
Amelogenin	X: p22.1–22.3 Y: p11.2	
D8S1179	8q24.13	
D21S11	21q11.2–q21	
D18S51	18q21.33	
DYS391	Yq11.21	<i>NEDTM</i>
D2S441	2p14	
D19S433	19q12	
TH01	11p15.5	
FGA	4q28	
D22S1045	22q12.3	<i>TAZTM</i>
D5S818	5q21–31	
D13S317	13q22–31	
D7S820	7q11.21–22	
SE33	6q14	
D10S1248	10q26.3	<i>SIDTM</i>
D1S1656	1q42.2	
D12S391	12p13.2	
D2S1338	2q35	

Tabella 12 | Elenco di tutti i marcatori presenti nel kit *GlobalFilerTM* con la loro localizzazione cromosomica e il fluoroforo associato (modificata da *GlobalFilerTM PCR Amplification Kits User Guide*; documents.thermofisher.com).

La miscela di reazione della *PCR* viene svolta in 12.5 µl così ripartiti:

- 3.75 µl di *Master Mix*;
- 1.25 µl di *Primers Set*;
- DNA estratto nella concentrazione ottimale per l'amplificazione in base alla quantificazione ottenuta (tra i 500 pg e 1 ng);
- H_2O fino a raggiungere il volume totale di 12.5 µl.

Le provette sono state inserite in un termociclatore utilizzando il seguente protocollo (Tabella 13).

<i>Initial incubation</i>	<i>Cycle (29 or 30 cycles)</i>		<i>Final extension</i>	<i>Final hold</i>
	<i>Denature</i>	<i>Anneal/Extend</i>		
<i>HOLD</i>	<i>CYCLE</i>		<i>HOLD</i>	<i>HOLD</i>
95 °C 1 minute	94 °C 10 seconds	59 °C 90 seconds	60 °C 10 minutes	4 °C ≤ 24 hours

Tabella 13 | Protocollo utilizzato per l'amplificazione impostato sul termociclatore (*GlobalFiler™ PCR Amplification Kits* User Guide) (documents.thermofisher.com).

Per ogni reazione di *PCR*, oltre ai campioni estratti di DNA si procede ad amplificare un campione contenente la miscela di reazione e l'acqua al posto del DNA, questo funge da controllo negativo della *PCR* per verificare che i componenti della miscela non siano contaminati. Inoltre si amplifica un altro controllo negativo, in questo caso è il negativo di estrazione che è il prodotto della estrazione fatta in precedenza, ma senza aver inserito il campione contenete DNA. Infine si inserisce un controllo positivo *2800M control DNA* (Promega) o *DNA Control 007* (ThermoFisher).

3.6 Elettroforesi capillare

L'elettroforesi capillare (*CE*) è la metodologia principale utilizzata per separare e rilevare gli alleli *STR* nei laboratori forensi di DNA in tutto il mondo. Questa tecnica sfrutta la presenza di cariche negative, che caratterizzano i frammenti di DNA, facendoli migrare in presenza di un campo elettrico creato da due elettrodi che sono un polo negativo (catodo) e un polo positivo (anodo). La velocità di migrazione dei frammenti di DNA amplificati è direttamente proporzionale alla loro dimensioni in paia di basi. Per ottenere una tipizzazione degli *STR* che sia affidabile, devono essere soddisfatte tre condizioni. In primo luogo, è necessaria una risoluzione spaziale per separare gli alleli *STR* che

potrebbero differire in dimensioni anche di un singolo nucleotide. In secondo luogo, è necessaria una risoluzione spettrale per separare tra loro i colori dei coloranti fluorescenti, in modo tale da poter distinguere i prodotti della *PCR* provenienti da *loci* marcati con fluorocromi differenti. In terzo luogo, la precisione nella determinazione delle dimensioni del DNA da una corsa all'altra deve essere sufficientemente costante, così da poter correlare i campioni ai *ladder* allelici utilizzati (Butler, 2011). Questa tecnica si serve di un capillare rivestito di poliammide e silice fusa del diametro interno di circa 50 μm e anti-connettivo, ovvero che limita la quantità di calore generato anche quando vengono applicate correnti molto elevate. Il capillare viene riempito da un polimero utilizzato come setaccio molecolare, la polidimetil-acrilammide (*POP*). Sul capillare è presente una finestra, unica parte trasparente e non flessibile del capillare, che viene utilizzata come cella di rivelazione: infatti, essa viene colpita da un raggio *laser*, il quale ha la funzione di eccitare i fluorocromi associati ai frammenti di DNA in modo da ottenere una fluorescenza che viene registrata dal detector. La rilevazione del passaggio dei frammenti all'interno del capillare viene operata da una camera *CCD* (*Charge Coupled Device*) attraverso l'unica porzione del capillare non rivestita da poliammide. In questa porzione la fluorescenza emessa dai fluorocromi legati agli ampliconi viene captata e il segnale trasferito al software associato allo strumento dove verrà analizzata. L'*output* della corsa elettroforetica è un elettroferogramma, ovvero un grafico che riporta sulle ascisse il peso molecolare, da cui si ricava la lunghezza in paia di basi (bp) dei frammenti amplificati, e sulle ordinate l'intensità della fluorescenza relativa (rfu) di ogni allele amplificato (Figura 23) (Durney, Carihfield, and Holland 2015).

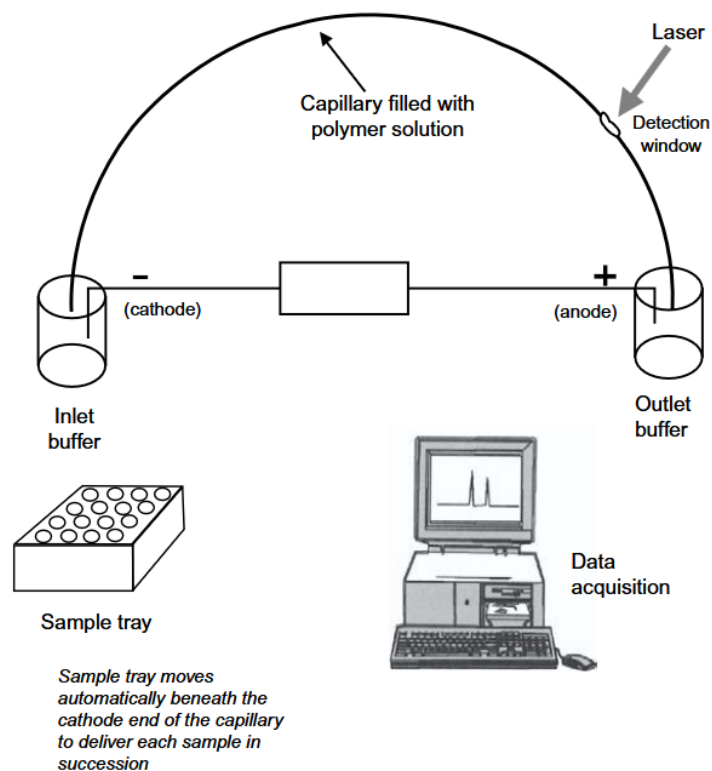


Figura 23 | Schema rappresentativo del funzionamento di uno strumento di elettroforesi capillare usato per tipizzare il DNA (Butler, 2011).

Dal tempo di migrazione degli ampliconi è possibile calcolarne l'esatto peso molecolare, usufruendo uno standard interno. Lo standard interno è un marcatore a peso molecolare noto che migra insieme al campione; in questo modo un software dedicato determina una curva di riferimento, la quale viene utilizzata per calcolare il peso molecolare dei campioni d'interesse interpolando il loro tempo di migrazione con la retta. Per svolgere questo lavoro, l'elettroforesi capillare, è stata condotta su un sequenziatore automatico *SeqStudio Genetic Analyzer* (AppliedBiosystems). Il *SeqStudio Genetic Analyzer* ha un sistema di elettroforesi a 4 capillari e rappresenta uno strumento molto moderno per l'analisi di frammenti per l'identificazione umana (*HID*). Lo strumento è composto da tre principali componenti che sono la cartuccia *All-in-one*, la piastra su cui vengono caricati i campioni e un computer con interfaccia a schermo (Figura 24) (*SeqStudio Genetic Analyzer*, AppliedBiosystems) (Thermofisher.com).



Figura 24 | Rappresentazione grafica delle principali componenti del sequenziatore automatico *SeqStudio Genetic Analyzer*, AppliedBiosystems (modificata da Thermofisher.com).

Il cuore funzionale dello strumento, estremamente facile da usare, è il risultato di un *design* innovativo della cartuccia che contribuisce a massimizzare l'efficienza e la praticità. L'analizzatore genetico *SeqStudio* utilizza una cartuccia *all-in-one* che contiene al suo interno la matrice di capillari, il serbatoio del polimero e il tampone anodico. Ogni cartuccia contiene un polimero esclusivo per il sistema *SeqStudio* che è l' *Applied Biosystems™ POP-1™ polymer* che consente di eseguire il sequenziamento *Sanger* e l'analisi di frammenti senza alcuna riconfigurazione. La cartuccia dispone di quattro capillari e può processare campioni provenienti sia da piastre standard a 96 pozzetti che da strisce di tubi a 8 pozzetti (*strip*) (Figura 25) (*SeqStudio Genetic Analyzer*, AppliedBiosystems) (Thermofisher.com).



Figura 25 | Rappresentazione grafica delle componenti della cartuccia *all-in-one* del sequenziatore *SeqStudio Genetic Analyzer* (*SeqStudio Genetic Analyzer*, AppliedBiosystems) (Thermofisher.com).

3.6.1 Preparazione della miscela di reazione per l'elettroforesi capillare

Per ciascun campione è stata preparata una provetta ciascuna contenente la seguente miscela di componenti:

- 1 µl di DNA amplificato;
- 10 µl di Formammide;
- 0.2 µl di *GeneScan™ 600 LIZ™ size standard v2.0*.

Dopo aver preparato la miscela per ciascun campione, si procede con la denaturazione dei campioni sottoponendoli ad una temperatura di 95 °C per 3 minuti. In seguito alla denaturazione vengono raffreddati brevemente in ghiaccio per 3 minuti al fine di bloccare la denaturazione tramite *shock* termico. Dopo la denaturazione i campioni sono stati caricati sulla piastra dello strumento *SeqStudio Genetic Analyzer* per la reazione elettroforetica.

Per l'analisi sono stati impostati sullo strumento i seguenti parametri operativi:

- Modulo di corsa: *HID Analysis*;
- Temperatura di corsa: 60°C;
- Voltaggio di iniezione: 1.2 kV;
- Tempo di iniezione: 10 secondi;
- Voltaggio di corsa: 11 kV;
- Tempo di corsa: 18 minuti.

Insieme ai campioni in studio vengono fatti correre anche uno standard interno e un *ladder* allelico. Lo standard interno è costituito da un insieme di frammenti di DNA a peso molecolare noto marcati con un fluorocromo differente rispetto a quelli associati agli altri campioni amplificati (Tabella 11). Il kit *GlobalFiler™ PCR Amplification* prevede l'utilizzo dello standard interno *GeneScan™ 600 LIZ™ size standard v2.0* (ThermoFisher) (Figura 26).

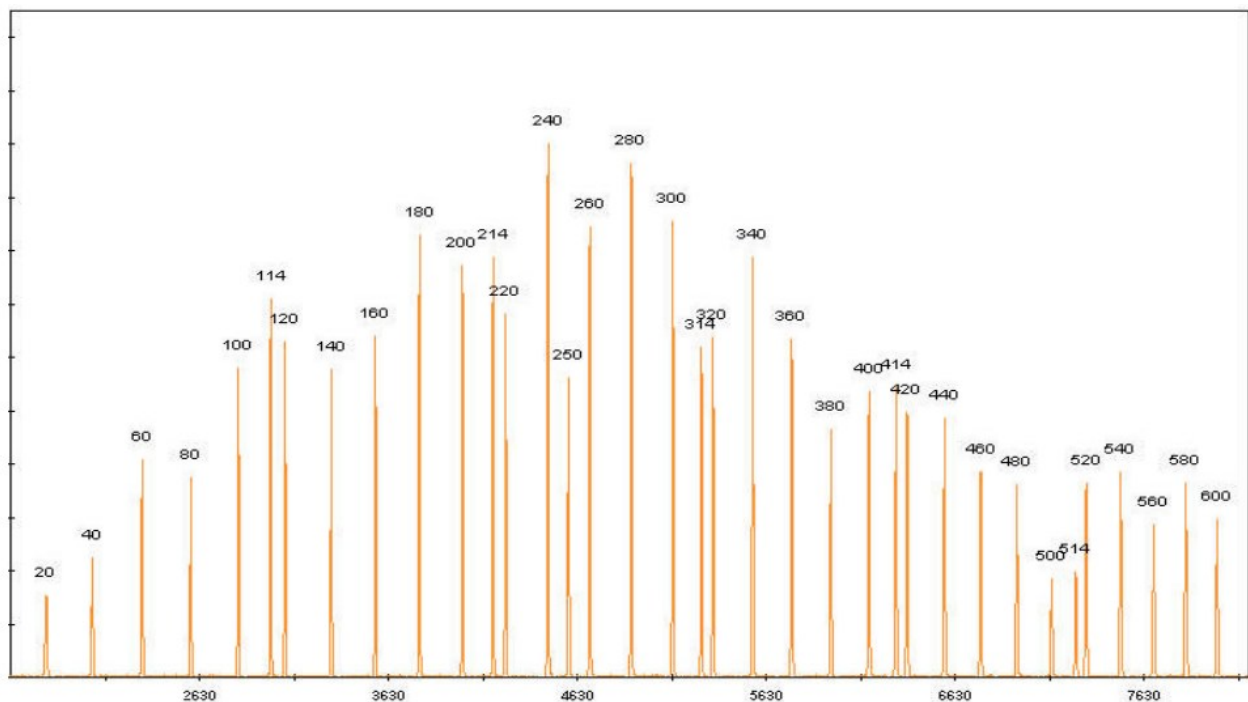


Figura 26| Elettroferogramma del *GeneScan™ 600 LIZ™ size standard v2.0*, i frammenti sono catalogati per lunghezza in bp (*GeneScan™ 600 LIZ™ size standard full sheet*) (assets.fishersci.com).

Il *ladder* allelico consiste invece in una miscela di tutti gli alleli più diffusi per ogni *locus STR* analizzato con il kit. Il *GlobalFiler™ allelic ladder* è fornito dall'azienda produttrice insieme al kit ed è composto da un totale di 589 alleli divisi per colore come i marcatori *STR*. Nel *SeqStudio Genetic Analyzer* i frammenti che vengono separati durante la corsa elettroforetica vengono rilevati mediante fluorescenza. I fluorocromi associati ai frammenti di DNA vengono eccitati da una sorgente luminosa che in questo caso è un *laser*. La fluorescenza emessa viene rilevata da un sensore *CCD* e questi segnali luminosi sono trasformati poi in segnali digitali che vengono mandati e analizzati dal software dello strumento. Il software è il *GeneMapper ID-X v 1.6* che genera un elettroferogramma in cui ogni allele amplificato è rappresentato da un picco colorato (Figura 27). La posizione di ogni picco corrisponde al peso molecolare dell'allele in bp e l'altezza riflette la quantità di DNA presente. Grazie all'analisi dello standard interno, il software assegna precisamente il peso molecolare di ciascun frammento di DNA per ogni *locus* e per confronto con il *ladder* allelico è in grado di assegnare il nome degli alleli dei frammenti amplificati, fornendo risultati molto accurati e soprattutto replicabili.

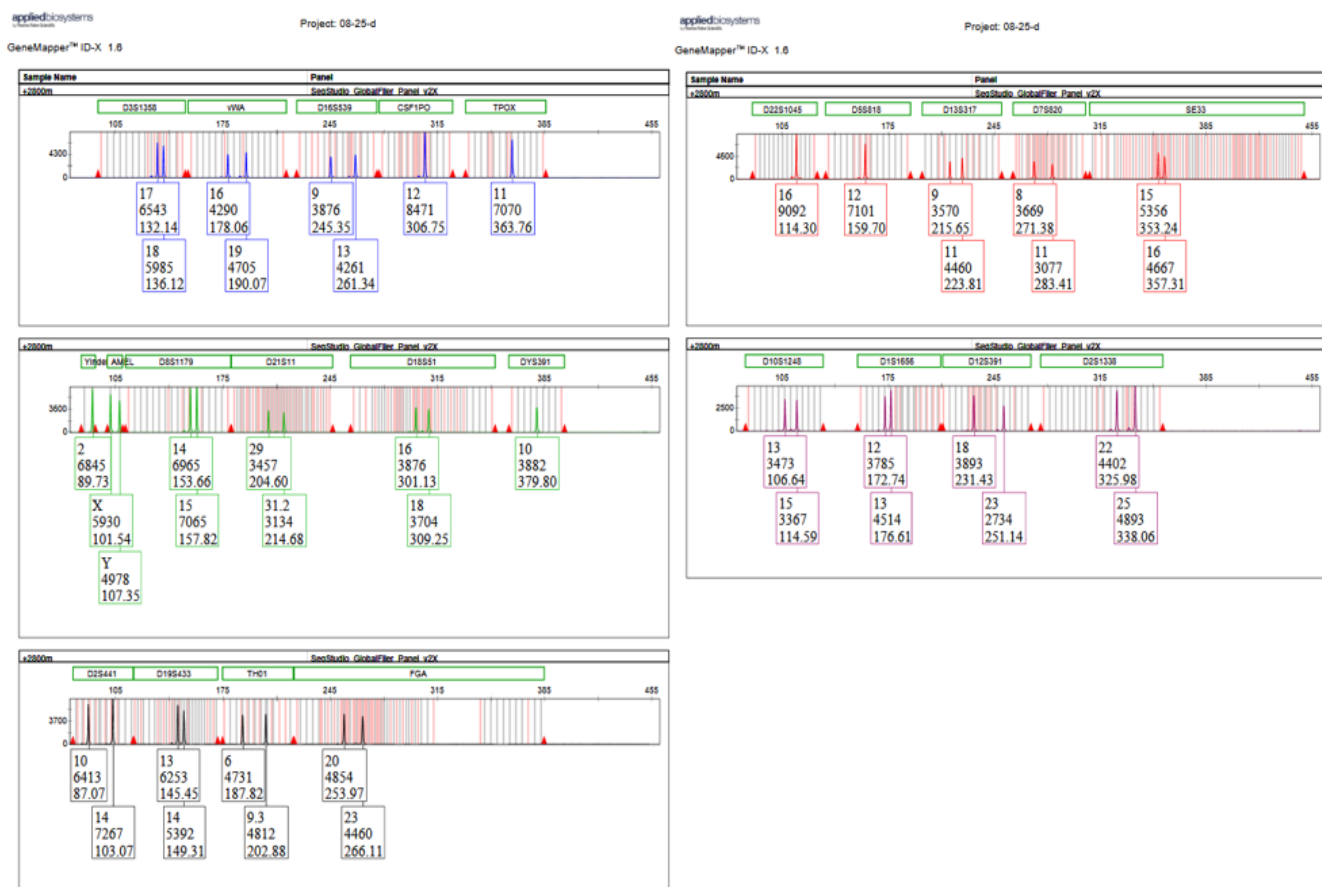


Figura 27 Esempio di tracciato elettroforetico del DNA di controllo 2800M (Promega) ottenuto tramite amplificazione con kit *GlobalFiler™ PCR Amplification* ed elaborato dal software *GeneMapper ID-X v 1.6*.

