



**Universitat Rovira i Virgili**

Departament de Bioquímica i Biotecnologia

**Màster en Genètica, Física i Química Forense**

Treball Final de Màster:

**Validació i avaluació del MiSeq FGx™ Forensic  
Genomics System (Illumina) a partir de mostres  
serialment degradades i diluïdes**

Ada Gomar Boscà

**Tutor acadèmic:** Dr. Santiago García Vallvé

**Tutores professionals:** Dra. Lúcia Martínez Priego i Dra. Ana Isabel Gil Tebar

Servei de Seqüenciació i Bioinformàtica (Fisabio)

*Setembre 2024*



## ÍNDIX DE CONTIGUTS

RESUM .....	3
ABSTRACT .....	4
1. INTRODUCCIÓ.....	5
1.1. ADN i variacions en el genoma humà .....	6
1.1.1. Estudi d'STRs en l'àmbit de genètica forense.....	8
1.1.2. Estudi d'SNP en l'àmbit de genètica forense.....	9
1.1.3. Ús del cromosoma Y en genètica forense .....	10
1.1.4. ADN mitocondrial en genètica forense .....	11
1.2. Estat de la qüestió: Electroforesis capil·lar (EC) vs. Seqüenciació paral·lela massiva (MPS).....	12
1.3. Programa de Fosses de la Comunitat Valenciana .....	13
1.4. MiSeq FGx™ Forensic Genomics System .....	15
2. OBJECTIUS .....	15
3. MATERIALS I MÈTODES.....	17
3.1. Presa de mostres.....	17
3.2. Extracció d'ADN .....	17
3.3. Quantificació.....	17
3.4. Experiment 1: Estudi de degradació + dilucions .....	19
3.5. Experiment 2: Estudi amb dilucions .....	20
3.6. Seqüenciació i anàlisi de dades.....	22
4. RESULTATS I INTEPRETACIÓ .....	22
4.1. Experiment 1 .....	22
4.2. Experiment 2.....	28
5. CONCLUSIONS .....	31
6. BIBLIOGRAFIA .....	35

## **RESUM**

Recentment, les tècniques de seqüenciació de nova generació (NGS) han revolucionat el camp de la genètica forense, millorant considerablement els anàlisis en comparació amb els mètodes convencionals. La seqüenciació massivament paral·lela (MPS) proporciona una cobertura extensa, permeten l'anàlisi simultània de múltiples fragments d'ADN, la detecció de variacions genètiques, incloent STRs, SNPs, X-STRs i Y-STRs, i així augmentant significativament la fiabilitat, sensibilitat i la precisió en mostres degradades o limitades, facilitant la resolució de casos forenses complexos.

Per avaluar la capacitat del sistema MiSeq FGx™ Forensic Genomics System (Illumina) en mostres associades a les restes humanes exhumades de fosses de la Guerra Civil Espanyola, que sovint es troben en estat avançat de degradació i fragmentació, amb quantitats mínimes de material genètic disponible, s'han dissenyat dos experiments. Per una banda, l'exposició de mostres de referència a una degradació serial a través de calor, combinada amb dilucions per a determinar el llindar a partir del qual el sistema pot oferir resultats fiables; i per altra, un experiment centrat en la dilució serial, amb diferents inputs d'ADN, amb l'objectiu de validar la precisió del sistema. Els resultats obtinguts han permès establir els primers límits operatius del MiSeq FGx™ en condicions extremes, oferint una validació primerenca per al seu ús en contextos forenses desafians i la seva aplicació en la identificació de les víctimes de la Guerra Civil Espanyola.

**PARAULES CLAU:** Seqüenciació de nova generació (NGS), Genètica Forense, ADN degradat, Seqüenciació massivament paral·lela (MPS), Guerra Civil Espanyola.

## **ABSTRACT**

Recently, next-generation sequencing (NGS) techniques have revolutionized the field of forensic genetics, significantly improving analyses compared to conventional methods. Massively parallel sequencing (MPS) provides extensive coverage, allowing the simultaneous analysis of multiple DNA fragments, the detection of genetic variations, including STRs, SNPs, X-STRs, and Y-STRs, thereby significantly increasing reliability, sensitivity, and precision in degraded or limited samples, facilitating the resolution of complex forensic cases.

To evaluate the capacity of the MiSeq FGx™ Forensic Genomics System (Illumina) on samples associated with human remains exhumed from the mass graves of the Spanish Civil War, which are often found in an advanced state of degradation and fragmentation with minimal amounts of genetic material available, two experiments were designed. On one hand, the exposure of reference samples to serial degradation through heat, combined with dilutions to determine the threshold at which the system can provide reliable results; and on the other hand, an experiment focused on serial dilution, with different DNA inputs, aimed at validating the system's accuracy. The results obtained have allowed the establishment of the initial operational limits of the MiSeq FGx™ under extreme conditions, providing early validation for its use in challenged forensic contexts and its application in the identification of victims of the Spanish Civil War.