4 RISULTATI E DISCUSSIONE

4.1 Quantificazione del DNA

4.1.1 Quantificazione del campione autoptico FFPE 1

A seguito della definizione del disegno sperimentale (Tabella 7), si è proceduto a pianificare le estrazioni di DNA dalle sezioni dei tessuti in paraffina selezionati. Le analisi sono iniziate dal primo campione autoptico, denominato "FFPE 1", applicando il protocollo estrattivo illustrato in Figura 16. Nello specifico, dal blocchetto in paraffina contenente il tessuto cardiaco sono state ricavate 26 sezioni di uguale spessore (10 μ m). Su tali sezioni, le estrazioni sono state eseguite in duplicato (2x13) per ciascuna delle soluzioni previste dall'*Experimental design* e riportate nella Tabella 7. Successivamente alle 13 estrazioni in duplicato, è stata effettuata un'analisi di quantificazione molecolare del DNA così estratto mediante *Real-Time PCR*, in accordo con le procedure descritte nel Capitolo 3 "Materiali e Metodi". Al fine di valutare l'efficacia del protocollo e individuare i componenti chimici più o meno performanti della soluzione, l'analisi si è basata sul parametro di riferimento della quantificazione definito [Auto]. Tale valore indica la concentrazione di DNA (espressa in ng/ μ l) riferibile alla sonda autosomica di dimensioni minori (come dettagliato a pagina 65 del capitolo "Materiali e Metodi"). I valori di [Auto] ottenuti sulla prima e la seconda serie di 13 estrazioni sono mostrati nella seguente Tabella 14.

Exp#	1° estrazione	2° estrazione
	[Auto]	[Auto]
1	0.4652	0.9951
2	0.2847	0.503
3	0.3089	0.624
4	0.5696	0.5855
5	0.005	0.0129
6	0.1143	0.1916
7	0.7349	1.0045
8	0.5812	1.0189
9	0.9395	0.8122
10	0.3876	0.3655
11	0.3339	0.3936
12	0.1852	0.2707
13	0.369	0.5174

Tabella 14 | Valori di quantità di DNA [Auto], espressi in ng/ μ l, ottenuti nelle prime 13 estrazioni effettuate a partire da sezioni di tessuto cardiaco del campione autoptico FFPE 1.

Una volta ottenuti questi risultati, si è proceduto a valutare dal punto di vista statistico i risultati, per comprendere quali fossero i componenti in grado di influire in maniera positiva o negativa sulla quantità di DNA ottenuta. L'analisi statistica impiegata è la Regressione Lineare Multipla, unitamente all'analisi della varianza (*ANOVA*). Sono state messe a confronto le risposte in termini di valori di [Auto] con l'aggiunta o la rimozione dei singoli componenti. Inoltre, l'analisi utilizza delle variabili fittizie definite e_1 , e_2 , e_3 . Non sono ingredienti reali, ma servono al sistema matematico per misurare l'errore sperimentale intrinseco e confrontarlo con l'effetto degli ingredienti reali. Nell'analisi statistica i valori di [Auto] della prima e della seconda estrazione sono stati accorpati. Il valore di R-quadrato (R^2) indica che percentuale dei dati originali è spiegata dal modello di regressione adottato. In questo caso il valore di R-quadrato è di 0.83324. Questo significa che circa l'83 % delle variazioni dell'effetto finale che stiamo misurando può essere spiegato dall'andamento della causa che abbiamo preso in esame. La restante percentuale delle oscillazioni sfugge a questa spiegazione ed è attribuibile ad altri fattori esterni o al puro caso. E' stato effettuato anche un test *ANOVA* della regressione che serve a valutare la validità dell'intero modello matematico creato dal disegno sperimentale. Nell'ambito dell'*ANOVA*, il valore che dà un'informazione interessante per capire se l'analisi è coerente con lo studio è il valore della statistica F , che in questo caso è pari a 0,0034. Questo valore è decisamente minore della soglia critica di significatività 0.05 (5%). Il valore della significatività "a posteriori" di F è definito *p-value* e indica che il modello è altamente significativo e affidabile. Questi risultati confortano, dal punto di vista scientifico, nel procedere con l'analisi successiva mirata alla comprensione di quali componenti della soluzione hanno migliorato o peggiorato la resa finale di estrazione. I risultati dell'analisi di regressione lineare multipla applicata ai dati ottenuti dal disegno sperimentale di *Plackett-Burman* sono riportati nel seguente grafico (Figura 28) e nella seguente tabella (Tabella 15).

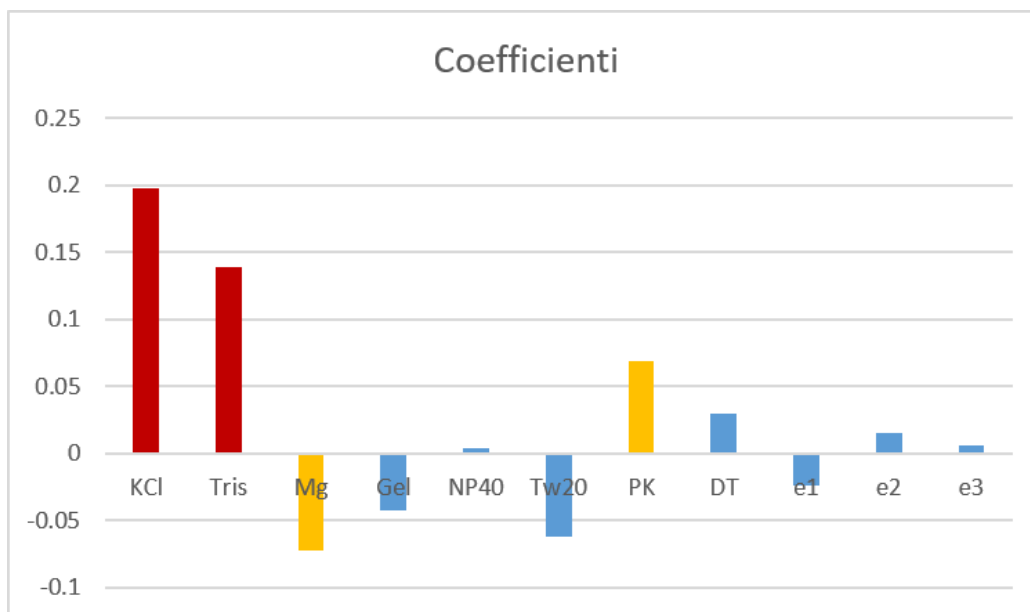


Figura 28| Rappresentazione grafica dei coefficienti del modello di *Plackett-Burman* ottenuti dall'analisi di regressione lineare multipla effettuata sulle due serie di estrazioni riportate nelle Tabelle 14 e 15.

Componente	Coefficienti	<i>P-value</i>
KCl	0.19734583	0.000124
Tris-HCl	0.1386875	0.002084
MgCl ₂	-0.0727458	0.062956
Gelatina	-0.0424125	0.255263
NP40	0.00367917	0.919166
Tween-20	-0.0615792	0.108384
Proteinasi K	0.06907917	0.075458
<i>D.Time</i>	0.0295625	0.421239
e1	-0.0237958	0.515344
e2	0.01537917	0.672534
e3	0.00582917	0.872302

Tabella 15| Valori dei coefficienti ottenuti per ogni termine del modello di *Plackett-Burman* e i *p-value* corrispondenti.

Per l'interpretazione del modello sono necessari due valori fondamentali: il primo valore è il coefficiente, il quale indica l'entità e la direzione dell'effetto del componente. Un coefficiente positivo indica che la presenza o l'aumento della concentrazione di quel componente della soluzione incrementa la resa di estrazione; al contrario, un coefficiente negativo suggerisce un effetto peggiorativo in termini di quantità di DNA estratto e quindi ne compromette in parte la resa finale. Il secondo valore è il *P-value* che rappresenta la probabilità che l'effetto osservato sia imputabile al caso quindi ad un errore casuale. Ai fini di questo studio, è stata adottata la soglia di significatività statistica standard $\alpha = 0,05$ (5%). Sulla base di questi parametri, i fattori analizzati si dividono in tre categorie:

- Fattori Altamente Significativi (Evidenziati in giallo brillante nella Tabella 15 e in rosso in Figura 28). L'analisi identifica due componenti essenziali che guidano l'efficienza dell'estrazione e sono il Cloruro di Potassio (KCl) e il Tris-HCl. Il KCl risulta essere il parametro più influente, mostrando il coefficiente positivo più alto in assoluto (0,197) e un *p-value* estremamente significativo ($p = 0,000124$). Il Tris-HCl mostra anch'esso un forte impatto positivo (coefficiente 0,138) supportato da un'elevata significatività statistica ($p = 0,002084$). Entrambi sono pertanto fattori critici e indispensabili per l'ottenimento di una buona resa di estrazione del DNA (Tabella 15);
- Fattori a Significatività Marginale (Evidenziati in giallo/arancione pallido), due variabili si attestano su valori vicini alla soglia limite di significatività ($0,05 < p < 0,1$), indicando delle tendenze che meritano una considerazione, pur non superando la soglia di significatività statistica fissata a priori $\alpha = 0.05$. Il $MgCl_2$ presenta un *p-value* di 0,062 e un coefficiente negativo (-0,072). Questo suggerisce una tendenza del magnesio a comportarsi come inibitore all'interno della soluzione, riducendo le performance di estrazione e questo si può notare bene nel grafico (Figura 28) dove la barra del magnesio, avendo un coefficiente negativo, è rivolta al di sotto della linea di soglia. L'altro componente con significatività marginale è la Proteinasi K con un *p-value* di 0,075 associato a un coefficiente positivo (0,069). Indica un probabile, seppur debole, effetto positivo sul processo (Tabella 15);
- Fattori Non Significativi, i restanti componenti della miscela, ovvero i detergenti (NP40, Tween-20), la gelatina e il tempo di denaturazione (*DT*), restituiscono *p-value* ampiamente superiori a 0,10. Dal punto di vista statistico, la loro variazione equivale a quella dei fattori fittizi, e1, e2, ed e3 (Tabella 15). Valutando i risultati ottenuti dall'analisi di regressione multipla effettuata sulla base del valore [Auto], si può riassumere quali sono i componenti che migliorano o peggiorano l'estrazione (Tabella 16).

KCl (mM)	100 (significativo)
Tris-HCl (mM)	20 (significativo)
MgCl₂ (mM)	0 (marginalmente significativo)
Gelatina (mg/mL)	0 (non significativo)
NP40 (%)	Indifferente
Tween-20 (%)	0 (non significativo)
Proteinasi K (mg/mL)	2 (marginalmente significativo)
D.Time (min)	10 (non significativo)

Tabella 16 | tabella riassuntiva dei componenti della miscela di estrazione del DNA testata e il loro impatto relativo sull'efficienza di recupero del DNA, per quanto riguarda il campione FFPE 1.

Per questo primo campione autoptico, l'analisi statistica è stata effettuata unicamente sulla base del valore [Auto], parametro di riferimento per calcolare la quantità totale di DNA estratto. Questo poiché, la *RT PCR* non ha fornito alcun valore di quantificazione della sonda [Deg] a più alto peso molecolare (294 bp). Da ciò si desume che non vi sia DNA nel range di peso molecolare della sonda [Deg], a motivo della degradazione/modificazione dell'acido nucleico ragionevolmente dovuta al tempo di fissazione in formalina del tessuto (tra 1 e 2 settimane, vedi Tabella 4). La conseguenza è che, per tale campione, non è stato possibile calcolare il rapporto [Auto]/[Deg], definito *Degradation Index* (DI), rapporto che fornisce una stima della qualità del DNA estratto.

4.1.2 Quantificazione del campione autoptico FFPE 2

Successivamente, al fine di acquisire dati genetici da un campione autoptico più "integro" rispetto a quelli forensi, in termini di inferiore tempo post-mortem e minore tempo di fissazione in formalina, è stato selezionato per le analisi tessuto cardiaco prelevato da un soggetto a 40 ore dalla morte e con tempo di fissazione in formalina di 30 ore. Da questo, sono state ricavate 13 sezioni di uguale spessore tra di loro. Su queste ultime è stata condotta una singola serie di estrazioni, testando le 12 soluzioni previste dal disegno sperimentale con l'aggiunta della soluzione originale, impiegata come riferimento (Tabella 7). Successivamente alle 13 estrazioni, è stata effettuata un'analisi di quantificazione molecolare mediante *Real-Time PCR*, in accordo con le procedure descritte nel Capitolo 3 "Materiali e Metodi". A differenza del primo campione autoptico, per il quale è stato valutato esclusivamente il parametro [Auto], per questo secondo campione l'analisi è stata estesa includendo anche il parametro [Deg] che, in relazione a tale campione ha fornito risultati. Tale valore rappresenta la concentrazione di DNA (espressa in ng/μl) riferibile alla sonda di degradazione (come

riportato a pagina 65 del capitolo "Materiali e Metodi"). La disponibilità congiunta dei parametri [Auto] e [Deg] ha permesso di integrare nella valutazione anche il parametro *Degradation Index* (DI), calcolato dal rapporto diretto tra i due valori ([Auto]/[Deg]). Più è vicino a 1 tale rapporto, più il DNA è qualitativamente integro. Il DI costituisce infatti il secondo parametro fondamentale su cui si basa lo studio, affiancando il valore [Auto]. I risultati di [Auto], [Deg] e del relativo rapporto ([Auto]/[Deg]), ottenuti a valle delle 13 estrazioni in doppio, sono riepilogati nella seguente Tabella 17.

Exp#	[Auto]	[Deg]	[Auto]/[Deg]
1	2.0356	0.8644	2.35
2	1.2930	0.5565	2.32
3	0.5800	0.1280	4.53
4	3.8019	1.5793	2.41
5	0.1556	0.0265	5.86
6	2.4332	0.4702	5.17
7	3.7579	0.9885	3.8
8	3.0666	1.1553	2.65
9	2.1891	0.8780	2.49
10	2.6120	0.9356	2.79
11	1.1883	0.4575	2.6
12	1.7549	0.5323	3.3
13	2.4223	0.7224	3.35

Tabella 17 | Valori di quantità di DNA [Auto], [Deg] e [Auto]/[Deg] ottenuti dall'analisi dei 13 campioni estratti da sezioni di tessuto cardiaco del campione autoptico FFPE 2.

A seguito dell'ottenimento di questi risultati, anche in questo caso sono state allestite analisi statistiche per andare a comprendere quali fossero i componenti che influissero in maniera positiva o negativa sulla quantità di DNA ottenuta. L'analisi statistica impiegata è la medesima applicata al primo campione autoptico riportato nel paragrafo precedente, la Regressione Lineare Multipla. In questo caso non è stato possibile effettuare l'*ANOVA* della regressione. In questa analisi avendo 0 gradi di libertà per l'errore, la varianza residua non esiste matematicamente. Di conseguenza, non è possibile generare la classica tabella *ANOVA* e non possono essere forniti dei "*P-value*".

I risultati dell'analisi di regressione lineare multipla applicata ai dati ottenuti dal disegno sperimentale di *Plackett-Burman* e riferiti al valore [Auto], sono riportati nel seguente grafico (Figura 29) e nella seguente tabella (Tabella 18).

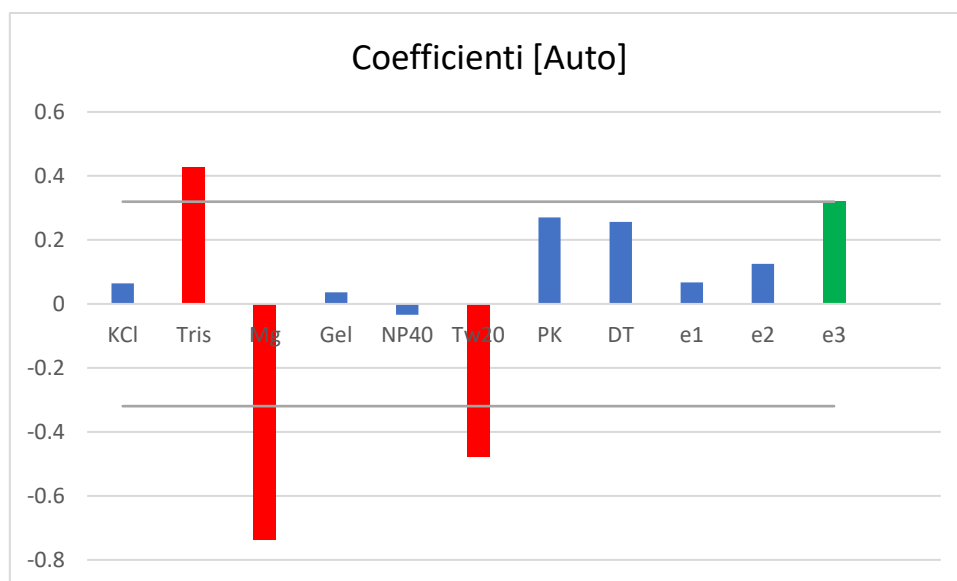


Figura 29| Rappresentazione grafica dei risultati ottenuti dall'analisi di regressione lineare multipla basata sul valore della sonda [Auto] ed effettuata sulle 13 estrazioni riportate nella Tabella 7.

Componente	Coefficienti
KCl	0.0639
Tris-HCl	0.4274
MgCl ₂	-0.736
Gelatina	0.0362
NP40	-0.0336
Tween-20	-0.4783
Proteinasi K	0.2702
<i>D.Time</i>	0.2567
e1	0.0668
e2	0.1249
e3	0.3195

Tabella 18| Rappresentazione dei coefficienti ottenuti per ogni componente per i valori di [Auto].

Per l'interpretazione di questi risultati è necessario osservare i coefficienti, ottenuti dall'analisi di regressione, per ogni componente. Da questi si evince che il Tris-HCl (+ 0.427) è il miglior componente della miscela; la sua presenza incrementa notevolmente la quantità di DNA estratto. Questo risultato è in accordo con quello ottenuto dal primo campione autoptico. All'esatto contrario invece, il magnesio (- 0.736) e il Tween-20 (- 0.478), hanno colonne negative più estese, dimostrando una spiccata azione negativa sul recupero del DNA. Per quanto riguarda il magnesio, in questo secondo campione risulta avere un effetto forte in senso negativo come già evidenziato per il campione precedente. Il Tween-20 in questo caso risulta essere un componente negativo, anche questo in accordo con il coefficiente in Tabella 16. Proteinasi K e tempo di denaturazione dal punto di vista biologico-sperimentale sono sicuramente importanti per la buona riuscita dell'estrazione,

anche se i loro coefficienti in queste prove non sono risultati elevati quanto quelli di Tris-HCl, magnesio e Tween-20. Il KCl, la gelatina e l'NP40, in questo campione sembrano avere un effetto irrilevante. Nonostante ciò, il KCl, avendo un coefficiente positivo, sembra migliorare la quantità finale di DNA estratta come anche mostrato in maniera significativa nel primo campione. La gelatina ha un coefficiente positivo, ma molto basso, risulta essere completamente ininfluyente. Per l'NP40, il coefficiente è negativo ma non in maniera così influente come il magnesio o il Tween-20. In questo secondo campione è stato valutato anche il valore [Deg] poiché il campione ha subito un tempo di fissazione ridotto. I risultati dell'analisi di regressione lineare multipla applicata ai dati ottenuti dal disegno sperimentale di *Plackett-Burman* e riferiti al valore [Deg], sono riportati nel seguente grafico (Figura 30) e nella seguente tabella (Tabella 19).

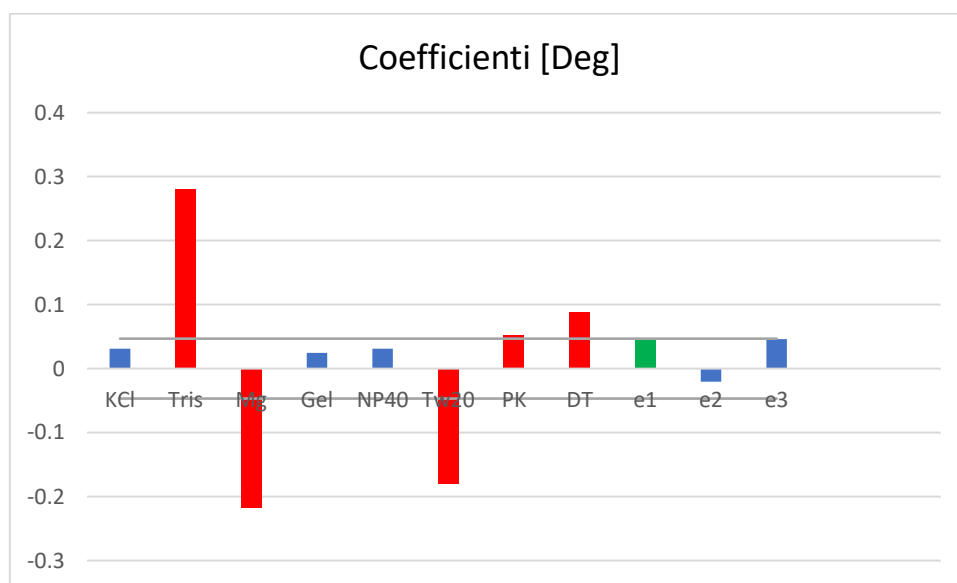


Figura 30| Rappresentazione grafica dei risultati ottenuti dall'analisi di regressione lineare multipla basata sul valore della sonda [Deg] ed effettuata sulle 13 estrazioni riportate nella Tabella 7.

Componente	Coefficienti
KCl	0.0309
Tris-HCl	0.2805
MgCl ₂	-0.2173
Gelatina	0.0248
NP40	0.0311
Tween-20	-0.1809
Proteinasi K	0.0524
<i>D.Time</i>	0.0879
e1	0.0468
e2	-0.0205
e3	0.0462

Tabella 19| Rappresentazione dei coefficienti ottenuti per ogni componente per i valori di [Deg].

Osservando i coefficienti ottenuti dall'analisi di regressione per ogni componente sulla base del valore della sonda [Deg], si può notare che il Tris-HCl (0.2805) è il fattore che ha l'effetto maggiore sulla risposta. Questo risultato è in accordo con quello ottenuto dall'analisi di regressione lineare effettuata sul valore della sonda [Auto]. Analogamente a quello osservato per la sonda [Auto], anche per quella [Deg], il magnesio (-0.2173) e il Tween-20 (-0.1809) hanno una influenza negativa in questa miscela estrattiva poiché portano ad ottenere valori di [Deg] più bassi. Per quanto riguarda la Proteinasi K e il tempo di denaturazione, entrambi risultano essere utili, con coefficienti positivi che hanno un valore superiore a quello della variabile *dummy* e1, nelle loro massime concentrazione e tempo. Il KCl, la gelatina e l'NP40, in questo campione sembrano avere un effetto irrilevante. Nonostante ciò, il KCl, avendo un coefficiente positivo, sembra migliorare la quantità finale di DNA estratta, come anche mostrato in maniera significativa nel primo campione. La gelatina ha un coefficiente positivo, ma molto basso, risulta essere completamente ininfluente come ottenuto già in precedenza per gli altri campioni. In questo caso, al contrario dell'[Auto], l'NP40 ha un coefficiente positivo rimanendo comunque sostanzialmente irrilevante perché hanno coefficienti piccoli numericamente e con valori inferiori a quelli della stima dell'errore casuale e1. Valutando i risultati ottenuti dall'analisi di regressione multipla effettuata sulla base del valore [Auto] e [Deg], si può riassumere quali sono i componenti che migliorano o peggiorano l'estrazione (Tabella 20).

Componente	[Auto]	[Deg]
KCl (mM)	Indifferente (meglio 100)	Indifferente (meglio 100)
Tris-HCl (mM)	20	20
MgCl ₂ (mM)	0	0
Gelatina (mg/mL)	Indifferente (probabilmente meglio 0.5)	Indifferente (probabilmente meglio 0.5)
NP40 (%)	Indifferente (meglio 0)	Indifferente (meglio 2)
Tween-20 (%)	0	0
Proteinasi K (mg/mL)	Indifferente (meglio 2)	Indifferente (meglio 2)
D.Time (min)	Indifferente (meglio 10)	Indifferente (meglio 10)

Tabella 20 | tabella riassuntiva dei componenti della miscela di estrazione del DNA testata e il loro impatto relativo sull'efficienza di recupero del DNA, per quanto riguarda il campione FFPE 2.

4.1.3 Quantificazione del campione autoptico FFPE 3

Le analisi del campione autoptico denominato "FFPE 3" sono state condotte su sezioni in paraffina di tessuto cardiaco. Tredici sezioni di identico spessore (10 μm) sono state ricavate per le analisi. Su queste ultime è stata condotta una singola serie di estrazioni, testando le 12 soluzioni previste dal disegno sperimentale con l'aggiunta della soluzione originale, impiegata come riferimento (Tabella 7). A conclusione delle 13 estrazioni, è stata effettuata un'analisi di quantificazione molecolare del DNA mediante *Real-Time PCR*, in accordo con le procedure descritte nel Capitolo 3 "Materiali e Metodi". A differenza del primo campione autoptico, per il quale è stato valutato esclusivamente il parametro [Auto] (espresso in $\text{ng}/\mu\text{l}$), al fine di individuare le componenti migliori e peggiori della soluzione, per questo terzo campione l'analisi è stata estesa includendo anche il parametro [Deg] come per il secondo campione. In questo terzo campione autoptico il valore [Deg] era presente, nonostante i tempi di fissazione prolungati (tra 1 e 2 settimane) cui era stato sottoposto il tessuto fossero simili a quelli del primo campione autoptico forense (Tabella 4). I risultati di [Auto], [Deg] e del relativo rapporto ([Auto]/[Deg]), ottenuti a valle delle 13 estrazioni, sono riepilogati nella Tabella 21.

Exp#	[Auto]	[Deg]	[Auto]/[Deg]
1	1.4125	0.0118	119.25
2	1.2991	0.0122	106.55
3	0.7331	0.0061	119.79
4	1.4222	0.0149	95.4
5	0.8702	0.0025	349.6
6	1.0918	0.0076	144.24
7	1.0940	0.0123	89.02
8	1.5483	0.0176	88.18
9	0.9429	0.0104	90.42
10	0.6087	0.0069	88.13
11	0.8648	0.0103	83.74
12	0.4245	0.0051	83.21
13	1.4610	0.0134	109.34

Tabella 21 | Valori di quantità di DNA [Auto], [Deg] e [Auto]/[Deg] ottenuti dall'analisi dei 13 campioni estratti da sezioni di tessuto cardiaco del campione autoptico FFPE 3.

A seguito dell'ottenimento di questi risultati, sono state condotte le analisi statistiche già descritte per i precedenti campioni *FFPE*. I risultati dell'analisi di regressione lineare multipla applicata ai dati ottenuti dal disegno sperimentale di *Plackett-Burman* sono riportati nel seguente grafico (Figura 31) e nella seguente tabella (Tabella 22).

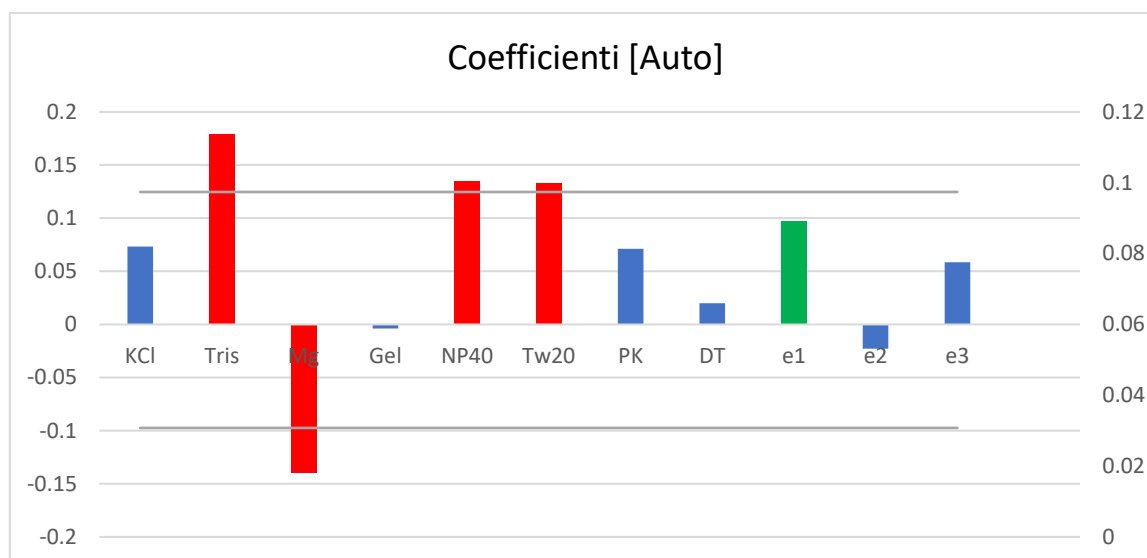


Figura 31 | Rappresentazione grafica dei risultati ottenuti dall'analisi di regressione lineare multipla basata sul valore della sonda [Auto] ed effettuata sulle 13 estrazioni riportate nella Tabella 7.

Componente	Coefficienti
KCl	0.0733
Tris-HCl	0.1796
MgCl ₂	-0.1395
Gelatina	-0.0038
NP40	0.1345
Tween-20	0.1332
Proteinasi K	0.0712
<i>D.Time</i>	0.0201
e1	0.0974
e2	-0.0227
e3	0.0584

Tabella 22 | Rappresentazione dei coefficienti ottenuti per ogni componente per i valori di [Auto].

Per l'interpretazione di questi risultati, è necessario osservare i coefficienti, ottenuti dall'analisi di regressione, per ogni componente. Da questi si evince che il Tris (+ 0.1796) è il fattore che ha l'effetto più importante sulla risposta sperimentale. Il Tris è un componente molto utile per ottenere valori di [Auto] alti, questo risultato ottenuto dal terzo campione autoptico è in accordo con quello ottenuto dai due precedenti. All'esatto contrario invece, il magnesio (- 0.1395) ha la colonna negativa più estesa, dimostrando una spiccata azione negativa sul recupero del DNA come già evidenziato in tutte le analisi precedenti sull'[Auto]. Il Tween-20 (+ 0.1332) e l'NP40 (+ 0.1345) in questo caso risultano avere coefficienti positivi, comportandosi come componenti favorevoli per il recupero del DNA, a differenza delle analisi sui precedenti campioni che, per contro, mostravano valori negativi o positivi ma molto vicini allo zero. Per quanto riguarda la Proteinasi K (+ 0.0712) e il tempo di denaturazione

(+ 0.0201), entrambi risultano essere utili al loro massimo livello, rispettivamente 2 mg/mL e 10 minuti anche se il coefficiente non è così elevato come quelli di Tris, magnesio, NP40 e Tween-20. Il KCl e la gelatina, in questo campione sembrano avere un effetto più trascurabile. Nonostante ciò, il KCl (+ 0.0733), che ha un coefficiente positivo, sembra migliorare la quantità finale di DNA estratta seppure non in maniera evidente come nel primo campione. La gelatina ha un coefficiente negativo (- 0.0038), ma molto basso, ed in questo caso sembra essere completamente ininfluenza. I risultati dell'analisi di regressione lineare multipla applicata ai dati ottenuti dal disegno sperimentale di *Plackett-Burman* e riferiti al valore [Deg], sono riportati nel seguente grafico (Figura 32) e nella seguente tabella (Tabella 23).

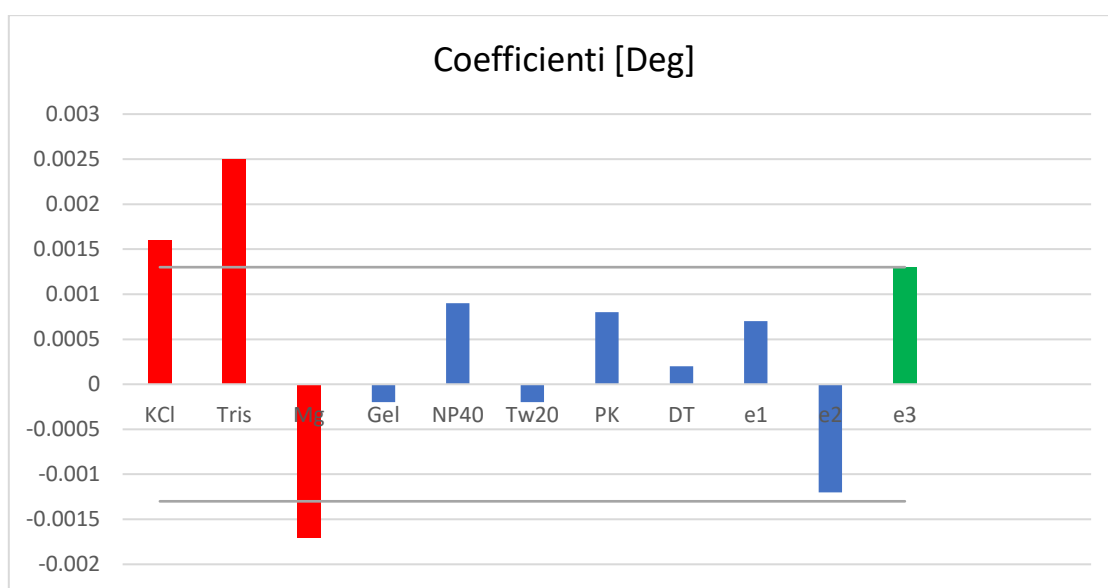


Figura 32 | Rappresentazione grafica dei risultati ottenuti dall'analisi di regressione lineare multipla basata sul valore della sonda [Deg] ed effettuata sulle 13 estrazioni riportate nella Tabella 7.

Componente	Coefficienti
KCl	0.0016
Tris-HCl	0.0025
MgCl ₂	-0.0017
Gelatina	-0.0002
NP40	0.0009
Tween-20	-0.0002
Proteinasi K	0.0008
<i>D.Time</i>	0.0002
e1	0.0007
e2	-0.0012
e3	0.0013

Tabella 23 | Rappresentazione dei coefficienti ottenuti per ogni componente per i valori della sonda [Deg].

Osservando i coefficienti ottenuti dall'analisi di regressione per ogni componente sulla base del valore della sonda [Deg], si può notare che il Tris (0.0025) ha l'influenza più elevata sui valori di [Deg]. Questo risultato è in accordo con quello ottenuto dall'analisi di regressione lineare effettuata sul valore [Auto] e [Deg] nel secondo campione autoptico. Analogamente a quello osservato per l'[Auto], anche per il [Deg], il magnesio (-0.0017) e il Tween-20 (-0.0002) hanno un'influenza negativa in questa miscela estrattiva poiché portano ad ottenere valori di [Deg] più bassi. Il magnesio anche in questo caso, come in tutti i casi precedenti, ha un forte impatto negativo, il Tween-20 invece, per questo terzo campione, riguardo al [Deg], risulta indifferente nonostante il coefficiente negativo indichi che ha un effetto che influisce riducendo la risposta sperimentale. Questo è contrario al risultato ottenuto per lo stesso campione analizzando il valore [Auto], ma in accordo con i risultati ottenuti per i campioni primo e secondo. Per quanto riguarda la Proteinasi K (0.0008) e il tempo di denaturazione (0.0002), entrambi risultano avere coefficienti positivi. In particolare, la Proteinasi K mostra un valore intorno alla variabile *dummy* e1 (0.0007), risultando utile nella sua massima concentrazione. Il tempo di denaturazione anche in questo caso non ha un valore molto importante, ma il coefficiente positivo è un dato coerente con tutti i casi visti in precedenza. A differenza di quanto osservato in precedenza, in questo campione il KCl (0.0016) ha un coefficiente marcatamente positivo, migliora la quantità finale di DNA estratta in maniera rilevante, come anche mostrato nel primo campione FFPE 1. La gelatina (-0.0002) ha un coefficiente negativo, ma molto basso, e risulta essere completamente ininfluenza come ottenuto già in precedenza per gli altri campioni. In questo caso, l'NP40 (0.0009) ha un coefficiente positivo dimostrando un'utilità nell'estrazione come in accordo con i risultati sull'[Auto] di questo stesso campione.

Valutando i risultati ottenuti dall'analisi di regressione multipla effettuata sulla base del valore [Auto] e [Deg], si può riassumere quali sono i componenti che migliorano o peggiorano l'estrazione (Tabella 24).

Componente	[Auto]	[Deg]
KCl (mM)	Indifferente (meglio 100)	100
Tris-HCl (mM)	20	20
MgCl ₂ (mM)	0	0
Gelatina (mg/mL)	Indifferente	Indifferente (meglio 0)
NP40 (%)	2	Indifferente (meglio 2)
Tween-20 (%)	2	Indifferente (meglio 0)
Proteinasi K (mg/mL)	Indifferente (meglio 2)	Indifferente (meglio 2)
D.Time (min)	Indifferente (meglio 10)	Indifferente (meglio 10)

Tabella 24 | tabella riassuntiva dei componenti della miscela di estrazione del DNA testata e il loro impatto relativo sull'efficienza di recupero del DNA, per quanto riguarda il campione FFPE3.

4.2 Analisi “Effetto tessuto”

A seguito della definizione del disegno sperimentale allestito per valutare la presenza o meno di un “effetto tessuto” (Tabella 8), cioè di una eventuale differenza nell'efficienza di recupero del DNA da differenti tessuti legata ad una diversa formulazione della miscela di estrazione, si è proceduto selezionando tre differenti tessuti, cardiaco, encefalico e polmonare provenienti da uno medesimo soggetto denominato "FFPE 3", applicando il protocollo estrattivo illustrato in Figura 16. Tale scelta è motivata dal fatto di diminuire la variabilità, già intrinsecamente elevata tra tutti i campioni presi in studio, legata a differenti tempistiche di fissazione dei tessuti in formalina. Nello specifico, sulla base delle elaborazioni dell'*Experimental design*, dai blocchetti in paraffina contenenti i tre tessuti sono state ricavate rispettivamente 4 sezioni di 10 µm cardiache, 5 sezioni encefaliche e 6 sezioni polmonari. Su queste sono state condotte le estrazioni pianificate nel disegno sperimentale, unitamente ad una estrazione per tessuto con la soluzione originale, impiegata come riferimento (Tabella 8). Successivamente, il DNA estratto dalle 15 prove è stato quantificato mediante *Real-Time PCR*, in accordo con le procedure descritte nel Capitolo 3 "Materiali e Metodi. I risultati di [Auto], [Deg] e del relativo rapporto ([Auto]/[Deg]) ottenuti sono riepilogati nella Tabella 25.

Exp#	Tessuto	[Auto]	[Deg]	[Auto]/[Deg]
1	Cardiaco	0.0683	<i>Undetermined</i>	<i>Undetermined</i>
2	Cardiaco	0.804	0.0042	190.61
3	Cardiaco	0.3971	0.0035	114.63
4	Encefalico	2.3602	0.0339	69.69
5	Encefalico	4.0906	0.0556	73.63
6	Encefalico	1.1689	0.0194	60.41
7	Encefalico	2.0604	0.0382	53.95
8	Polmonare	0.9311	0.0247	37.64
9	Polmonare	2.0968	0.0579	36.2
10	Polmonare	2.6792	0.07	38.28
11	Polmonare	2.3713	0.0656	36.15
12	Polmonare	1.8378	0.0304	60.49
13	Cardiaco	0.5779	0.0043	134.63
13	Encefalico	2.7115	0.054	50.26
13	Polmonare	2.0882	0.056	37.28

Tabella 25 | Valori di [Auto], [Deg] e [Auto]/[Deg] ottenuti nelle estrazioni svolte su sezioni di tessuto cardiaco, encefalico e polmonare paraffinato del soggetto FFPE 3.