**KEY WORKS:** Next-generation sequencing (NGS), Forensic Genetics, Degraded DNA, Massively parallel sequencing (MPS), Spanish Civil War.

## **1. INTRODUCCIÓ**

La Genètica Forense és una especialitat de la genètica aplicada a la medicina legal i la criminologia. S'utilitzen tota una sèrie de tècniques per a poder analitzar l'ADN de restes biològiques per a la investigació i persecució de delictes, en la identificació de cadàvers o de persones desaparegudes, així com en l'ordre civil per a la resolució de processos de filiació. Fins al 1984, la identificació i la individualització d'evidències biològiques es basava en aplicacions d'histologia, microscòpia, immunologia, bioquímica i serologia. Tot i la contribució significativa de marcadors de proteïnes a la justícia criminal, hi havia una desavantatge amb el limitat grau de variabilitat (Udogadi *et al.*, 2020). En aquest any, el genetista britànic Alec Jeffreys desenvolupa la tècnica d'identificació a través d'ADN (amb empremtes digitals), la qual cosa permet que l'àmbit de la genètica forense canviï el seu enfocament inicial i comenci a evolucionar metodològicament al llarg del temps (García i Sánchez, 2018).

Dins els laboratoris d'anàlisi d'ADN es troben freqüentment una sèrie de casuístiques complicades que fan que aquesta disciplina s'enfronti constantment a objectius desafians. Les mostres genètiques de restes biològiques normalment es recullen de l'escena forense amb una qualitat pobre i una quantitat limitada, degut a factors ambientals com el pH, humitat, temperatura, acidesa del sol o activitat enzimàtica (Senst *et al.*, 2022), a la qual cosa se li afegeix l'alta possibilitat de trobar barreges de perfils de més de dos contribuïdors. A més, aquesta degradació dificulta la polimerització i per tant, les mostres son en gran mesura reticents a l'amplificació per PCR (Fattorini *et al.*, 2017), afectant a la correcta tipificació dels genotips.

La identificació genètica, però, no només s'aplica a la investigació criminal, sinó que la generació de perfils d'ADN de restes humanes és essencial també en la identificació de persones desaparegudes o víctimes de desastres en massa. La recuperació de l'ADN s'obté generalment d'ossos i dents, ja que son normalment l'únic material biològic que es manté després d'estar exposat a diferents condicions ambientals, fauna microbiana, esdeveniments traumàtics o inclús a temps *postmortem* elevats (Latham i Miller, 2019). Degut a la dificultat d'obtenir perfils genètics clars dels individus desconeguts, és necessària la comparació amb mostres de referència (especialment de familiars), per a poder comparar l'ADN i confirmar o descartar l'herència genètica.

Tot i les complicacions durant la seqüenciació de perfils derivades de la qualitat de les mostres, no és l'únic factor que produeix pèrdua o errors en la informació, sinó que també hi és present l'efecte estocàstic. Aquest fet es produeix com a resultat de l'atzar en el pas de l'amplificació de quantitats d'ADN escasses o de qualitat deficient i es caracteritza per l'aparició de diferents fenòmens com l'amplificació desproporcionada dels al·lels, que es manifesta com un desequilibri intralocus i/o en situacions severes de drop-out al·lèlic complet, és a dir, pèrdua d'al·lels en un perfil genètic, normalment per ampliació d'un al·lel en detriment de l'altre, detectant-se per tant, falsos homozigots. Altres efectes que també es produeixen serien el drop-out de loci (és a dir la desaparició de tot el marcador), així com guanys al·lèlics (drop-in), que apareixen en un perfil genètic com a conseqüència de contaminacions esporàdiques<sup>1</sup>.

### 1.1. ADN i variacions en el genoma humà

La base genètica de qualsevol ésser humà resideix en el seu genoma, on l'àcid desoxiribonucleic (ADN), realitza la funció de molècula portadora de la informació genètica d'un individu. Aquest compost orgànic s'emmagatzema dins les cèl·lules que componen el cos humà, i a excepció de les cèl·lules sexuals i els glòbuls vermells sense nucli, s'encarreguen de contenir el mateix conjunt d'instruccions genètiques per a la síntesis de proteïnes responsables de la fisiologia i morfologia de teixits i òrgans (Alonso, 2004).