L'analisi dei risultati dell'esperimento "*effetto tessuto*" è stata condotta per verificare non solo quali fossero i componenti della miscela con effetti positivi o negativi sull'estrazione, ma anche se le diverse concentrazioni dei componenti modificassero la resa di DNA, in funzione dei differenti tessuti. Il modello aveva anche l'obiettivo di misurare la "linea di base" intrinseca di ciascun tessuto, ovvero la loro naturale differenza di partenza in termini di caratteristiche dell'estratto utilizzando come riferimento di base il tessuto cardiaco. I risultati dell'analisi di regressione lineare multipla applicata ai dati ottenuti dal disegno sperimentale di *Plackett-Burman* sono riportati nel seguente grafico (Figura 33) e nella seguente tabella (Tabella 26).

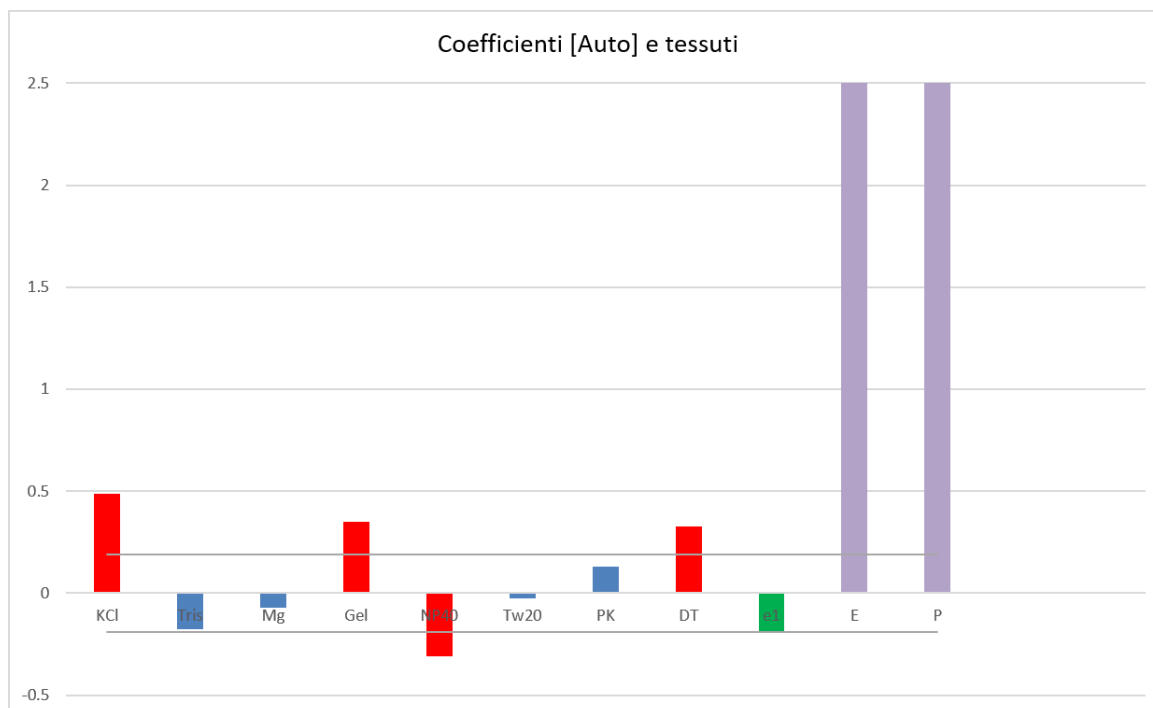


Figura 33 | Rappresentazione grafica dei risultati ottenuti dall'analisi di regressione lineare multipla basata sul valore [Auto] ed effettuata sulle estrazioni riportate nella Tabella 8. Rappresentazione anche delle due variabili tessuto "E" e "P".

Componente	Coefficienti
KCl	0.486
Tris-HCl	-0.176
MgCl ₂	-0.072
Gelatina	0.35
NP40	-0.308
Tween-20	-0.025
Proteinasi K	0.128
<i>D.Time</i>	0.325
e1	-0.189
E	2.165
P	2.296

Tabella 26 | Rappresentazione dei coefficienti ottenuti per ogni componente per i valori di [Auto] e le variabili tessuto.

Per l'interpretazione di questi risultati, è necessario osservare i coefficienti, ottenuti dall'analisi di regressione, per ogni componente. Da questi si evince innanzitutto il peso predominante della matrice biologica: il polmone P (+ 2.296) e l'encefalo E (+ 2.165) presentano coefficienti positivi molto alti, dimostrando che il tipo di tessuto ha un'influenza intrinseca preponderante sui valori di [Auto] rispetto al tessuto cardiaco di riferimento. Spostando l'attenzione sui componenti della miscela di lisi, il KCl (+ 0.486) è il componente che ha l'effetto più rilevante nella soluzione; la sua presenza incrementa notevolmente la quantità di DNA estratto ed è molto utile per ottenere valori di [Auto] alti.

Oltre al KCl, anche la gelatina (+ 0.35) e il tempo di digestione DT (+ 0.325) in questo caso risultano avere coefficienti alti e positivi, e sono interpretabili come fattori che contribuiscono in modo importante a massimizzare la resa. All'esatto contrario invece, l'NP40 (- 0.308) ha la colonna negativa più estesa, dimostrando un'azione contraria sul recupero del DNA. Anche il Tris-HCl (- 0.176) agisce in senso negativo in questo esperimento, dimostrando di avere un effetto dannoso sulla resa di estrazione. Questo risultato non è in accordo con quelli ottenuti in precedenza probabilmente a motivo di una variabilità inter e intra individuale dei campioni non completamente esplorata nello studio condotto in questa tesi. Per quanto riguarda la Proteinasi K (+ 0.128), questo fattore ha un effetto positivo e utile nella sua massima concentrazione, anche se il coefficiente non è così elevato come quelli di KCl, gelatina e DT, rimanendo al di sotto della soglia del rumore di fondo stimata dalla variabile *dummy* e1 (- 0.189). Il magnesio e il Tween-20, in questo *screening*, sembrano avere un effetto trascurabile. Valutando i risultati ottenuti dall'analisi di regressione multipla effettuata sulla base del valore [Auto], si può riassumere quali sono i componenti che migliorano o peggiorano l'estrazione (Tabella 27)

KCl (mM)	100
Tris-HCl (mM)	Indifferente (meglio 5)
MgCl₂ (mM)	Indifferente (meglio 0)
Gelatina (mg/mL)	0.5
NP40 (%)	0
Tween-20 (%)	Indifferente
Proteinasi K (mg/mL)	2
D. Time (min)	10

Tabella 27 tabella riassuntiva dei componenti della miscela di estrazione del DNA testata e il loro impatto relativo sull'efficienza di recupero del DNA, per quanto riguarda il campione FFPE 3.

4.3 Miscela di estrazione finale ottimizzata

A seguito delle sperimentazioni condotte sui differenti campioni FFPE 1, 2 e 3 e sui tre tessuti selezionati (cuore, encefalo, polmone), si è infine giunti a stabilire quali componenti della miscela di estrazione abbiano un effetto positivo o negativo sull'estrazione del DNA, in termini di resa quantitativa, come misurato dalla sonda Auto, l'unica per cui erano disponibili i valori di quantificazione per tutti e tre i campioni analizzati.

Di seguito è riassunto il peso e la rilevanza dei parametri presi in considerazione nella formulazione delle miscele, dopo ogni serie di estrazione (Tabella 28).

[Auto]				
Componenti	Campione			
	FFPE 1	FFPE 2	FFPE 3	Effetto tessuto
KCl (mM)	100	Indiff. (100)	Indiff. (100)	100
Tris-HCl (mM)	20	20	20	Indiff. (5)
MgCl₂ (mM)	0	0	0	Indiff. (0)
Gelatina (mg/mL)	Indiff. (0)	Indiff. (0.5)	Indiff.	0.5
NP40 (%)	Indiff.	Indiff. (0)	2	0
Tween-20 (%)	Indiff. (0)	0	2	Indiff.
Proteinasi K (mg/mL)	2	Indiff. (2)	Indiff. (2)	2
D. Time (min)	Indiff. (10)	Indiff. (10)	Indiff. (10)	10

Tabella 28| Tabella riassuntiva dell’impatto dei parametri considerati nella formulazioni delle miscele di estrazione ottenute dall’analisi di regressione lineare multipla sul valore [Auto,] in relazione ai tessuti cardiaci dei campioni autoptici FFPE 1, 2 e 3 e delle analisi svolte nell’esperimento “Effetto tessuto”.

Utilizzando quindi tutti i coefficienti calcolati dall’analisi di tutti i campioni presenti nella tabella 29, è stato costruito un grafico che riporta tutti valori dei coefficienti per ogni campione e la media di tutti questi per ogni fattore della miscela (Figura 34).

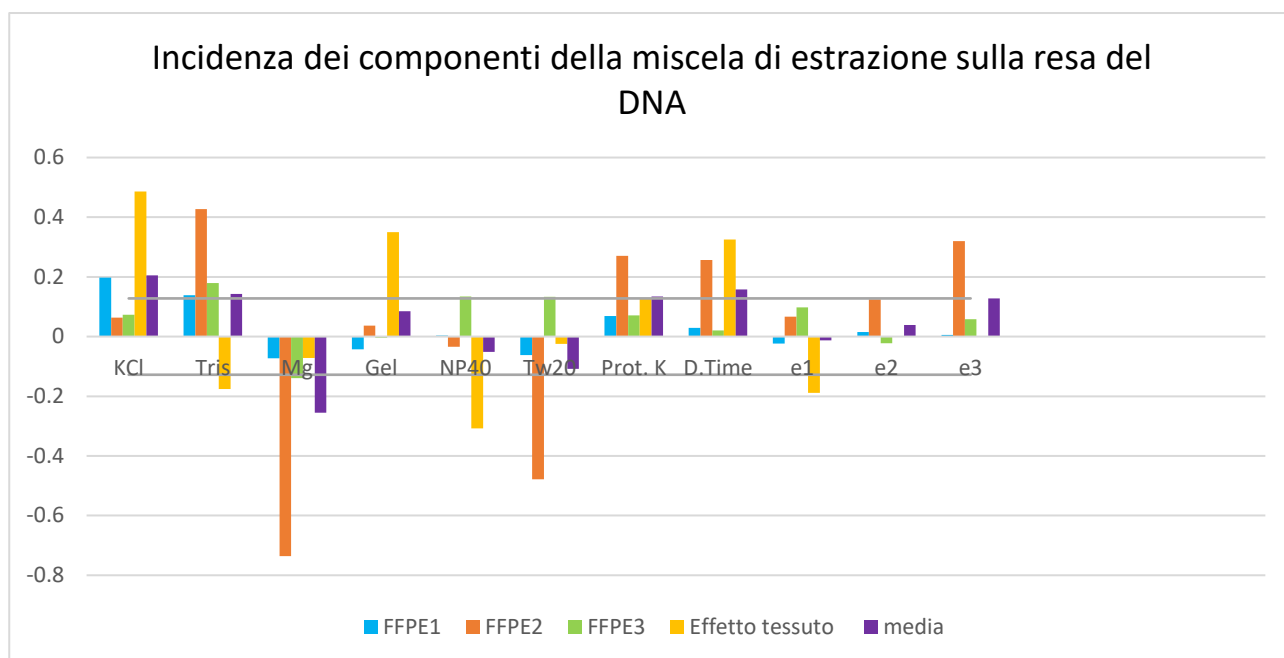


Figura 34| Rappresentazione grafica dell'incidenza dei componenti della miscela di estrazione sulla resa del DNA. I risultati sono stati ottenuti dall'analisi di regressione lineare multipla basata sul valore [Auto] per i campioni autoptici FFPE 1, FFPE 2, FFPE 3 e per l'esperimento "Effetto tessuto". E' rappresentata anche la media dei coefficienti per ogni componente della soluzione.

L'analisi della media dei coefficienti (viola) consente di superare le fluttuazioni intrinseche dei singoli campioni, permettendo di definire la composizione più promettente della soluzione di estrazione. L'andamento della media dei coefficienti (colonna viola) traccia una chiara linea di demarcazione tra i fattori che migliorano per l'efficienza della lisi e quelli dannosi o influenti. I risultati mostrano una spiccata e sistematica efficacia del KCl, che registra il coefficiente medio positivo più alto in assoluto (0.2051). La sua azione, essenziale per il mantenimento dell'equilibrio osmotico e strutturale, si è dimostrata importante in tutti gli scenari testati. Parallelamente, la Proteinasi K e il Tempo di Denaturazione confermano il loro ruolo cruciale: le loro colonne medie spiccatamente positive evidenziano come una digestione enzimatica prolungata e un tempo di denaturazione lungo possano migliorare la resa finale di DNA totale estratto. Il Tris, sebbene abbia mostrato una fluttuazione nel test "Effetto tessuto", mantiene un coefficiente medio numericamente grande e positivo, confermandosi un tampone che determina un contributo positivo nella maggior parte dei campioni analizzati. Dal grafico, non solo si possono definire quali sono i fattori che aumentino la resa di DNA, ma si può individuare quali abbiano un effetto negativo. Il dato sperimentale inequivocabile dell'intero studio riguarda il Magnesio Cloruro. Questo ione bivalente ha costantemente un coefficiente medio grande e numericamente negativo (-0.255), particolarmente evidente nel caso della drastica inibizione registrata nel campione FFPE 2. La sua presenza ad alte concentrazioni si

conferma quindi controproducente per l'efficienza di recupero del DNA, probabilmente a causa dell'attivazione di nucleasi endogene $MgCl_2$ dipendenti. Allo stesso modo, i detergenti non ionici Tween-20 e NP40 restituiscono coefficienti medi negativi; il loro impiego ad alte concentrazioni non solo non migliora la lisi cellulare rispetto ai protocolli base, ma sembra interferire con la resa finale. La Gelatina (Gel) ha un effetto modesto in generale, che sembra dipendere dal tipo di tessuto: debolmente negativo o ininfluyente nei tessuti cardiaci mentre si rivela invece benefica nel test su encefalo e polmone, portando la sua media complessiva in valore positivo; tuttavia non è stata aggiunta alla soluzione finale ottimizzata, in quanto risulta indifferente e non migliorativa in termini significativi di resa finale di estrazione.

Sulla base delle evidenze sperimentali emerse dall'analisi dei dati acquisiti, la formulazione finale ottimizzata per il tampone di estrazione relativo a campioni autoptici è la seguente (Tabella 29):

KCl (mM)	100
Tris-HCl (mM)	20
MgCl₂ (mM)	0
Gelatina (mg/mL)	0
NP40 (%)	0
Tween-20 (%)	0
Proteinasi K (mg/mL)	2
D.Time (min)	10

Tabella 29 | Componenti della miscela finale ottimizzata a seguito delle analisi effettuate sui tessuti cardiaci dei campioni autoptici FFPE 1, 2, 3 e sui tessuti encefalico e polmonare dell'esperimento "Effetto tessuto".

4.4 Confronto tra estrazione con la miscela ottimizzata ed estrazione automatizzata con kit

A seguito della definizione della formulazione ottimale della miscela di estrazione, è stato condotto uno studio comparativo per valutare le prestazioni del protocollo di laboratorio ottimizzato rispetto a quelle di un sistema di estrazione automatizzato basato su kit commerciale. A tal fine, sono state analizzate parallelamente con entrambe le metodologie 24 sezioni di tessuto complessive, suddivise in modo omogeneo tra tre diverse tipologie di matrice biologica: 8 campioni di tessuto cardiaco; 8 campioni di tessuto encefalico; 8 campioni di tessuto polmonare. Per azzerare la variabilità biologica inter-individuale e garantire il massimo rigore nel confronto metodologico, tutte le matrici tissutali sono state prelevate da un unico soggetto (il medesimo impiegato per la caratterizzazione del terzo campione autoptico FFPE 3 e per le analisi sull'"Effetto tessuto").

Al termine dei processi estrattivi, ciascun eluito è stato sottoposto a quantificazione molecolare mediante *Real-Time PCR*, secondo le procedure operative dettagliatamente descritte nel Capitolo 3 ("*Materiali e Metodi*"). Per tutti i campioni di ogni tessuto, sia estratti con la metodica di laboratorio ottimizzata che con quella automatizzata sono stati valutati i valori delle sonde [Auto] (espressi in ng/μl), [Deg] e [Auto]/[Deg]. Tali valori, unitamente a quello calcolato della resa totale di DNA estratto espressa in ng (ottenuto moltiplicando il valore [Auto] per il volume finale di eluizione pari a 150 μl per quella ottimizzata e 50 μl per la metodica automatizzata) ottenuti, sono riportati nella Tabella 30. I medesimi valori ottenuti per il tessuto encefalico e polmonare sono riportati rispettivamente nella Tabella 31 e nella Tabella 32.

Exp#	[Auto]	[Deg]	[Auto]/[Deg]	DNA totale estratto
C1	0.9383	0.0081	115.56	140.75
C2	1.3685	0.0135	101.25	205.275
C3	0.8107	0.0100	80.74	121.605
C4	1.2024	0.0137	87.47	180.36
C5	1.1530	0.0134	85.97	172.95
C6	1.2769	0.0107	119.59	191.535
C7	0.7484	0.0063	118.67	112.26
C8	1.8313	0.0143	127.98	274.70
C1 pro	1.0388	0.0159	65.17	51.94
C2 pro	1.2728	0.0317	40.13	63.64
C3 pro	1.3566	0.0316	42.92	67.83
C4 pro	0.9523	0.0280	34.04	47.615
C5 pro	1.2047	0.0384	31.37	60.235
C6 pro	1.5896	0.0427	37.20	79.48
C7 pro	1.0623	0.0231	46.03	53.12
C8 pro	1.2688	0.0287	44.21	63.44

Tabella 30 | Valori di [Auto], [Deg], [Auto]/[Deg] e DNA totale estratto ottenuti nelle 8 estrazioni condotte utilizzando la miscela ottimizzata e le 8 con il kit automatizzato su sezioni di tessuto cardiaco.

Exp#	[Auto]	[Deg]	[Auto]/[Deg]	DNA totale estratto
E1	3.9692	0.0682	58.24	595.38
E2	2.8012	0.0533	52.52	420.18
E3	3.1131	0.0695	44.81	466.97
E4	3.0101	0.0665	45.27	451.52
E5	2.1004	0.0639	32.88	315.06
E6	3.1075	0.0779	39.91	466.13
E7	2.4176	0.0494	48.91	362.64
E8	3.6833	0.0648	56.82	552.50
E1 pro	2.9486	0.1022	28.86	147.43
E2 pro	4.9250	0.1713	28.75	246.25
E3 pro	3.8978	0.1626	23.98	194.89
E4 pro	3.6942	0.1600	23.08	184.71
E5 pro	4.0904	0.1934	21.15	204.52
E6 pro	3.3897	0.1381	24.54	169.49
E7 pro	3.7551	0.1511	24.85	187.76
E8 pro	4.4101	0.1519	29.02	220.51

Tabella 31 | Valori di [Auto], [Deg], [Auto]/[Deg] e DNA totale estratto ottenuti nelle 8 estrazioni condotte utilizzando la miscela ottimizzata e le 8 con il kit automatizzato su sezioni di tessuto encefalico.

Exp#	[Auto]	[Deg]	[Auto]/[Deg]	DNA totale estratto
P1	0.7763	0.0254	30.60	116.45
P2	2.2272	0.0611	36.47	334.08
P3	2.4644	0.0855	28.83	369.66
P4	2.0495	0.0716	28.60	307.43
P5	2.7379	0.0934	29.31	410.69
P6	2.4843	0.0882	28.18	372.65
P7	2.5395	0.0677	37.49	280.93
P8	2.9776	0.0690	43.17	446.64
P1 pro	4.6053	0.1089	42.29	230.27
P2 pro	3.6525	0.1438	25.39	182.63
P3 pro	2.9802	0.1540	19.35	149.01
P4 pro	4.4186	0.2177	20.30	220.93
P5 pro	4.2057	0.2111	19.92	210.29
P6 pro	2.5305	0.1207	20.97	126.53
P7 pro	3.2875	0.1455	22.60	164.38
P8 pro	4.3455	0.1407	30.88	217.28

Tabella 32 | Valori di [Auto], [Deg], [Auto]/[Deg] e DNA totale estratto ottenuti nelle 8 estrazioni condotte utilizzando la miscela ottimizzata e le 8 con il kit automatizzato su sezioni di tessuto polmonare.

L'analisi dei dati è stata impostata secondo i criteri di indipendenza dei campioni. I campioni estratti con il metodo automatizzato sono indicati dal suffisso “Pro”, a differenza di quelli estratti con il metodo ottimizzato privi di suffisso. La valutazione delle prestazioni di un metodo rispetto ad un altro richiede un'indagine che vada oltre la semplice verifica dell'uguaglianza tra le medie. L'adozione combinata del test di ipotesi classico per la differenza (*t-Student/Welch*) e del test di equivalenza *Two*

One-Sided Tests (TOST) ha consentito di distinguere con precisione due concetti analitici ben differenti: la presenza di un bias sistematico (scostamento delle medie) e l'effetto della dispersione analitica (precisione e variabilità del dato). Per il confronto statistico è stato applicato un semplice test di equivalenza, il *TOST*. Si tratta di un modello ampiamente descritto in letteratura per gli ambiti clinico, farmaceutico, farmacocinetico e chimico (Limentani et al., 2005). L'equivalenza tra due metodi di misura, di estrazione o altro, viene valutata calcolando l'intervallo di confidenza per la differenza tra le medie, $|\mu_{\text{kit}} - \mu_{\text{DoE}}|$. Se tale intervallo cade interamente entro un limite di accettabilità $[-\delta, +\delta]$, prefissato stabilito arbitrariamente pari al $\pm 20\%$ del valore medio del kit, i due metodi si considerano equivalenti. In formula, $|\mu_{\text{kit}} - \mu_{\text{DoE}}| < 0.2 \mu_{\text{kit}}$. Il risultato è un esito binario (SI/NO) che descrive se i due metodi sono equiparabili o meno. Il *TOST* è stato valutato come soluzione ottimale per analizzare i dati ottenuti. A differenza dell'approccio di *Mullen (FSI, 2026)*, basato sul classico test d'ipotesi per la differenza tra medie (*difference test*), il *TOST* ha come scopo principale quello di dimostrare se i due metodi sono equivalenti, in particolare verifica se la differenza tra i risultati dei due metodi è trascurabile ai fini pratici. Questo permette di dimostrare l'equivalenza di fatto dei protocolli, anche nel caso in cui il test della differenza prova che le medie di popolazione μ_{kit} e μ_{DoE} sono differenti tra loro. I risultati del test di equivalenza sui parametri di DNA totale e $[\text{Auto}]/[\text{Deg}]$ per tessuto cardiaco, encefalico e polmonare sono riportati nella Tabella 33.

	DNA totale	[Auto]/[Deg]	Tessuto
Equivalenza	NO	NO	Cardiaco
<i>p(Difference test)</i>	2.96E-05	7.97E-07	
Significato	c > c pro	c > c pro	
Equivalenza	NO	NO	Encefalico
<i>p(Difference test)</i>	4.52E-05	9.08E-05	
Significato	e > e pro	e > e pro	
Equivalenza	NO	NO	Polmonare
<i>p(Difference test)</i>	0.004904	0.04416	
Significato	p > p pro	p > p pro	

Tabella 33 | Rappresentazione dei risultati dell'analisi statistica combinata sui tre diversi tessuti per i parametri DNA totale e $[\text{Auto}]/[\text{Deg}]$.

L'analisi congiunta sui tre tessuti ha evidenziato come le prestazioni dei metodi varino sensibilmente in funzione della matrice biologica e del parametro considerato. Il parametro DNA totale rappresenta l'indice quantitativo primario per valutare l'efficienza estrattiva, la capacità di recupero degli acidi nucleici da parte dei due metodi, cioè la resa quantitativa a partire da una sezione di spessore identico. I risultati ottenuti sull'estrazione del DNA totale mostrano che l'equivalenza statistica tra i due metodi

è difficilmente raggiungibile in modo trasversale su tutte le matrici. Questo fenomeno è ascrivibile a differenze sistematiche nelle rese di estrazione. Quando il test di differenza rileva un $p < 0.05$ significa che le due chimiche di estrazione hanno un'affinità e una capacità di recupero intrinsecamente e significativamente diverse. In presenza di una reale differenza di resa, il test *TOST* esclude rigorosamente l'equivalenza.

In conclusione, si può affermare che il livello di DNA totale estratto è maggiore utilizzando la miscela ottimizzata rispetto all'estrazione con il kit in tutti i tessuti (Figura 35). In particolare, la resa totale del DNA è circa 3 volte più alta con il metodo ottimizzato rispetto al kit per il tessuto cardiaco, 2,3 volte per quello encefalico e 1,75 volte per quello polmonare. E' stato valutato anche il valore $[Auto]/[Deg]$. Questo rapporto è il *Degradation Index* (DI) ed è un parametro qualitativo fondamentale, in quanto fornisce una informazione per quanto riguarda la qualità del DNA; più il valore DI è alto più il DNA è degradato. Valutare questo rapporto permette di capire se il metodo ottimizzato produce un estratto più integro o più degradato rispetto alla metodica commerciale. Anche in questo caso, i risultati dell'analisi statistica non mostrano una significativa equivalenza tra i due metodi, come per la resa del DNA estratto. Dalla analisi risulta che i valori di $[Auto]/[Deg]$ siano significativamente maggiori in tutti i campioni estratti con la miscela ottimizzata rispetto a quelli estratti con il kit per tutti e tre i tessuti (Figura 35). In particolare, il valore di $[Auto]/[Deg]$ ottenuto per il tessuto cardiaco con l'estrazione con la miscela sperimentale risulta essere circa 2,5 maggiore di quello ottenuto con il kit commerciale. Per quando riguarda il tessuto encefalico la miscela di estrazione ottimizzata risulta fornire valori di circa due volte superiori a quello ottenuti con il kit, e per il tessuto polmonare circa 1.5 volte. Questi dati sono univocamente concordi nell'indicare che il DNA estratto con la miscela ottimizzata sia più degradato rispetto a quello estratto con il kit commerciale.

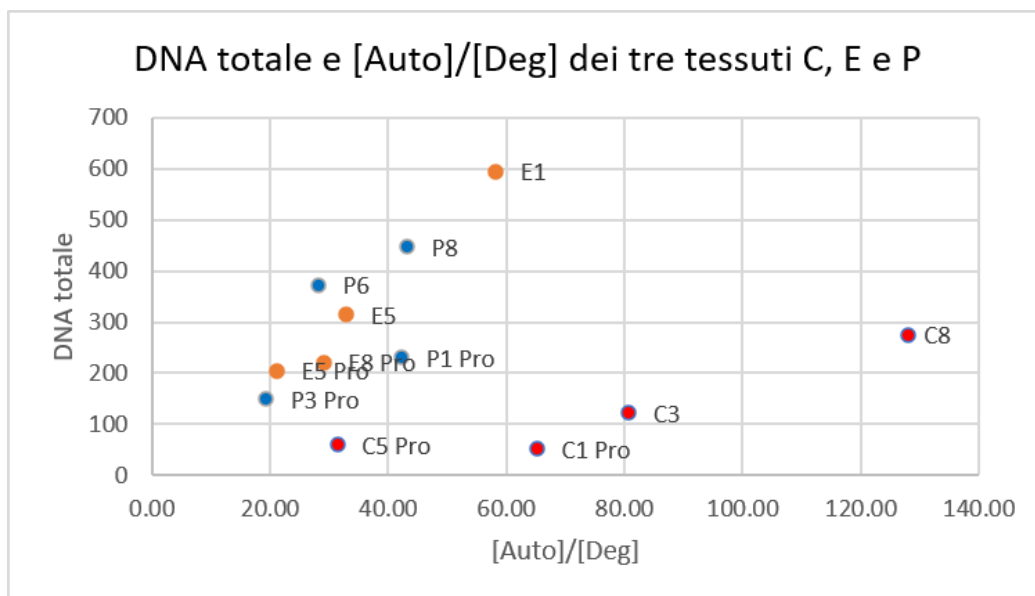


Figura 35| Rappresentazione grafica a due variabili: DNA totale estratto (ng) e valori di $[Auto]/[Deg]$. Sono riportati quattro campioni per ognuno dei tre tessuti. I campioni di tessuto cardiaco (rosa), tessuto encefalico (arancione) e tessuto polmonare (azzurro). Per ognuno dei tre tessuti sono presenti due campioni estratti con la miscela ottimizzata (senza suffisso pro) e due estratti con il kit commerciale (con suffisso pro). I campioni di tessuto cardiaco mostrano valori di $[Auto]/[Deg]$ alti, maggiori rispetto agli altri due tessuti. Il tessuto encefalico e il tessuto polmonare mostrano valori più elevati di DNA totale estratto e valori relativamente bassi di $[Auto]/[Deg]$.

4.5 Valutazione della qualità dei profili genetici

La qualità di un profilo genetico è strettamente correlata alla quantità di DNA utilizzato per la reazione di *PCR* e alla sua qualità, in termini di maggiore o minore degradazione e/o modificazione della struttura primaria. In presenza di DNA degradato, è possibile che, nel corso della reazione di amplificazione enzimatica, si originino gli artefatti di amplificazione già in precedenza descritti nell'Introduzione a questo elaborato di tesi. Se il DNA è danneggiato nella sua struttura primaria, sarà difficile ottenere una tipizzazione per tutti i marcatori, in particolare per quelli ad alto peso molecolare, producendo così profili genetici parziali, ove non tutti i marcatori del pannello utilizzato forniscono un genotipo. L'analisi di campioni biologici, quali quelli selezionati e oggetto del presente studio, potrebbero presentare, quindi, complessità interpretative, dovute alle peculiari caratteristiche di degradazione dei tessuti inclusi in paraffina. Al fine di valutare la qualità dei profili genetici, sono stati selezionati 12 campioni riportati nella sottostante tabella, precedentemente analizzati nella valutazione comparativa delle metodiche, selezionati secondo i criteri qui di seguito riportati (Tabella 34).

Campione	Tessuto	Degradation index ([Auto]/[Deg])	Tipo di estrazione
C3	Cardiaco	80.74	Ottimizzata
C5 pro	Cardiaco	31.37	Kit
C8	Cardiaco	127.98	Ottimizzata
C1 pro	Cardiaco	65.17	Kit
E5	Encefalico	32.88	Ottimizzata
E5 pro	Encefalico	21.15	Kit
E1	Encefalico	58.24	Ottimizzata
E8 pro	Encefalico	29.02	Kit
P6	Polmonare	28.18	Ottimizzata
P3 pro	Polmonare	19.35	Kit
P8	Polmonare	43.17	Ottimizzata
P1 pro	Polmonare	42.29	Kit

Tabella 34 | Rappresentazione dei 12 campioni estratti, selezionati per la valutazione qualitativa dei profili genetici.

I 12 campioni analizzati sono stati suddivisi in base al metodo di estrazione, al tessuto biologico e agli score del *Degradation Index*, cioè alla qualità del DNA. In relazione a questo ultimo parametro, sono stati selezionati i campioni dei tre tessuti estratti con le due differenti metodiche che presentavano rispettivamente il più alto indice di degradazione (quindi i campioni con DNA qualitativamente peggiore, in arancione nella Tabella 34) e quelli con il DI più basso (quelli con DNA qualitativamente migliore, in verde nella Tabella 34).

Tale criterio è stato adottato per valutare, in parallelo per le due metodiche di estrazione, la qualità dei profili genetici ottenuti amplificando il DNA più degradato e quello più integro estratto dai tre differenti tessuti considerati. Al fine di poter comparare i profili genetici, è stata normalizzata la quantità di DNA input per l'amplificazione in *PCR* dei 21 marcatori STR autosomici contenuti in un kit commerciale per la determinazione del profilo genetico, il kit *GlobalFiler™ PCR Amplification* (ThermoFisher Scientific). 1 ng di DNA è stato amplificato per tutti i campioni; gli amplificati sono stati successivamente sottoposti ad elettroforesi capillare su un sequenziatore automatico. I risultanti profili sono stati valutati utilizzando il software *GeneMapper ID-X v 1.6* settando una soglia analitica (*AT*) a 50 *RFU*.

I genotipi di riferimento per i 21 marcatori caratterizzati per il soggetto da cui sono stati prelevati i tessuti biologici utilizzati per la presente fase sperimentale (campione FFPE 3) sono stati determinati dalla comparazione dei risultati di tutte le 12 amplificazioni effettuate. Per la gran parte dei marcatori, è stato utilizzato il metodo del profilo *consenso*, cioè genotipi concordi in più della metà delle tipizzazioni ottenute (Gill P. et al., 2000). Per i due marcatori per cui non era possibile ottenere un profilo di questo tipo, i genotipi sono stati determinati con il metodo *composito*, considerando solo gli alleli presenti in misura quantitativamente superiore (Gill P. et al., 2000).

Dalla disamina dei profili sono stati, quindi, individuati i fenomeni artefattuali causati dalla degradazione/modificazione della struttura primaria del DNA. Sono stati valutati: il numero di marcatori completi tipizzati, i marcatori eterozigoti affetti da *Drop-out* di un allele, il numero di alleli sovrannumerari (*Drop-in*) individuati e i picchi *stutter* più alti rispetto al loro normale valore soglia. Tutti questi dati sono riportati nella Tabella 35.

Campione	Marcatori tipizzati	Allele Drop-out	Allele Drop-in	Stutter sopra soglia
C3	19 (2)	2	4	0
C5 pro	21 (0)	4	5	1
C8	18 (3)	4	1	0
C1 pro	20 (1)	2	3	1
E5	20 (1)	2	0	0
E5 pro	20 (1)	1	0	1
E1	21 (0)	1	0	2
E8 pro	21 (0)	1	2	1
P6	21 (0)	2	1	0
P3 pro	20 (1)	1	4	1
P8	19 (2)	0	1	1
P1 pro	21 (0)	3	1	1

Tabella 35 | Tabella riassuntiva delle caratteristiche dei profili genetici ottenuti dall'amplificazione dei 12 campioni selezionati. Sono riportati il numero di marcatori tipizzati (tra parentesi i loci privi di tipizzazione, ovvero i locus-dropout), il numero dei *Drop-out allelici*, il numero di *Allele Drop-in* e le *stutter* sopra soglia.

Qui di seguito, vengono mostrati, per confronto, i migliori e i peggiori profili genetici ottenuti relativamente ad ogni tessuto analizzato, indipendentemente dal metodo di estrazione e dal valore di *Degradation Index*. Questo al fine di esplicitare quanto sopra riportato in relazione alla valutazione della qualità dei profili. Nel dettaglio, si apprezza la maggiore completezza dei profili (minore presenza di *locus out*, LO), la maggiore altezza dei picchi (scala verticale nei profili), la minore presenza di artefatti quali il *Drop-out* allelico (DO), il *Drop-in* allelico (DI) e le *stutter* sopra soglia in nero (ST) nei profili migliori (in alto nelle figure 36, 37 e 38). In generale, in tutti e tre i tessuti sono presenti artefatti dovuti alla bassa qualità del DNA ad alto peso molecolare. Gli artefatti sono presenti tuttavia sia nei campioni estratti con la miscela sperimentale che quelli estratti in maniera automatizzata con il kit (campioni denominati pro). La degradazione generale del DNA di tutti i campioni è facilmente comprensibile anche visualizzando i tracciati elettroforetici poiché è evidente come l'altezza dei picchi allelici sulla scala delle *RFU* (asse Y) decresca all'aumentare delle dimensioni degli alleli. I marcatori con alleli a peso molecolare alto, a causa della degradazione del DNA che ne riduce le dimensioni, presentano alleli con altezze inferiori rispetto a quelli a basso peso molecolare (a sinistra nel tracciato). Infatti, tutti gli eventi di *Drop-out* allelico, in tutti i tessuti e in entrambe le modalità di estrazione, avvengono in *loci* ad alto peso molecolare, da circa 250 bp in poi.

Dai tracciati elettroforetici dei campioni riportati in Figura 36, 37 e 38 si può notare inoltre una generale differenza in termini di numero di marcatori tipizzati tra i diversi tessuti. Nel tessuto cardiaco il numero di marcatori (tra quelli riportati) completamente tipizzati è minore rispetto a quelli analoghi del tessuto encefalico e polmonare che sono simili. Ciò indica che il tessuto cardiaco presenta una minore qualità di DNA generale indipendentemente dalla miscela di estrazione utilizzata.

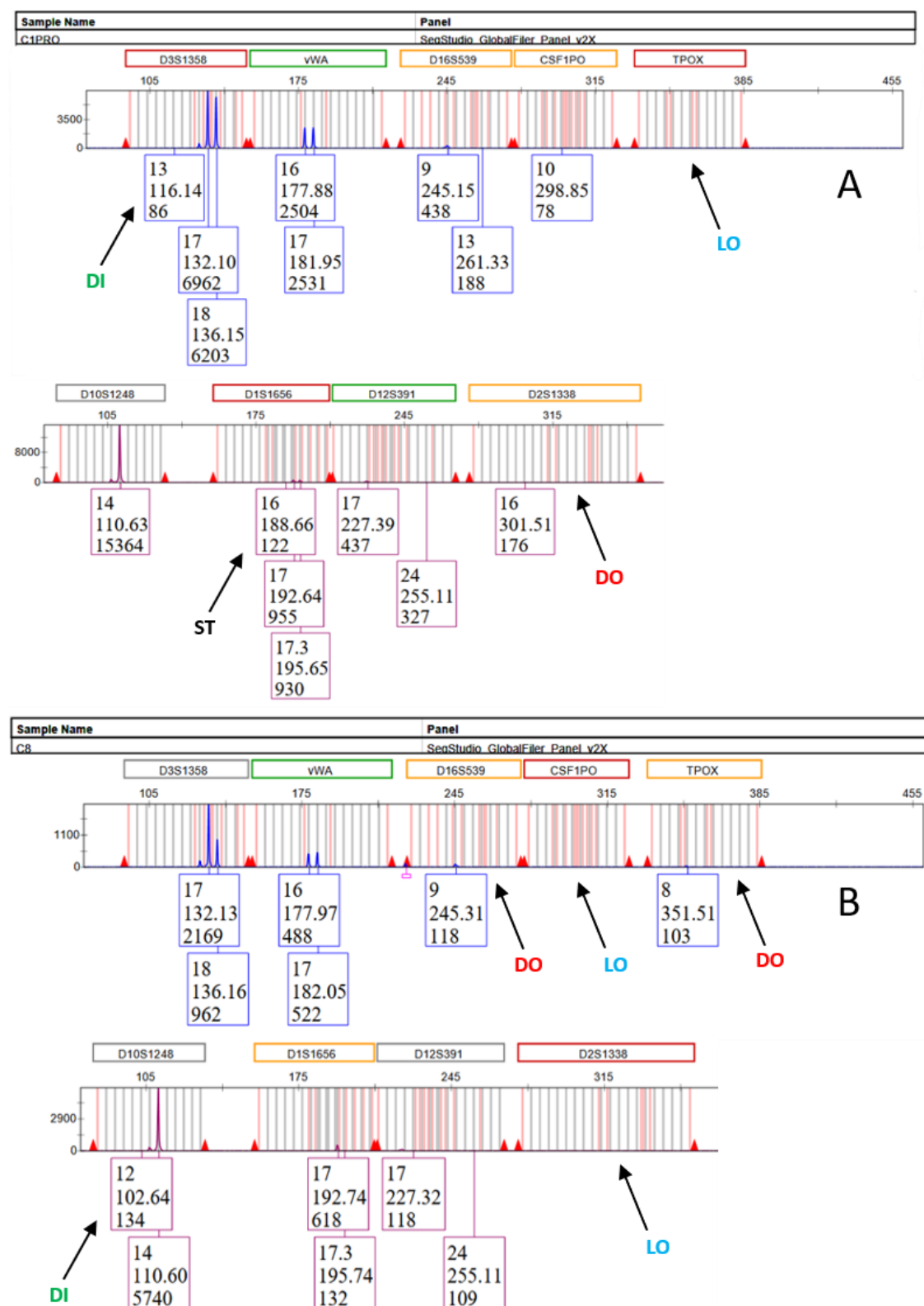


Figura 36 Confronto tra i profili genetici di due campioni di tessuto cardiaco. Nello specifico, sono riportati, il migliore (A) e il peggiore (B) in termini di numero di marcatori tipizzati. Gli artefatti individuati sono indicati dalle frecce: in rosso i *Drop-out* allelici (DO), in verde i *Drop-in* allelici (DI), in nero le *stutter* (ST) sopra soglia e in azzurro i *Locus-out* (LO), cioè l'assenza completa di genotipo. L'altezza dei picchi allelici del profilo A è generalmente maggiore rispetto a quella del profilo B.