El genoma humà es troba dins les cèl·lules en forma de cromosomes, estructures que contenen una molècula d'ADN. Aquesta filament en forma de cadena doble, el componen un ordre concret de nucleòtids en forma de bases nitrogenades que determinen el tipus d'informació genètica i les instruccions per a la síntesis d'una proteïna dins un segment determinat anomenat gen (d'entre 3000 i 100.000 nucleòtids) (Alonso, 2004; Watson *et al.*, 2018). El genoma humà està repartit en 22 parells de cromosomes autosòmics o somàtics i una parella de cromosomes sexuals, i presenten dos còpies de cada un (anomenats cromosomes homòlegs), un heretat del progenitor masculí i l'altre del femení.

---

<sup>1</sup> Comissió Nacional per a l'ús forense de l'ADN. Ministeri de Justícia. Disponible a: [https://www.mjusticia.gob.es/gl/ElMinisterio/OrganismosMinisterio/Documents/1292428320473-Glosario\\_Recomendaciones\\_Tecnicas\\_Perfiles\\_Mezcla\\_STRs.PDF](https://www.mjusticia.gob.es/gl/ElMinisterio/OrganismosMinisterio/Documents/1292428320473-Glosario_Recomendaciones_Tecnicas_Perfiles_Mezcla_STRs.PDF)

Tot i que al genoma humà conté aproximadament 20.000 gens, només una fracció reduïda (1,5% aproximadament) representen l'ADN total d'una cèl·lula (Watson *et al.*, 2016). La resta, és ADN no codificant basat en seqüències d'ADN repetides i amb una gran variabilitat entre els individus de la població (Alonso, 2004). Per tant, les diferències entre el genoma humà dels diferents individus són degudes a polimorfismes de diferent origen entre els quals es troben:

- SNPs: polimorfismes d'un sol nucleòtid (Figura 1). Son molt estables al llarg de les generacions (a diferència dels STRs) i presenten caràcter bial·lèlic. Aquests tipus de polimorfismes són utilitzats per a anàlisis biogeogràfics i per a determinar l'expressió del fenotip d'un individu (Freire-Aradas *et al.*, 2019).
- Minisatèlits: Tradicionalment s'havien emprat en el camp de la identificació aquests tipus de seqüències repetitives de major longitud (també anomenades VNTR), més localitzades en regions subtèlomèriques dels cromosomes amb motius repetitius de 15 a 50 parells de bases, generant seqüències d'entre 500 parells de bases fins a varies kilobases. Per al seu anàlisi, però, es necessiten grans quantitats d'ADN i per això s'ha limitat el seu ús en anàlisis forenses (Pifarré, 2019).
- Microsatèlits: són les anomenades repeticions curtes en tàndem o Short Tandem Repeats (STRs). Es caracteritzen per ser fragments d'ADN abundants i dispersos per tot el genoma (aproximadament es troben cada 7-8 kilobases), que formen part de l'ADN repetitiu, el qual constitueix un 30 % de l'ADN no codificant. Presenten unitats de repetició que oscil·len entre 1 i 8 parells de bases en un nombre comprès entre les 3 i 60 unitats i generant fragments d'entre poques desenes i centenars de parells de bases (Figura 1) (Pifarré, 2019).
- Anomalies estructurals aleatòries que afecten als nucleòtids de la seqüència, com insercions, delecions, inversions, duplicacions, etc.



**Figura 1.** Comparació entre un SNP i un STR en una seqüència d'ADN. Tot i que els SNPs representen canvis en una única base, els STRs difereixen en el nombre de repeticions en un locus determinat. Extret de Novroski i Cihlar, 2022.

Degut a totes aquestes característiques, es defineix el genoma humà com a universal, divers i estable, i és eficaçment aplicable en el camp de la identificació humana forense. Per una banda, totes les cèl·lules que componen els teixits i òrgans d'un individu son portadores del mateix conjunt d'instruccions genètiques, i per tant, independentment d'on s'extregui l'ADN, el perfil genètic serà el mateix. A partir d'aquest principi es basa l'anàlisi genètic comparatiu que es du a terme entre una mostra dubitada (la procedència de la qual és desconeguda) i una altra mostra de referència indubitada (la seva procedència i autenticitat es coneixen amb certesa) extreta de l'individu del que es sospita que hagi pogut deixar l'indici biològic dubitat. Per altra banda, degut a que l'ADN presenta variacions entre els diferents individus de la població, en permet la seva distinció de forma precisa. Finalment, i gràcies a que en condicions normals, l'ADN es caracteritza per tenir una elevada estabilitat tant en els fluids biològics com en els diferents teixits humans, es possible obtenir i genotipar la seqüència genètica a partir d'un gran nombre i diversitat de vestigis forenses (Alonso, 2004).