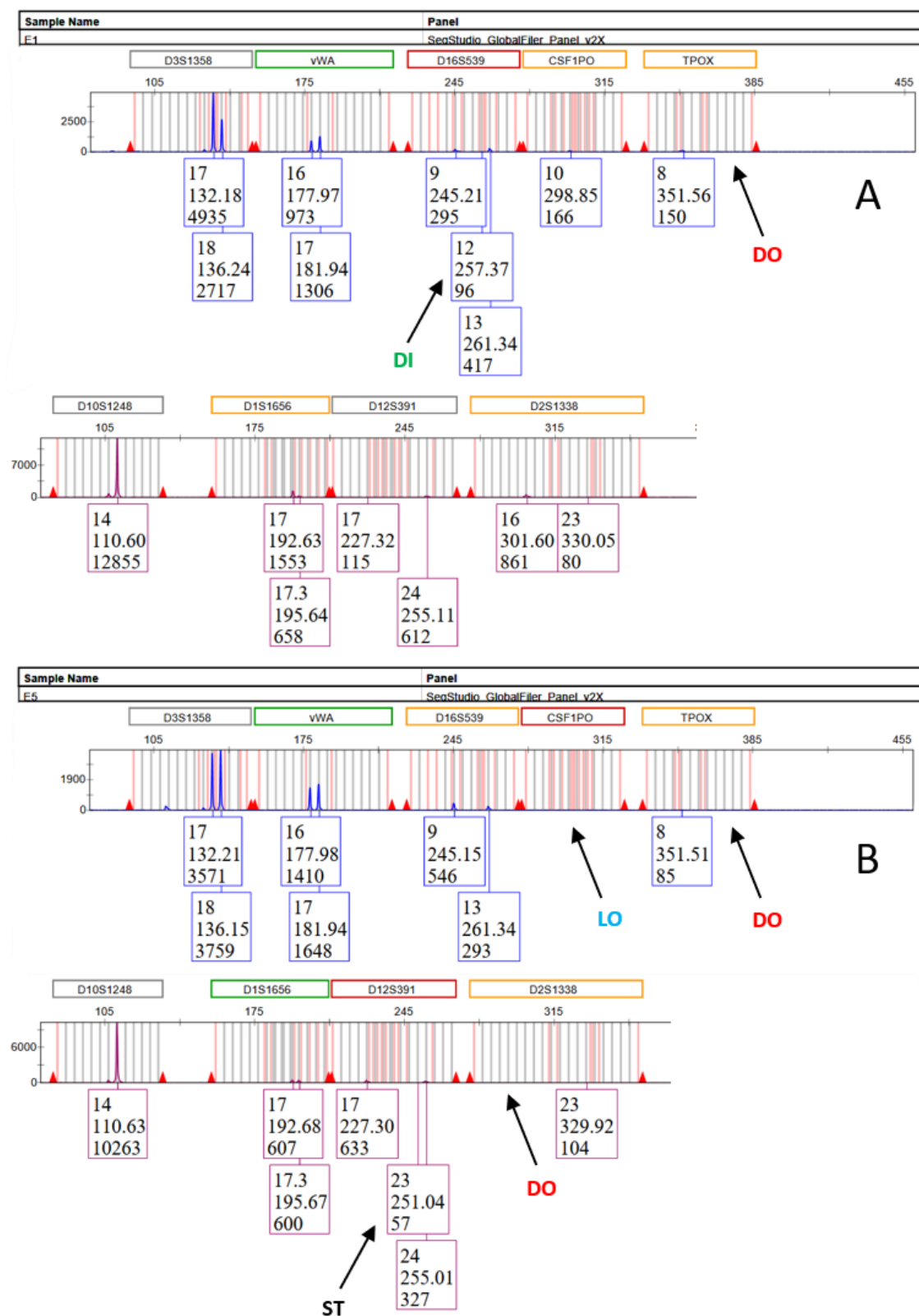


Figura 37 Confronto tra i profili genetici di due campioni di tessuto encefalico. Nello specifico, sono riportati, il migliore (A) e il peggiore (B) in termini di numero di marcatori tipizzati. Gli artefatti individuati sono indicati dalle frecce: in rosso i *Drop-out* allelici (DO), in verde i *Drop-in* allelici (DI), in nero le *stutter* (ST) sopra soglia e in azzurro i *Locus-out* (LO), cioè l'assenza completa di genotipo. L'altezza dei picchi allelici del profilo A è generalmente maggiore rispetto a quella del profilo B.

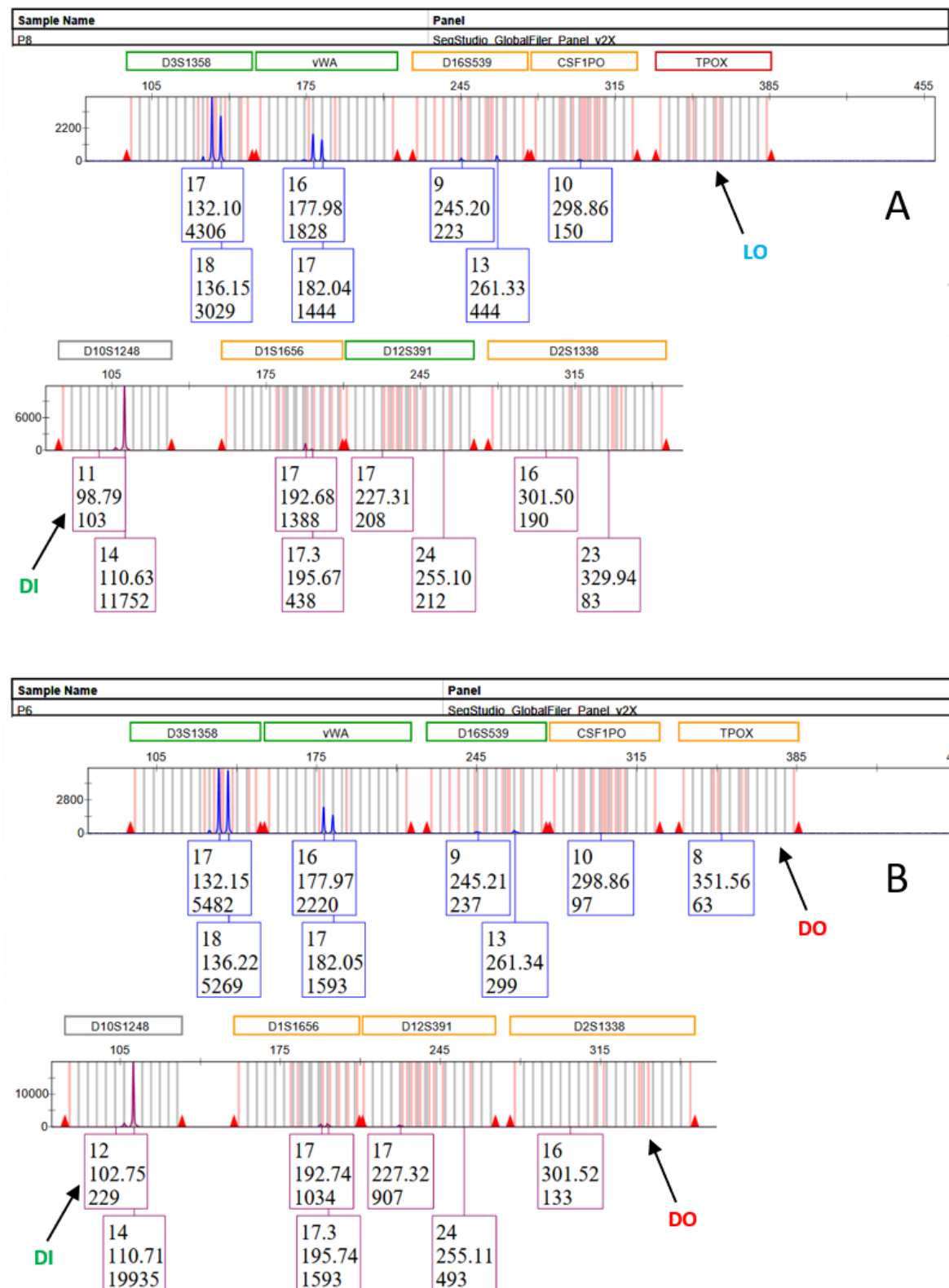


Figura 38| Confronto tra i profili genetici di due campioni di tessuto polmonare. Nello specifico, sono riportati, il migliore (A) e il peggiore (B) in termini di numero di marcatori tipizzati. Gli artefatti individuati sono indicati dalle frecce: in rosso i *Drop-out* allelici (DO), in verde i *Drop-in* allelici (DI), in nero le *stutter* (ST) sopra soglia e in azzurro i *Locus-out* (LO), cioè l'assenza completa di genotipo. L'altezza dei picchi allelici del profilo A è generalmente maggiore rispetto a quella del profilo B.

Dopo aver valutato i tracciati elettroforetici è stata condotta un'analisi multivariata dei dati tramite l'Analisi delle Componenti Principali (PCA) con tre scopi principali: diminuire la dimensionalità dei dati, individuare le relazioni tra le variabili e valutare l'impatto dei metodi di estrazione sui diversi tessuti. Nel *dataset* sono state misurate sette variabili distinte per ogni campione. La prima è il *Degradation slope* (DS) che è stato calcolato per ogni profilo utilizzando il software *EuroForMix*. Il *Degradation slope* è un valore che indica la qualità di un profilo genetico poiché esprime in maniera precisa l'andamento del profilo, cioè quando il profilo decresca in termini di altezza dei picchi allelici del tracciato all'aumentare del peso molecolare dei marcatori. Più è vicino all'unità il valore, più il profilo genetico è di buona qualità. Sono stati valutati anche il valore di [Auto]/[Deg] (AD), il numero di marcatori che hanno fornito un genotipo completo (M), il numero di *Drop-out* (Dout), il numero di *Drop-in* (Din), il numero di *stutter* sopra soglia (S) e la somma dei *Drop-in* e le *stutter* sopra soglia (DinSt). L'introduzione del parametro combinato DinSt si è resa necessaria a causa dei limiti intrinseci del software *GeneMapper*, il quale rileva ed etichetta indistintamente come "alleli" tutti i picchi che superano la soglia "stutter" impostata. Ciò rende spesso complessa la discriminazione univoca e certa tra veri e propri eventi di *Drop-in* e *stutter*, rendendo la variabile DinSt un indicatore cumulativo altamente rappresentativo dei segnali sovrannumerari complessivi.

La PCA consente di studiare questa mole di dati, descritti in uno spazio a sette dimensioni (una per ciascuna variabile), in poche dimensioni sintetiche, e queste sono le Componenti Principali. La componente principale 1, PC1 descrive il 42.1 % della varianza di tutti i dati. La componente principale 2, PC2, racchiude il 31.3 % della varianza. Insieme, PC1 e PC2 spiegano il 73.4 % di tutta l'informazione contenuta nei dati (Figura 39). Il restante 26.6 % della varianza è distribuito tra le restanti componenti principali PC3, PC4 e PC5 e contribuiscono in modo rilevante a definire fluttuazioni residuali (rumore di fondo) e sfumature secondarie tra i campioni che non sono descritte dal quadro generale dei risultati delle prime due componenti principali. L'analisi dei dati proiettati nello spazio PC1 e PC2 spiega oltre il 73 % della varianza dei dati. Tale potere esplicativo conferma che la proiezione bidimensionale descrive gran parte della variabilità biologica e tecnica dei campioni, validando le osservazioni sulla distinzione tra i tessuti e tra i metodi di estrazione.

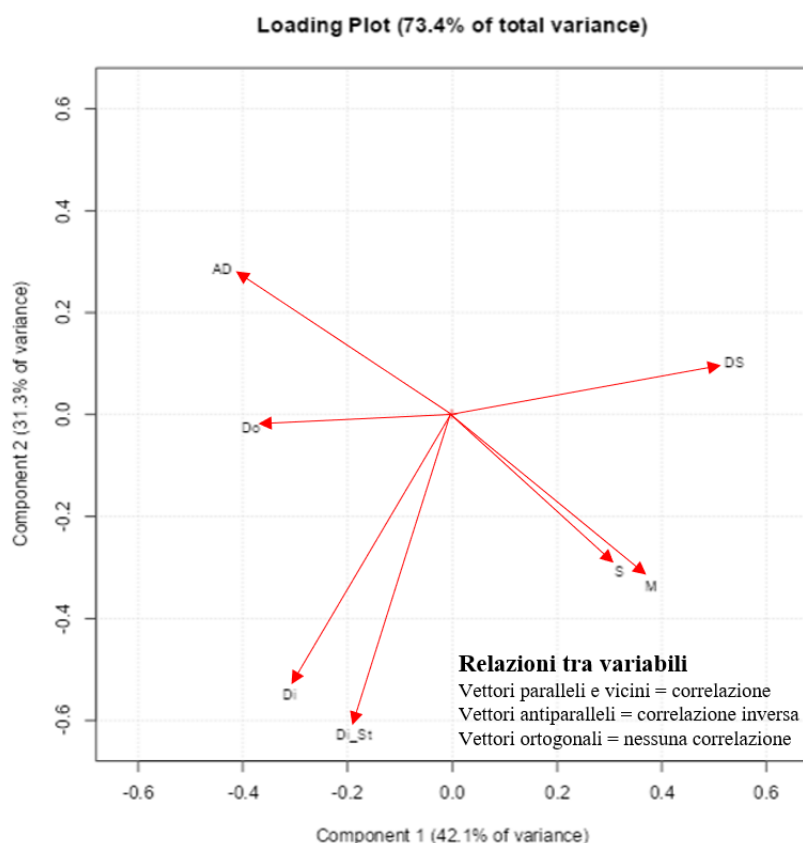


Figura 39| Rappresentazione grafica (*Loading Plot*) delle sette variabili proiettate nello spazio delle prime due componenti principali, PC1 e PC2. Le variabili rappresentate sono: *Drop-out* allelico (Do), *Drop-in* allelico (Di), *Stutter* (S), *Drop-in + Stutter* (Di_St), numero di marcatori (M), *Degradation slope* (DS), *Degradation index* o valore [Auto]/[Deg] (AD).

Dalla matrice di correlazione dei dati emergono relazioni fondamentali tra le variabili analizzate. Il Primo Componente PC1 è fortemente influenzato da due gruppi di variabili che agiscono in direzioni opposte: a valori alti positivi di PC1 si nota il *Degradation Slope* (DS) (+0.5394) e a valori leggermente inferiori il numero dei marcatori (M) (+ 0.3796) e il numero di *Stutter* sopra soglia (S) (+ 0.3202); a valori negativi di PC1 invece ci sono [Auto]/[Deg] (AD) (-0.4348), *Allele Drop-out* (Do) (0.3803), *Allele Drop-in* (Di) (-0.3073) e la variabile *Drop-in + Stutter* (Di_St) (-0.1848), e questa ultima è fortemente correlata per sua costruzione con *Drop-in*. Questo quadro significa che se un campione ha un punteggio alto nel PC1, avrà valori alti di DS e nel numero di marcatori e bassi valori di Di, Do e Di_St. Il Secondo Componente PC2 separa caratteristiche diverse: a valori fortemente negativi isola l'influenza delle variabili *Drop-in* (Di) (-0.5498) e *Drop-in + Stutter* (Di_St) (- 0.6295), a valori poco più positivi sono presenti le variabili M (-0.3365) ed S (-0.3078). Mentre a valori positivi di PC2 compaiono le variabili DS (+0.1039) e AD (+0.2863). Le variabili Marcatori (M) e numero di *Stutter* (S) mostrano pesi fortemente positivi lungo l'asse PC1 e sono tra di loro

direttamente correlate. Ciò significa che più è alto il numero di marcatori tipizzati più *Stutter* sopra soglia vengono individuate dal software. I *Drop-out* (Do) e l'indice [Auto]/[Deg] (AD) presentano dei pesi nettamente negativi lungo PC1. Ciò significa che quando il valore di [Auto]/[Deg] è basso (cioè quando il DNA è poco degradato) vi è un numero limitato di dropout allelici. La contrapposizione quasi antiparallela tra i vettori M e AD evidenzia una forte correlazione inversa: all'aumentare dell'indice di degradazione del DNA (AD), si verifica una contestuale riduzione del numero di marcatori completamente profilati (M). Lo stesso tipo di correlazione antiparallela è presente tra il numero dei *Drop-out* e i valori di *Degradation slope*, questo è coerente con il fatto che all'aumentare il valore di DS diminuiscono i *Drop-out* allelici e quindi il profilo è più completo. *Drop-in* (Di) e la variabile combinata (Di_St): sono caratterizzati da carichi fortemente negativi lungo l'asse PC2 e sono correlate positivamente tra di loro. Il passaggio dai campioni estratti con metodo ottimizzato a quelli estratti con il kit sposta i valori delle coordinate dei campioni lungo PC2 verso punteggi negativi. Lo spostamento è correlato con le variabili Di e Di_St: quindi il passaggio dall'estrazione con il metodo sperimentale a quella con il metodo commerciale corrisponde a valori di Di e Di_St più elevati. Dall'analisi inoltre PC1 separa i campioni in tre gruppi distinti (Figura 40).

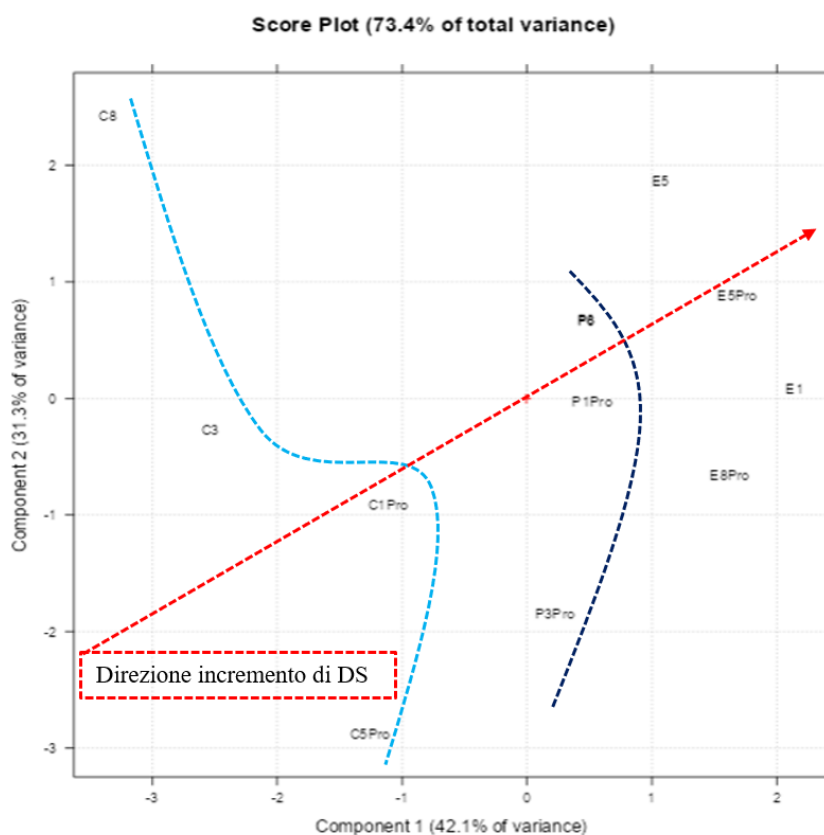


Figura 40 | Rappresentazione grafica (*Score Plot*) dei tre gruppi distinti di tessuti valutati secondo le sette variabili sulla base della PC1 e PC2. (I campioni P6 e P8 hanno score identici e nel grafico sono sovrapposti).

Dallo *Score Plot* in Figura 40, emerge chiaramente che la qualità dei profili genetici dipende fortemente dal tessuto di origine. Si può vedere che i campioni del gruppo C, che rappresentano i tessuti cardiaci, si raggruppano in valori negativi di PC1 (inferiori o uguali a -1), è il tessuto che presenta le maggiori criticità in termini di degradazione con AD alto, marcatori ridotti e dropout elevato. I campioni del gruppo E (tessuto encefalico) hanno valori positivi di PC1 dimostrando una qualità del DNA e una percentuale di tipizzazione nettamente superiori ai tessuti cardiaci. I campioni del gruppo P (tessuto polmonare) occupano una zona intermedia tra C ed E con valori positivi di PC1. La PCA permette di verificare graficamente se e come il metodo di estrazione ottimizzato differisca da quello automatizzato tramite kit. Inoltre, come si vede nel grafico (Figura 40) il *Degradation slope* aumenta nella direzione che va dai tessuti cardiaci ai tessuti encefalici.

Nel tessuto cardiaco, più degradato, tutti i quattro campioni, si posizionano nella parte sinistra del grafico in Figura 40, a valori di PC1 negativi caratteristici di un numero elevato di Do, Di, Di_St e valori elevati di AD. I campioni estratti con la miscela ottimizzata come C3 e C8 si posizionano in alto, a valori positivi di PC2, perché hanno valori di AD elevati a contrario dei corrispettivi estratti con kit C5 Pro e C1 Pro che si collocano più in basso perché hanno valori AD più bassi della media e allo stesso tempo valori alti di Di_St e M. , Questo indica che il metodo automatizzato ha garantito una stabilità o una qualità del profilo leggermente superiore su questo tessuto critico in termini di degradazione. I tessuti encefalico e polmonare mostrano invece prestazioni comparabili, posizionandosi nella medesima area di alta qualità complessiva.

Il *Degradation slope* è l'indice di degradazione più accurato del profilo genetico del campione in quanto tiene conto delle altezze degli alleli lungo tutto il l'elettroferogramma. Nella figura 41 si possono vedere le differenze significative di *Degradation slope* tra i diversi tessuti.

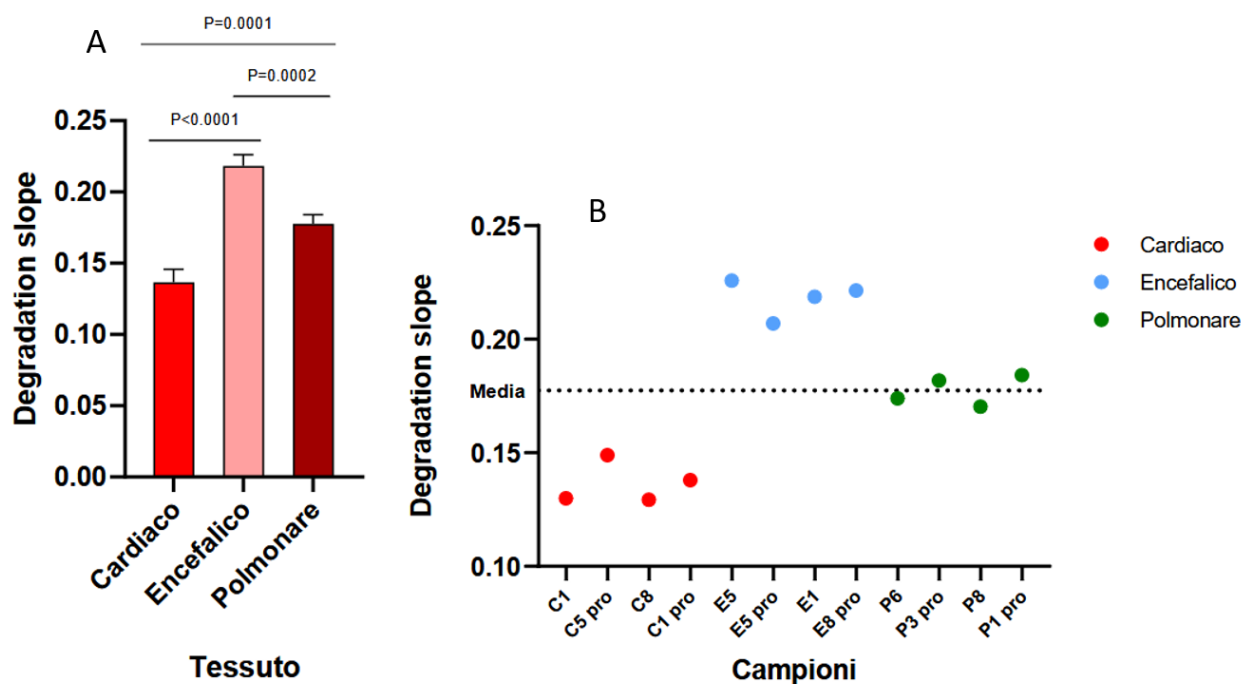


Figura 41| A) Istogramma che mostra le differenze del valore di *Degradation slope* tra i tre diversi tessuti; B) Scatterplot che mostra la distribuzione dei valori di *Degradation slope* per ogni campione, tessuto cardiaco (C) in rosso, encefalico (E) in azzurro e polmonare (P) in verde.

Come si può evincere in Figura 41, il tessuto encefalico mostra valori di *Degradation slope* significativamente più elevati (media 0.2182, mediana 0.2200) rispetto a quelli degli altri tessuti, confermandosi la matrice biologica caratterizzata dalla migliore qualità del DNA e della minore frammentazione rispetto al tessuto cardiaco e a quello polmonare. Al contrario, il tessuto cardiaco ha registrato i valori più bassi (media 0.1365, mediana 0.1339), evidenziando una qualità relativa meno elevata, con differenze significative anche rispetto al polmone che si colloca in una posizione intermedia (media 0.1776, mediana 0.1779). L'analisi, inoltre, dei singoli campioni nello Scatterplot (Figura 41 B) conferma una spiccata riproducibilità all'interno di ciascun gruppo e una netta stratificazione dei tre distretti tissutali rispetto alla media generale dei dati (0.1774). La migliore qualità riscontrata nel tessuto encefalico rispetto al cardiaco evidenzia come i diversi tessuti rispondano in modo differente ai processi di fissazione in formalina, inclusione e lisi enzimatico-chimica. Tale tessuto ha consentito in questo contesto una migliore preservazione o resa delle molecole di DNA ad alto peso molecolare, riducendo l'impatto della frammentazione chimica sui loci a dimensioni maggiori e generando dei profili qualitativamente migliori.

Questi risultati confermano l'esistenza di una diversa qualità di DNA intrinseca ai tessuti al di là della modalità di estrazione utilizzata, dimostrando come la scelta della matrice d'origine rappresenti una variabile critica per l'ottimizzazione dell'analisi STR in ambito identificativo forense. Occorre precisare che i dati utilizzati per questo studio comparativo si riferiscono a matrici tissutali differenti prelevate da un singolo soggetto. Le prospettive future di questo lavoro prevedono l'estensione dell'analisi a una coorte più ampia di campioni (circa 20 casi autoptici), con l'obiettivo di valutare l'impatto della variabilità inter-individuale oltre a quella intra-individuale già documentata.

5 CONCLUSIONI

Il presente elaborato di tesi si è proposto di affrontare le complessità analitiche legate al recupero di DNA da tessuti biologici inclusi in paraffina prelevati da soggetti sottoposti ad autopsia, a fini di identificazione personale. L'attività sperimentale si è articolata in tre fasi.

La prima fase ha avuto come scopo l'ottimizzazione di una miscela di estrazione del DNA convenzionalmente utilizzata nel laboratorio di Genetica forense, attraverso un approccio di *Design of Experiments (DoE)* basato su matrici ortogonali di *Plackett-Burman* e l'analisi di regressione lineare multipla. Questo approccio ha consentito di modificare la formulazione "originale" e ricavare una miscela essenziale ed ottimizzata, contenente solo componenti che aumentano la resa di estrazione e priva di componenti ridondanti o inibitorie. Le analisi sono state condotte su tessuti inclusi in paraffina relativi a tre differenti soggetti sottoposti ad autopsia giudiziaria o a riscontro diagnostico (campioni *FFPE* 1, 2 e 3). Tali campioni presentavano differenti tempistiche per il prelievo dei tessuti e tempistiche di fissazione in formalina. I campioni autoptici (*FFPE* 1 e 3) erano stati prelevati a 4-5 giorni dal decesso del soggetto e sottoposti a fissazione in formalina per un tempo variabile tra 1 e 2 settimane mentre il campione prelevato nel corso del riscontro diagnostico (*FFPE* 2) era stato prelevato a 40 ore dal decesso e fissato per 30 ore in formalina. La selezione di campioni con tali differenti tempistiche era motivata dal fatto di verificare gli effetti della degradazione post-mortale dei tessuti (legata al tempo trascorso dalla morte, cioè al *post-mortem interval*) e della fissazione in formalina (reagente noto per introdurre *cross-link* proteine-acidi nucleici che indeboliscono la struttura primaria del DNA), sul recupero quantitativo e sulla qualità del DNA, anche in termini di valutazione dei risultanti profili genetici.

I parametri di riferimento da monitorare nel corso delle analisi a tale scopo sono stati i valori delle sonde definite [Auto] e [Deg], e il loro rapporto, ottenuti in un saggio di *Real-time PCR*. Mentre la sonda [Auto] amplifica un frammento di DNA di dimensioni limitate (84 bp) la sonda [Deg] caratterizza un frammento di dimensioni maggiori (294 bp). La valutazione della quantità di DNA espressa in ng/μl avviene sulla base della sonda [Auto], in rapporto a standard quantitativi forniti dall'azienda produttrice del kit commerciale *PowerQuant® System* (Promega). Il rapporto [Auto]/[Deg] è indice della qualità del DNA: idealmente in un DNA di ottima qualità tale valore si avvicina a 1 mentre la qualità del DNA diminuisce tanto più il valore del rapporto incrementa.

Le analisi condotte secondo lo schema predisposto dall'*Experimental design* sono state inizialmente condotte su sezioni di tessuto cardiaco provenienti dai tre soggetti selezionati, individuando i componenti della miscela di estrazione del DNA con impatto positivo e quelli con impatto negativo rispetto alla resa quantitativa di DNA e alla sua qualità relativa. Successivamente, si è proceduto a verificare l'esistenza di un possibile impatto sulla miscela di estrazione legato al tipo di tessuto

biologico utilizzato. A tale scopo, l'analisi è stata estesa a differenti matrici biologiche, nello specifico tessuto encefalico e polmonare, prelevate da uno stesso soggetto. I risultati hanno evidenziato la netta prevalenza della matrice anatomica sulla resa finale, con l'encefalo ed il polmone che hanno mostrato rese quantitative significativamente superiori rispetto al tessuto cardiaco di riferimento. La serie di esperimenti condotti utilizzando l'*Experimental design* e l'analisi di regressione lineare multipla ha consentito di identificare la formulazione ottimizzata della miscela di estrazione del DNA composta da 100 mM KCl, 20 mM Tris-HCl, 2 mg/mL di Proteinasi K e 10 minuti di denaturazione termica a 95°C, azzerando ed eliminando alcuni reagenti della miscela convenzionalmente utilizzata nel laboratorio di Genetica Forense, tra i quali il MgCl₂, la Gelatina, l'NP40 e il Tween-20. In conclusione, l'*Experimental design* ha consentito di trasformare una ricetta empirica in un protocollo essenziale, economico e altamente performante. La rimozione dei reagenti superflui ha garantito rese quantitative superiori rispetto a quelle della miscela in precedenza utilizzata.

Sulla scorta di tali risultati, si è proceduto alla seconda fase di analisi legate ad uno studio comparativo delle prestazioni della miscela di estrazione del DNA da FFPE ottimizzata rispetto a quelle di un kit commerciale automatizzato (*Maxwell® RSC DNA FFPE Kit*), a partire da diverse tipologie di tessuti. L'esperimento è stato anch'esso pianificato utilizzando l'*Experimental design* e ha previsto l'analisi comparativa di 48 sezioni istologiche di tessuto cardiaco, encefalico e polmonare, prelevate da un singolo soggetto per eliminare la variabilità biologica inter-individuale, metà delle quali sono state estratte con la metodica ottimizzata e metà con il kit commerciale. I risultati hanno evidenziato che la miscela ottimizzata ha superato in termini di resa totale di DNA in modo significativo il kit commerciale in tutte le matrici analizzate da 1,75 a 3 volte nei differenti tessuti testati. Per contro, il DNA estratto con la metodica ottimizzata ha mostrato una peggiore qualità del DNA, presentando valori di *Degradation index* da 1.5 a 2.5 volte più elevati rispetto a quelli kit.

La terza fase ha, infine, previsto la valutazione della qualità dei profili genetici ottenuti dai campioni sottoposti ad estrazione con le due differenti metodiche. In questo caso, sono stati selezionati per ognuno dei tre tessuti biologici analizzati (cuore, encefalo, polmone) i campioni che presentavano il migliore e il peggiore indice di degradazione del DNA per le due differenti metodiche di estrazione. Da tali campioni è stato determinato il profilo genetico amplificando i marcatori STR contenuti nel kit commerciale *GlobalFiler™* (Promega) ed è stato verificato il numero di marcatori determinati, quelli mancanti (*locus out*), quelli affetti da artefatti di amplificazione quali perdita di un allele del genotipo eterozigote (*allele dropout*) o inserimento di alleli sovrannumerari (*allele dropin*) o altezza anomala del picco accessorio *stutter*. Tali artefatti si originano principalmente a causa di una limitazione quantitativa del DNA o a causa di modificazioni della struttura primaria del DNA e

possono creare ambiguità nell'interpretazione del profilo genetico. Oltre al riscontro di tali artefatti, la qualità dei profili genetici è stata determinata utilizzando il *Degradation slope* parametro che, considerando l'altezza dei picchi allelici, ne valuta la diminuzione in presenza di DNA degradato.

L'analisi dei profili *STR* ottenuti con kit *GlobalFiler™* ha evidenziato che il danno molecolare primario è legato alla fissazione in formalina, con artefatti riscontrati esclusivamente a carico di marcatori ad alto peso molecolare (> 250 bp). Tra i tessuti biologici analizzati, quello cardiaco ha confermato le maggiori criticità in termini di qualità del DNA, mentre encefalo e polmone hanno fornito tracciati qualitativamente superiori. La qualità dei profili è comunque tale da consentire una inequivoca attribuzione ad un soggetto di comparazione, con altissima significatività statistica.

In conclusione, la fase di ottimizzazione con l'*Experimental design* ha consentito di sviluppare un protocollo di estrazione del DNA da tessuti inclusi in paraffina (*FFPE*) semplice, in singola provetta, performante ed economico. Questo protocollo ha dimostrato migliori rese quantitative di DNA più alte rispetto a quelle ottenute con un kit commerciale e un estrattore automatico di DNA. Anche se quest'ultimo approccio fornisce un DNA di migliore qualità rispetto al metodo ottimizzato, sulla base del *Degradation Index* e *Degradation Slope*, tuttavia questa differenza qualitativa non risulta essere significativa confrontando i profili genetici da questi campioni determinati.

In prospettiva, le analisi comparative tra i due metodi di estrazioni verranno ampliate selezionando ulteriori campioni inclusi in paraffina di differenti soggetti, per esplorare una eventuale variabilità inter-individuale e confermare o meno i risultati ottenuti in questo set di esperimenti.

6 BIBLIOGRAFIA E SITOGRAFIA

- Alaeddini, Reza et al. "Forensic implications of genetic analyses from degraded DNA--a review." *Forensic science international. Genetics* vol. 4,3 (2010): 148-57. doi:10.1016/j.fsigen.2009.09.007;
- Amankwaa, Aaron Opoku, and Carole McCartney. "The effectiveness of the UK national DNA database." *Forensic science international. Synergy* vol. 1 45-55. 21 Mar. 2019, doi:10.1016/j.fsisyn.2019.03.004;
- Amory, Christina et al. "COMBO: An AmpliSeq-based trilogy for autosomal and uniparental (Y & mito) biogeographical ancestry and appearance testing." *Forensic science international. Genetics* vol. 81 (2026): 103349. doi:10.1016/j.fsigen.2025.103349;
- Amosova, Olga et al. "Self-catalyzed site-specific depurination of guanine residues within gene sequences." *Proceedings of the National Academy of Sciences of the United States of America* vol. 103,12 (2006): 4392-7. doi:10.1073/pnas.0508499103;
- Antheunisse, J. "Decomposition of nucleic acids and some of their degradation products by microorganisms." *Antonie van Leeuwenhoek* vol. 38,3 (1972): 311-27. doi:10.1007/BF02328101;
- Anzai, N et al. "Types of nuclear endonuclease activity capable of inducing internucleosomal DNA fragmentation are completely different between human CD34+ cells and their granulocytic descendants." *Blood* vol. 86,3 (1995): 917-23;
- Arya, Mani et al. "Basic principles of real-time quantitative PCR." *Expert review of molecular diagnostics* vol. 5,2 (2005): 209-19. doi:10.1586/14737159.5.2.209;
- Bacher, J. W., et al. (1999). "Pentanucleotide repeats: highly polymorphic genetic markers displaying minimal stutter artifact". *Proceedings of the ninth international symposium on human identification* (pp. 24–37). Madison, Wisconsin: Promega Corporation;
- Barry, M A, and A Eastman. "Identification of deoxyribonuclease II as an endonuclease involved in apoptosis." *Archives of biochemistry and biophysics* vol. 300,1 (1993): 440-50. doi:10.1006/abbi.1993.1060;
- Bauer, Sabine et al. "Optimized High Quality DNA Extraction from Formalin-Fixed Paraffin-Embedded Human Atherosclerotic Lesions." *Journal of visualized experiments : JoVE*,168 10.3791/61452. 12 Feb. 2021, doi:10.3791/61452;
- Britten, R J, and D E Kohne. "Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms." *Science (New York, N.Y.)* vol. 161,3841 (1968): 529-40. doi:10.1126/science.161.3841.529;
- Brookes, A J. "The essence of SNPs." *Gene* vol. 234,2 (1999): 177-86. doi:10.1016/s0378-1119(99)00219-x;
- Bruskov, Vadim I et al. "Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA." *Nucleic acids research* vol. 30,6 (2002): 1354-63. doi:10.1093/nar/30.6.1354;
- Bruskov, Vadim I et al. "Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA." *Nucleic acids research* vol. 30,6 (2002): 1354-63. doi:10.1093/nar/30.6.1354;
- Budowle, B., et al. (1998). "CODIS and PCR-based short tandem repeat loci: law enforcement tools". *Proceedings of the second European symposium on human identification* (pp. 73–88). Madison, Wisconsin: Promega Corporation;
- Budowle, Bruce et al. "Validity of low copy number typing and applications to forensic science." *Croatian medical journal* vol. 50,3 (2009): 207-17. doi:10.3325/cmj.2009.50.207;
- Butler, J. M. (2010). "Fundamentals of forensic DNA typing". San Diego: Elsevier Academic Press;
- Butler, J. M., et al. (2008). "Report on ISFG SNP panel discussion. Forensic Science International". *Forensic Science International: Genetics Supplement Series*, 1, 471–472;
- Butler, J.M (2011). "Advanced Topics in Forensic DNA Typing: Methodology". Elsevier Academic Press;
- Butler, John M et al. "STRs vs. SNPs: thoughts on the future of forensic DNA testing." *Forensic science, medicine, and pathology* vol. 3,3 (2007): 200-5. doi:10.1007/s12024-007-0018-1;
- Butler, John M. "Short tandem repeat typing technologies used in human identity testing." *BioTechniques* vol. 43,4 (2007): ii-v. doi:10.2144/000112582;