### 1.1.1. Estudi d'STRs en l'àmbit de genètica forense

Els STRs formen part de les seqüències repetitives en tàndem, fragments curts que tenen un ratio de mutació 100.000 vegades més elevat que en altres llocs genòmics, la qual cosa els aporta un alt grau de polimorfisme. A partir de la seva herència seguint patrons de tipus mendelià i la seva gran recombinació, els individus de la població difereixen tant en el número de repeticions com en l'ordre i tipus d'aquestes, causant també variacions en els al·lels autosòmics. Degut a aquest fet, i juntament amb la seva abundància i compatibilitat amb els anàlisis a través de PCR (reacció en cadena de la polimerasa), aquests polimorfismes han estat profundament investigats en estudis

forenses (Udogadi *et al.*, 2020). Aquesta aplicació va produir un establiment i expansió centralitzada de les bases de dades nacionals d'ADN criminal amb l'aparició per exemple de NDIS (National DNA Index System) o el CODIS (Combined DNA Index System), que proposen un conjunt bàsic de marcadors apropiats per poder discriminar de manera adequada els individus analitzats (Udogadi *et al.*, 2020).

La gran abundància d'STRs ha estat vital per a considerar la seva incorporació en kits comercials i ha estat la metodologia més utilitzada per a la identificació humana en els laboratoris, el diagnòstic molecular, estudis de filiació i paternitat així com estudis poblacionals o d'ascendència ancestral. Tot i que el seu procediment d'anàlisi depèn de mètodes establerts per cada casa comercial, la seva tipificació sempre conté uns passos generals en un ordre en concret que van des de l'extracció de l'ADN fins la generació d'electroferogrames i el seu anàlisi amb softwares específics (Figura 2).



**Figura 2.** Procediment per a l'anàlisi d'STRs. Extret de Udogadi *et al.*, 2020.

### 1.1.2. Estudi d'SNP en l'àmbit de genètica forense

Tot i que fa temps que s'utilitzen en genètica mèdica, l'ús d'SNPs per a la identificació d'individus en un context forense s'ha tornat cada vegada més popular en els darrers anys. De manera tradicional, la tipificació de l'ADN es realitza examinant repeticions curtes en tàndem a través d l'electroforesi capil·lar per separar els al·lels en funció de la seva mida (Budowle *et al.*, 1998; Clayton *et al.*, 1995; Edwards *et al.*, 1991). Tot i que s'ha demostrat que aquesta metodologia proporciona un alt poder de discriminació, encara hi ha limitacions relacionades amb el nombre de loci que s'analitzen en una sola reacció, així com problemes amb la correcta amplificació d'amplicons més grans quan s'analitza ADN degradat on existeixi la probabilitat de fragmentació del genoma.

Degut a que els SNPs presenten variabilitat d'un únic nucleòtid, son abundants en el genoma humà i tenen una mida reduïda dels seus amplicons, proporcionen una alternativa o un suplement per poder fer front a les limitacions anteriorment esmentades. A més, els SNPs tenen una taxa de mutació més baixa en comparació amb els STRs (de

l'ordre de  $10^{-8}$  per als SNPs (Nachman *et al.*, 2000), en comparació amb  $10^{-3}$  per als STRs (Brinkmann *et al.*, 1998)), cosa que pot ser especialment beneficiosa per a diverses aplicacions, com l'anàlisi de parentiu, la identificació humana, barreges genètiques i finalitats genealògiques (Novroski i Cihlar, 2022). Ara bé, com la majoria dels SNPs són bial·lèlics, es necessita analitzar un major nombre de loci per aconseguir el mateix poder de discriminació que amb els STR multial·lèlics. Tot i que es pot utilitzar la electroforesis capil·lar (EC) per a la seva identificació, aquest mètode requereix múltiples anàlisis per tipificar més d'entre 25 i 40 marcadors, la qual cosa requereix volums creixents de mostra i temps (Davenport *et al.*, 2023).

### 1.1.3. Ús del cromosoma Y en genètica forense

El cromosoma Y és un dels dos cromosomes sexuals humans, que conté dos regions pseudoautosòmiques (PARs), amb un braç curt (Yp) i un braç llarg (Yq). En aquest últim es on es troben la major part de les seqüències polimòrfiques repetitives (Y-STRs), que acostumen a no recombinar-se durant la meiosi amb el cromosoma X, per tant s'hereten conjuntament formant un haplotip que es transmet de pares a fills (Sydercombe, 2021b) i permet traçar la línia paterna d'un individu.