- Carracedo, A., & Lareu, M. V. (1998). "Development of new STRs for forensic casework: criteria for selection, sequencing and population data and forensic validation". *Proceedings of the ninth international symposium on human identification* (pp. 89–107). Madison, Wisconsin: Promega Corporation;
- Chakraborty, R et al. "The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems." *Electrophoresis* vol. 20,8 (1999): 1682-96. doi:10.1002/(SICI)1522-2683(19990101)20:8<1682:AID-ELPS1682>3.0.CO;2-Z;
- Coble, Michael D, and John M Butler. "Characterization of new miniSTR loci to aid analysis of degraded DNA." *Journal of forensic sciences* vol. 50,1 (2005): 43-53;
- Deans, Andrew J, and Stephen C West. "DNA interstrand crosslink repair and cancer." *Nature reviews. Cancer* vol. 11,7 467-80. 24 Jun. 2011, doi:10.1038/nrc3088;
- Durney, Brandon C et al. "Capillary electrophoresis applied to DNA: determining and harnessing sequence and structure to advance bioanalyses (2009-2014)." *Analytical and bioanalytical chemistry* vol. 407,23 (2015): 6923-38. doi:10.1007/s00216-015-8703-5;
- Edwards, A et al. "DNA typing and genetic mapping with trimeric and tetrameric tandem repeats." *American journal of human genetics* vol. 49,4 (1991): 746-56;
- Ellegren, Hans. "Microsatellites: simple sequences with complex evolution." *Nature reviews. Genetics* vol. 5,6 (2004): 435-45. doi:10.1038/nrg1348;
- Fan, Hao, and Jia-You Chu. "A brief review of short tandem repeat mutation." *Genomics, proteomics & bioinformatics* vol. 5,1 (2007): 7-14. doi:10.1016/S1672-0229(07)60009-6;
- Filipski, J et al. "Periodicity of DNA folding in higher order chromatin structures." *The EMBO journal* vol. 9,4 (1990): 1319-27. doi:10.1002/j.1460-2075.1990.tb08241.x;
- Gaffney, E F et al. "Factors that drive the increasing use of FFPE tissue in basic and translational cancer research." *Biotechnic & histochemistry : official publication of the Biological Stain Commission* vol. 93,5 (2018): 373-386. doi:10.1080/10520295.2018.1446101;
- Gavrieli, Y et al. "Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation." *The Journal of cell biology* vol. 119,3 (1992): 493-501. doi:10.1083/jcb.119.3.493;
- Ge, Jianye et al. "Haplotype block: a new type of forensic DNA markers." *International journal of legal medicine* vol. 124,5 (2010): 353-61. doi:10.1007/s00414-009-0400-5;
- Giannakis, C et al. "Ca²⁺/Mg²⁺-dependent nuclease: tissue distribution, relationship to inter-nucleosomal DNA fragmentation and inhibition by Zn²⁺." *Biochemical and biophysical research communications* vol. 181,2 (1991): 915-20. doi:10.1016/0006-291x(91)91278-k;
- Gill, P et al. "A new methods of STR interpretation using inferential logic--development of a criminal intelligence database." *International journal of legal medicine* vol. 109,1 (1996): 14-22. doi:10.1007/BF01369596;
- Gill, P et al. "An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA." *Forensic science international* vol. 112,1 (2000): 17-40. doi:10.1016/s0379-0738(00)00158-4;
- Gill, P et al. "Report of the European DNA profiling group (EDNAP)--towards standardisation of short tandem repeat (STR) loci." *Forensic science international* vol. 65,1 (1994): 51-9. doi:10.1016/0379-0738(94)90299-2;
- Gill, Peter et al. "New multiplexes for Europe-amendments and clarification of strategic development." *Forensic science international* vol. 163,1-2 (2006): 155-7. doi:10.1016/j.forsciint.2005.11.025;
- Gill, Peter, and John Buckleton. "A universal strategy to interpret DNA profiles that does not require a definition of low-copy-number." *Forensic science international. Genetics* vol. 4,4 (2010): 221-7. doi:10.1016/j.fsigen.2009.09.008;
- Heid, C A et al. "Real time quantitative PCR." *Genome research* vol. 6,10 (1996): 986-94. doi:10.1101/gr.6.10.986;
- Higuchi R., (1989). "PCR Technology, Principles and applications for DNA Amplification", Chapter 4. *M stockton Press*;

- Higuchi, R et al. "Kinetic PCR analysis: real-time monitoring of DNA amplification reactions." *Bio/technology (Nature Publishing Company)* vol. 11,9 (1993): 1026-30. doi:10.1038/nbt0993-1026;
- Higuchi, R et al. "Simultaneous amplification and detection of specific DNA sequences." *Bio/technology (Nature Publishing Company)* vol. 10,4 (1992): 413-7. doi:10.1038/nbt0492-413;
- Hill, Carolyn R et al. "Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples." *Journal of forensic sciences* vol. 53,1 (2008): 73-80. doi:10.1111/j.1556-4029.2008.00595.x;
- Hofreiter, M et al. "Ancient DNA." *Nature reviews. Genetics* vol. 2,5 (2001): 353-9. doi:10.1038/35072071;
- Holland, P M et al. "Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase." *Proceedings of the National Academy of Sciences of the United States of America* vol. 88,16 (1991): 7276-80. doi:10.1073/pnas.88.16.7276;
- Hoque, Mainul et al. "Extraction, Purification, and Next-Generation Sequencing (NGS) Analysis of DNA and RNA from Formalin-Fixed and Paraffin-Embedded (FFPE) Tissue." *Methods in molecular biology (Clifton, N.J.)* vol. 2866 (2025): 207-227. doi:10.1007/978-1-0716-4192-7_12;
- Huang, P et al. "Biochemical characterization of the protein activity responsible for high molecular weight DNA fragmentation during drug-induced apoptosis." *Cancer research* vol. 57,16 (1997): 3407-14;
- International Human Genome Sequencing Consortium. "Finishing the euchromatic sequence of the human genome." *Nature* vol. 431,7011 (2004): 931-45. doi:10.1038/nature03001;
- Jäger, Richard. "New Perspectives for Whole Genome Amplification in Forensic STR Analysis." *International journal of molecular sciences* vol. 23,13 7090. 25 Jun. 2022, doi:10.3390/ijms23137090;
- Jeffreys, Alec J. "Genetic fingerprinting." *Nature medicine* vol. 11,10 (2005): 1035-9 doi:10.1038/nm1005-1035;
- Jeffreys, Alec J. "Genetic fingerprinting." *Nature medicine* vol. 11,10 (2005): 1035-9. doi:10.1038/nm1005-1035;
- Jin, L et al. "The exact numbers of possible microsatellite motifs." *American journal of human genetics* vol. 55,3 (1994): 582-3;
- Karkar, Slim et al. "Statistical modeling of STR capillary electrophoresis signal." *BMC bioinformatics* vol. 20,Suppl 16 584. 2 Dec. 2019, doi:10.1186/s12859-019-3074-0;
- Kasai, K et al. "Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science." *Journal of forensic sciences* vol. 35,5 (1990): 1196-200;
- Kazmi Sara et al. "Protocol Development for Yielding High Quality DNA and RNA from Archived Formalin-Fixed Paraffin Embedded Tissues." *Translation: The University of Toledo Journal of Medical Sciences* vol. 11 No. 3 (2023): 2469-6706. doi:10.46570/utjms.2469-6706;
- Kennedy-Darling, Julia, and Lloyd M Smith. "Measuring the formaldehyde Protein-DNA cross-link reversal rate." *Analytical chemistry* vol. 86,12 (2014): 5678-81. doi:10.1021/ac501354y;
- Kimpton, C et al. "Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci." *International journal of legal medicine* vol. 106,6 (1994): 302-11. doi:10.1007/BF01224776;
- Kimpton, C P et al. "Automated DNA profiling employing multiplex amplification of short tandem repeat loci." *PCR methods and applications* vol. 3,1 (1993): 13-22. doi:10.1101/gr.3.1.13;
- Kondo, Natsuko et al. "DNA damage induced by alkylating agents and repair pathways." *Journal of nucleic acids* vol. 2010 543531. 21 Nov. 2010, doi:10.4061/2010/54353;
- Koshiba, M et al. "The effect of formalin fixation on DNA and the extraction of high-molecular-weight DNA from fixed and embedded tissues." *Pathology, research and practice* vol. 189,1 (1993): 66-72. doi:10.1016/S0344-0338(11)80118-4;

- Ku, Chee Seng et al. "The discovery of human genetic variations and their use as disease markers: past, present and future." *Journal of human genetics* vol. 55,7 (2010): 403-15. doi:10.1038/jhg.2010.55;
- Lander, E S et al. "Initial sequencing and analysis of the human genome." *Nature* vol. 409,6822 (2001): 860-921. doi:10.1038/35057062;
- Leardi R., "D-Optimal Designs, in Chemometrics, Encyclopedia of Analytical Chemistry: Applications, Theory, and Instrumentation". Wiley, 2018. <https://doi.org/10.1002/9780470027318.a9646>;
- Lee, L G et al. "Allelic discrimination by nick-translation PCR with fluorogenic probes." *Nucleic acids research* vol. 21,16 (1993): 3761-6. doi:10.1093/nar/21.16.3761;
- Limentani, Giselle B et al. "Beyond the t-test: statistical equivalence testing." *Analytical chemistry* vol. 77,11 (2005): 221A-226A. doi:10.1021/ac053390m;
- Lindahl, T, and B Nyberg. "Heat-induced deamination of cytosine residues in deoxyribonucleic acid." *Biochemistry* vol. 13,16 (1974): 3405-10. doi:10.1021/bi00713a035;
- Lindahl, T, and B Nyberg. "Rate of depurination of native deoxyribonucleic acid." *Biochemistry* vol. 11,19 (1972): 3610-8. doi:10.1021/bi00769a018;
- Lindahl, T. "Instability and decay of the primary structure of DNA." *Nature* vol. 362,6422 (1993): 709-15. doi:10.1038/362709a0;
- Lisman, Dagmara et al. "Forensic DNA Recovery from FFPE Tissue Using the Maxwell® RSC Xcelerate Kit: Optimization, Challenges, and Limitations." *Genes* vol. 16,9 1074. 12 Sep. 2025, doi:10.3390/genes16091074;
- Lu, Kun et al. "Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers." *Journal of the American Chemical Society* vol. 132,10 (2010): 3388-99. doi:10.1021/ja908282f;
- Lygo, J E et al. "The validation of short tandem repeat (STR) loci for use in forensic casework." *International journal of legal medicine* vol. 107,2 (1994): 77-89. doi:10.1007/BF01225493;
- Marafioti, Luparia (2010). "Banca dati del DNA e accertamento penale". *Giuffrè editore*;
- McNulty, Shannon M, and Beth A Sullivan. "Alpha satellite DNA biology: finding function in the recesses of the genome." *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* vol. 26,3 (2018): 115-138. doi:10.1007/s10577-018-9582-3;
- Morozova, Olena, and Marco A Marra. "Applications of next-generation sequencing technologies in functional genomics." *Genomics* vol. 92,5 (2008): 255-64. doi:10.1016/j.ygeno.2008.07.001;
- Mullen, Lauren E et al. "A quantitative method to assess DNA extraction efficiency." *Journal of forensic sciences* vol. 71,3 (2026): 1398-1404. doi:10.1111/1556-4029.70302;
- Mullis, K., et al. (1986). "Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction". *Cold Spring Harbor Symposium on Quantitative Biology*, 51, 263–273;
- Nakamura, M et al. "Purification and characterization of the Ca²⁺ plus Mg²⁺-dependent endodeoxyribonuclease from calf thymus chromatin." *Journal of biochemistry* vol. 89,1 (1981): 143-52. doi:10.1093/oxfordjournals.jbchem.a133175;
- Nicklas, Janice A, and Eric Buel. "Quantification of DNA in forensic samples." *Analytical and bioanalytical chemistry* vol. 376,8 (2003): 1160-7. doi:10.1007/s00216-003-1924-z;
- Palchaudhuri, Rahul, and Paul J Hergenrother. "DNA as a target for anticancer compounds: methods to determine the mode of binding and the mechanism of action." *Current opinion in biotechnology* vol. 18,6 (2007): 497-503. doi:10.1016/j.copbio.2007.09.006;
- Panchal, Nagesh Kishan et al. "PCR Amplifiable DNA from Breast Disease FFPE Section for Mutational Analysis." *Journal of biomolecular techniques : JBT* vol. 31,1 (2020): 1-6. doi:10.7171/jbt.20-3101-001;
- Phillips, C et al. "Building a forensic ancestry panel from the ground up: The EUROFORGEN Global AIM-SNP set." *Forensic science international. Genetics* vol. 11 (2014): 13-25. doi:10.1016/j.fsigen.2014.02.012;

- Puers, C et al. "Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01[AATG]_n and reassignment of alleles in population analysis by using a locus-specific allelic ladder." *American journal of human genetics* vol. 53,4 (1993): 953-8;
- R. L. Plackett, J. P. Burman. "The Design Of Optimum Multifactorial Experiments." *Biometrika*, vol. 33, (June 1946): 305–325, doi:10.1093/biomet/33.4.305;
- Rajske, Scott R., and Robert M. Williams. "DNA Cross-Linking Agents as Antitumor Drugs." *Chemical reviews* vol. 98,8 (1998): 2723-2796. doi:10.1021/cr9800199;
- Rastogi, Rajesh P et al. "Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair." *Journal of nucleic acids* vol. 2010 592980. 16 Dec. 2010, doi:10.4061/2010/592980;
- Saiki, R K et al. "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." *Science (New York, N.Y.)* vol. 239,4839 (1988): 487-91. doi:10.1126/science.2448875;
- Saini, Saurabh, and Prajwal Gurung. "A comprehensive review of sensors of radiation-induced damage, radiation-induced proximal events, and cell death." *Immunological reviews* vol. 329,1 (2025): e13409. doi:10.1111/imr.13409;
- Schneider, P. M. (2009). "Expansion of the European Standard Set of DNA database loci—the current situation". *Profiles in DNA*, 12(1), 6–7;
- Soldati, Giulia et al. "How to combine the exigencies of histopathological and forensic-genetic investigation: a preliminary study on the use of medicolegal FFPE samples for DNA analysis." *Forensic science, medicine, and pathology*, 10.1007/s12024-026-01277-8. 28 May. 2026, doi:10.1007/s12024-026-01277-8;
- Southern, E M. "Detection of specific sequences among DNA fragments separated by gel electrophoresis." *Journal of molecular biology* vol. 98,3 (1975): 503-17. doi:10.1016/s0022-2836(75)80083-0;
- Sparkes, R et al. "The validation of a 7-locus multiplex STR test for use in forensic casework. (II), Artefacts, casework studies and success rates." *International journal of legal medicine* vol. 109,4 (1996): 195-204. doi:10.1007/BF01225518;
- Stanta G. (2011). "Guidelines for molecular analysis in archive tissues" *Springer Heidelberg Dordrecht London New York*. DOI 10.1007/978-3-642-17890-0;
- Sun, X M, and G M Cohen. "Mg(2+)-dependent cleavage of DNA into kilobase pair fragments is responsible for the initial degradation of DNA in apoptosis." *The Journal of biological chemistry* vol. 269,21 (1994): 14857-60;
- Tagliabracci A. et al. (2009). "Introduzione alla genetica forense. Indagini di identificazione personale e di paternità". *Springer-Verlag Italia*;
- Tautz, D. "Notes on the definition and nomenclature of tandemly repetitive DNA sequences." *EXS* vol. 67 (1993): 21-8. doi:10.1007/978-3-0348-8583-6_2;
- Terje Ahlquist & Ragnhild A. Lothe. (2011). "Encyclopedia of Cancer". Microsatellite Instability. *Springer*;
- Tucker, Valerie C et al. "Developmental validation of the PowerPlex(®) ESI 16 and PowerPlex(®) ESI 17 Systems: STR multiplexes for the new European standard." *Forensic science international. Genetics* vol. 5,5 (2011): 436-48. doi:10.1016/j.fsigen.2010.09.004;
- Urquhart, A et al. "Sequence variability of the tetranucleotide repeat of the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus." *Human genetics* vol. 92,6 (1993): 637-8. doi:10.1007/BF00420957;
- Urquhart, A et al. "Variation in short tandem repeat sequences--a survey of twelve microsatellite loci for use as forensic identification markers." *International journal of legal medicine* vol. 107,1 (1994): 13-20. doi:10.1007/BF01247268;
- Van Pelt-Verkuil et al (2008). "Principles and Technical Aspects of PCR Amplification". *Springer*. DOI:10.1007/978-1-4020-6241-4;
- Walsh, P S et al. "Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA." *Nucleic acids research* vol. 24,14 (1996): 2807-12. doi:10.1093/nar/24.14.2807;
- Walsh, Susan et al. "DNA-based eye colour prediction across Europe with the IrisPlex system." *Forensic science international. Genetics* vol. 6,3 (2012): 330-40. doi:10.1016/j.fsigen.2011.07.009;

- Wang X. Et al. “Specific recognition of DNA depurination by a luminescent terbium(III) complex” *Chemical Science* (2013) 4(9):3748. DOI:10.1039/c3sc51781k;
- Weber, J L. “Human DNA polymorphisms and methods of analysis.” *Current opinion in biotechnology* vol. 1,2 (1990): 166-71. doi:10.1016/0958-1669(90)90026-h;
- White, Blánaid et al. “Oscillating formation of 8-oxoguanine during DNA oxidation.” *Journal of the American Chemical Society* vol. 125,22 (2003): 6604-5. doi:10.1021/ja0343252;
- Xavier, Catarina et al. “Evaluation of the VISAGE basic tool for appearance and ancestry inference using ForenSeq® chemistry on the MiSeq FGx® system.” *Forensic science international. Genetics* vol. 58 (2022): 102675. doi:10.1016/j.fsigen.2022.102675;
- Yoca O. U. “Ancient DNA Research: Ongoing Challenges and Contribution to Medical Sciences” *Journal of Basic and Clinical Health Sciences* (2021): 5(2):182-189 DOI:10.30621/jbachs.854258;
- Zhang, Jinpei et al. “DNA quality and STR success rate in different formalin-fixed tissues.” *International journal of legal medicine* vol. 139,3 (2025): 995-1003. doi:10.1007/s00414-024-03391-0;
- https://assets.fishersci.com/TFS-Assets/LSG/manuals/cms_082407.pdf;
- <https://documents.thermofisher.com/TFS-Assets/LSG/manuals/MAN0029969-GlobalFilerIQC-PCRAmpKit-UG.pdf>;
- https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf;
- <https://ita.promega.com/products/forensic-dna-analysis-ce/human-specific-dna-quantitation/powerquant-system/?catNum=PQ5002>;
- <https://ita.promega.com/products/lab-automation/automated-dna-rna-extraction-purification-maxwell/maxwell-csc-instrument/?tabset0=0>;
- <https://ita.promega.com/resources/profiles-in-dna/2011/new-control-dna-for-powerplex-systems/>;
- https://strbase.nist.gov/Information/General_Overview;
- <https://www.thermofisher.com/it/en/home/life-science/sequencing/sanger-sequencing/genetic-analyzers/models/seqstudio.html>;
- <https://www.thermofisher.com/it/en/home/virtual/seqstudio-3d-tour.html>;
- [Maxwell%20RSC%20DNA%20FFPE%20Kit%20TM437.pdf](#);