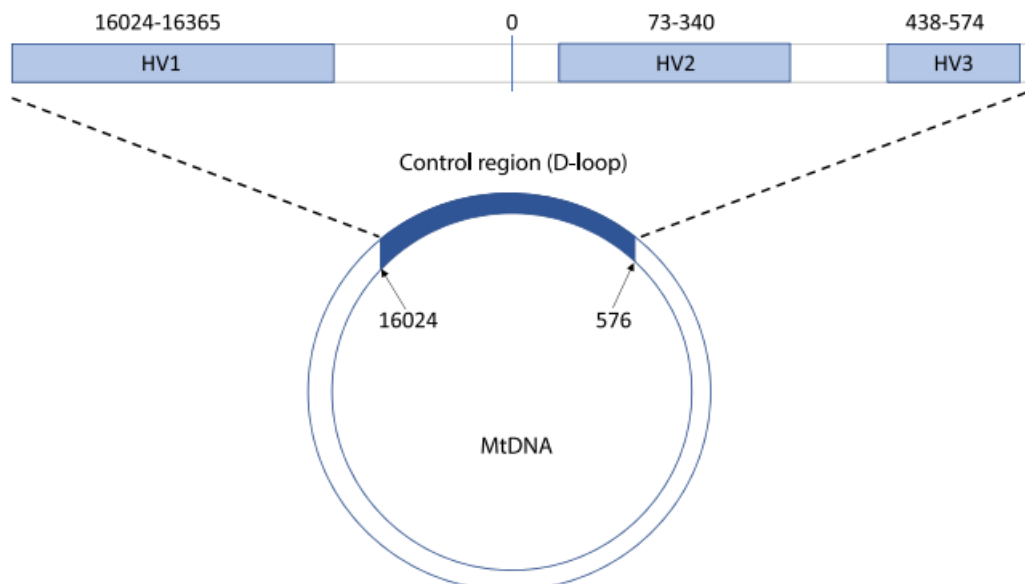
La publicació del primer STR polimòrfic descobert en la part no recombinant del cromosoma Y es va produir l'any 1992, i és a partir d'aquí quan es desenvolupem més marcadors per al seu ús en casos forenses (Gopinath *et al.*, 2016), fins l'actualitat, on ja hi ha aproximadament 27 marcadors inclosos en diferents kits comercials com l'Yfiler Plus de Thermo Fisher Scientific (Gopinath *et al.*, 2016).

L'aplicació dels Y-STRs en la investigació de la escena del crim s'utilitza principalment per excloure sospitosos masculins de la seva implicació, identificar la línia paterna de l'agressor i ressaltar múltiples contribuents. A més de ser utilitzat el seu anàlisi principalment en casos d'agressions sexuals, també és freqüent trobar la seva aplicació en proves de parentiu paterna, incloent casos històrics, així com en casos especials de persones desaparegudes i identificació de víctimes de desastres (Kayser, 2017). A més, els polimorfismes del cromosoma Y s'apliquen per inferir l'ascendència bio-geogràfica paterna de donants desconeguts o persones desaparegudes, en casos on el perfil a partir de l'DNA autosòmic no és informatiu (Kayser, 2017).

#### 1.1.4. ADN mitocondrial en genètica forense

L'ADN mitocondrial (ADNmt) està situat dins els orgànuls anomenats mitocòndries, delimitats per una membrana i que es troben en quantitats variables en el citoplasma de les cèl·lules humanes, on generen la major part del subministrament de trifosfat d'adenosina (ATP). El genoma mitocondrial va ser seqüenciat completament per primera vegada l'any 1981, i més endavant, va ser revisat per la Seqüència de Referència de Cambridge per a corregir-ne els errors, que actualment és el model per a tots els estudis (Syndercombe, 2021a).

Les mitocòndries contenen el seu propi genoma, organitzat en una estructura circular bicentenària (amb una cadena lleugera i una pesada) que conté entre 2 i 5 còpies completes de la seqüència genòmica (Syndercombe, 2021a). Aquesta consta de 37 gens que codifiquen tant per a rRNA i tRNA juntament amb 13 polipèptids. La resta de la molècula conté regions no codificants i una regió de control que s'encarrega de la transcripció. Dins d'aquesta, és on es troben les tres regions hipervariables, que contenen la major part de la variabilitat de la seqüència del genoma, i son utilitzades freqüentment en anàlisis forenses (Syndercombe, 2021a) (Figura 3). A diferència dels STRs, la seqüència d'ADNmt no és exclusiva d'un individu, sinó a un grup d'individus de la mateixa línia materna, transmesa de mares a fills (Syndercombe, 2021a).



**Figura 3.** Diagrama de l'estructura bàsica del genoma mitocondrial amb les 3 regions hipervariables. Extret de Syndercombe, 2021.

La seqüenciació de l'ADNmt és utilitzada avui dia en laboratoris forenses perquè aporta una eina essencial per a caracteritzar les evidències biològiques. Mentre que és difícil identificar a un individu degut a la seva limitada recombinació, ofereix avantatges per a confirmar llinatges materns o quan l'ADN nuclear està limitat com en casos d'anàlisis d'ossos, dents i cabell, degut al nombre elevat de còpies que presenta. A més, també és utilitzat normalment en casos d'anàlisis d'ADN ancestral i en identificacions de víctimes de grans desastres en massa (Syndercombe, 2021a).

## 1.2. Estat de la qüestió: Electroforesis capil·lar (EC) vs. Seqüenciació paral·lela massiva (MPS)

La genotipació de perfils d'ADN s'ha utilitzat en l'àmbit forense des de 1985, basant-se en els anàlisis de minisatèlits i la transferència Southern<sup>2</sup>. És més tard, cap als anys 90 (Jagger *et al.*, 2017), quan s'introdueixen les repeticions curtes en tàndem (STR) com un nou tipus de marcador d'ADN polimòrfic. A partir d'aquí, per a la identificació d'individus en casos penals, anàlisis de paternitat i identificació de persones desaparegudes, s'ha utilitzat el mètode d'electroforesis capil·lar, donant en alguns casos perfils clars a través d'anàlisis de STRs però en altres, resultats amb poca qualitat i utilitzant ADNmt (Senst *et al.*, 2022).

L'EC constitueix una tècnica de separació basada en la migració de molècules com l'ADN, proteïnes o ions inorgànics, que es sotmeten a un camp elèctric a través d'un capil·lar de menys de 50 µm de diàmetre (Magaña *et al.*, 2009). En el cas de l'ADN, els fragments objecte d'anàlisi s'uneixen a marcadors fluorescents i es mouen a través del gel d'agarosa per l'aplicació d'un camp elèctric cap al pol positiu. Aquests, es separen d'acord amb la longitud del fragment, de tal manera que els de menor pes molecular viatgen més ràpidament a través del capil·lar i són detectats per un làser d'argó, excitat a diferents longituds d'ona. Aquests factors caracteritzen la tècnica d'electroforesis com un mètode eficient i econòmic, capaç de separar diversos components de manera simultània amb mínimes quantitats de mostra i reactius (Magaña *et al.*, 2009).

Tot i el seu ús estandarditzat, la tècnica d'EC presenta limitacions relacionades amb la mida mínima necessària dels amplicons que s'han de seqüenciar, amb la separació

---

<sup>2</sup> Pàgina web del National Human Genome Research Institute. Disponible a: <https://www.genome.gov/es/genetics-glossary>

mínima dels loci per a una correcta resolució (Sharma *et al.*, 2020) i a la necessitat de tenir fragments d'ADN de mida similar per a poder ser etiquetats amb diferents marcadors fluorescents, normalment entre 25-30 loci (Senst *et al.*, 2022).

És per aquesta raó, que aquesta tècnica, si s'aplica a ADN degradat, presenta una limitació accentuada, ja que els loci més grans són menys probables de ser amplificats que els petits, generant així una falta de detecció de seqüències específiques que poden contenir informació significativa per millorar les dades de la tipificació al·lèlica individual (Sharma *et al.*, 2020). A més, aquest anàlisi no es podria aplicar en casos d'exhumacions on no se sap la identitat de l'individu ni es tenen mostres de referència de possibles familiars, especialment de primer grau, limitant-ne així la identificació (Calafell *et al.*, 2016).

El sorgiment de les tecnologies de seqüenciació de nova generació (NGS), basades en la MPS, s'han endinsat dins el camp de la investigació genòmica, incrementant en gran mesura la quantitat de seqüències produïdes en un *run* i reduint de forma dràstica el temps i cost de seqüenciació (Calafell *et al.*, 2016). A més, com s'ha demostrat per nombrosos estudis, la seqüenciació base per base, enforteix el poder de discriminació ja que es detecten variants en les regions flanquejants i possibiliten la multiplexació en un únic assaig de X-STRs i Y-STRs, SNPs informatius d'identitat (iiSNPs), SNPs informatius d'ascendència (aiSNPs) i SNPs fenotípics (piSNPs) (Senst *et al.*, 2022).

Una altre dels beneficis d'utilitzar múltiples amplicons, són les mesures més reduïdes d'aquests, que permeten augmentar els *ratios* genotípics en les mostres degradades o amb poca concentració d'ADN, ja que són menys susceptibles de ser sensibles a la degradació. A més, i juntament amb la possibilitat de resoldre barreges, la MPS genera informació tant de la llargària dels al·lèls com de la seqüència, permetent discriminar al·lèls isomèrics produïts per la presència de SNPs o repeticions incompletes o complexes dels STRs (Calafell *et al.*, 2016).

### 1.3. Programa de Fosses de la Comunitat Valenciana

Dins la Comunitat Valenciana, i a partir de l'entrada en vigor de la Llei 14/2017, de 10 de novembre, de la Generalitat, de memòria democràtica i per a la convivència de la Comunitat Valenciana, s'han estat desenvolupant una sèrie de polítiques de caràcter públic que garanteixen els drets de les víctimes de la Guerra Civil Espanyola i el

Franquisme, així com dels seus familiars. En aquest marc, s'ha elaborat un programa de Fosses de la Comunitat Valenciana, que inclou, entre altres coses, l'elaboració d'un mapa de zones on es localitzen víctimes de la dictadura franquista, la confecció d'un cens d'aquestes, així com l'exhumació de les diferents fosses i la identificació dels individus que s'hi troben sepultats<sup>3</sup>.

En relació amb l'últim punt, l'article 12 de la Llei 14/2017, estableix l'obligació de realitzar proves bioantropològiques i genètiques que permeten la identificació de les restes òssies exhumades i la de gestionar un banc de dades ADN, com a centre d'investigació i depòsit temporal de totes les restes humanes exhumades fins al lliurament definitiu per a la seva reinhumació. Per al desenvolupament d'aquestes competències és necessari utilitzar una metodologia interdisciplinària, genètica, forense i històrica, i és per això, que el 12 de maig de 2021 es va firmar un acord entre la Conselleria, la Fundació per al Foment de la Investigació Sanitària i Biomèdica de la Comunitat Valenciana (fundació del sector públic, adscrita a la Conselleria de Sanitat), i la Universitat de València-Estudi General<sup>4</sup>.

Gràcies a aquests acords, s'han desenvolupat fins a l'actualitat, una sèrie de projectes d'investigació amb la finalitat de millorar les taxes d'identificació, posant en marxa noves tècniques, així com la seva transferència i difusió científica i social. És el Servei de Seqüenciació i Bioinformàtica de Fisabio, qui s'encarrega de la identificació històrica d'individus, utilitzant els serveis de genètica forense basats en la seqüenciació massiva de segona generació. Es disposa d'un seqüenciador MiSeq FGx (Illumina) desenvolupat i validat per genòmica forense que permet obtenir perfils genètics que inclouen més de 200 marcadors d'SNPs autosòmics, STRs autosòmics i sexuals (Y-STR y X-STR) i perfils genètics de l'ADN mitocondrial<sup>5</sup>.

---

<sup>3/4</sup> Pàgina Web de la Generalitat Valenciana de Memòria Democràtica. Conselleria de Participació, Transparència, Cooperació i Qualitat Democràtica. Disponible a: <https://participacio.gva.es/va/web/memoria/qu%C3%A8-hi-fem->

<sup>5</sup> Pàgina Web del Servei de Seqüenciació i Bioinformàtica de FISABIO. Salut Pública de la Generalitat Valenciana. Disponible a: [https://www.fisabio-ngs.com/genetica\\_forense/](https://www.fisabio-ngs.com/genetica_forense/)

#### 1.4. MiSeq FGx™ Forensic Genomics System

El MiSeq FGx™ Forensic Genomics System (Illumina) ha estat desenvolupat específicament per a la identificació d'individus i per a línies d'investigació, permetent als laboratoris forenses augmentar el número de marcadors analitzats en comparació amb els mètodes d'electroforesis capil·lar, sent necessàries quantitats d'ADN, només d'1ng (Calafell *et al.*, 2016). El sistema consisteix en 4 components: el ForenSeq™ DNA Signature Prep Kit, el MiSeq FGx™ Reagent Kit, l'instrument de seqüenciació MiSeq FGx™, i el software ForenSeq™ Universal Analysis (Jäger *et al.*, 2017).

El ForenSeq™ DNA Signature Prep Kit es basa en una PCR d'amplificació i en la creació de llibreries, amb la majoria d'amplicons resultants amb una mesura de menys de 200 bp (Senst *et al.*, 2022). L'assaig conté dos mix de primers segons el tipus d'anàlisi que es realitzi (Calafell *et al.*, 2016; Jäger *et al.*, 2017; Zhang *et al.*, 2018):

- DNA Primer Mix A (DPMA) que marca l'Amelogenina, 27 STRs autosòmics, 24 Y-STRs, 7 X-STRs, 94 SNPs informatius d'identitat (iiSNPs)
- DNA Primer Mix B (DPMB) que a més de marcar tots els components del DPMA, també s'afegeix l'anàlisi de 56 SNPs informatius de l'ascendència biogeogràfica (aiSNPs) i 22 SNPs fenotípics més que són informatius del color de cabell i d'ulls (piSNPs).

Per tant, el MiSeq FGx™ Forensic Genomics System facilita la seqüenciació simultània incrementant el poder de discriminació en relació a les mostres difícils. En els últims anys, s'han realitzat diferents assajos per provar la seva eficiència, demostrant així que la tècnica i el sistema compleixen amb les garanties de qualitat de l'ADN forense, proporcionant un rendiment fiable i reproducible (Zhang *et al.*, 2018).

## **2. OBJECTIUS**

Tot i els avantatges i els avanços generats gràcies a la seqüenciació massivament paral·lela (MPS) en el marc de la investigació forense, encara falta un llarg camí per a que s'implementi en tots els laboratoris. Tanmateix l'anàlisi de mostres degradades com els ossos, i en particular aquells que porten llarg temps soterrats com és el cas de les víctimes de fosses comunes de la Guerra Civil i del Franquisme continua sent un dels principals reptes dels laboratoris de genètica forense. És per això que es necessita

temps, personal qualificat i validacions internes per a poder desplaçar la tecnologia dominant de la electroforesis capil·lar (Senst *et al.*, 2022).

La tècnica de seqüenciació massiva aplicada a aquestes mostres ha demostrat millorar les taxes d'identificació al augmentar el nombre de marcadors analitzats alhora i a reduir la llargària dels fragments d'estudi. Tanmateix al ser una tècnica de recent introducció i tal i com indica el Scientific Working Group on DNA Analysis Methods (SWGDM), és necessari establir uns llindars interpretatius i d'anàlisi tant interns com compartits. L'objectiu principal, per tant, serà establir aquests llindars a nivell intern per al projecte de millora de les taxes d'identificació de les víctimes de la Guerra Civil i del Franquisme que porta a terme FISABIO per encàrrec de la Generalitat Valenciana.

Tot i que s'han realitzat un seguit de validacions de la tècnica per MPS, amb mostres humanes, per tal d'avaluar la sensitivitat, repetibilitat, reproductibilitat, etc., i així optimitzar el flux de treball (Senst *et al.*, 2022), la majoria d'aquests no reflecteixen el que es troba realment a les escenes forenses amb poca qualitat i quantitat d'ADN. És per això, que aquest treball té com a objectiu fer un disseny experimental per tal de **reproduir allò que es troba en contextos arqueològics**, especialment ADN de teixits *postmortem* o cadàvers alterats que poden estar altament degradats i inhibits degut a processos de descomposició.

A partir d'aquesta premissa, es realitzaran dos experiments amb diferents condicions als que seran exposades mostres de referència, per tal **d'avaluar els llindars interpretatius i de detecció d'anàlisi de STRs i SNPs obtinguts mitjançant el kit ForenSeq™ DNA Signature Prep kit (Verogen) per seqüenciació Illumina MiSeq™ FGx**. El disseny experimental serà el següent:

- Experiment 1: 2 mostres de referència amb una concentració inicial de 10 ng/ul han estat gradualment degradades a través del termociclador Mastercycler X50s (Eppendorf) a 98°C durant diferents intervals de temps. A més, en aquest experiment també s'ha combinat la fragmentació amb un seguit de dilucions per afegir-hi una nova variable semblant al que experimenten les mostres òssies arqueològiques. El principal objectiu d'aquest assaig és **avaluar el llindar interpretatiu de mostres fragmentades** i el límit en el qual els STRs comencen a estar afectats per drop-outs segons el grau d'afectació (baix, moderat o intens), així com el nivell de detecció al·lèlica segons el temps d'exposició al calor.

- Experiment 2: 2 mostres de referència han estat exposades a diferents dilucions per tal **d'establir el límit interpretatiu en relació amb la concentració d'ADN que es pot trobar en un context forense**, el qual ha estat parcialment afectat per factors externs. Se sap que una quantitat insuficient d'ADN provoca un augment en la formació de dímers adaptadors, la qual cosa pot afectar negativament a la qualitat de la seqüenciació (Senst *et al.*, 2022).

### **3. MATERIALS I MÈTODES**

#### 3.1. Presa de mostres

Per la presa de mostres s'ha fet ús dels serveis del Biobanc per a la Investigació Biomèdica i en Salut Pública de la Comunitat Valenciana (IBSP-CV). S'han agafat 8 mostres en total de 2 voluntaris: 4 mostres d'un individu femení (EXT – 001) i 4 mostres d'un individu masculí (EXT – 004). S'ha fet servir un kit de presa de mostra forense de l'empresa Sarstedt, Inc., amb un hisop de viscosa, que s'ha guardat dins un tub amb membrana de ventilació. Cada mostra s'ha etiquetat correctament amb les sigles: EXT – 001 (1, 2, 3, 4) i EXT – 004 (1, 2, 3, 4) i s'emmagatzemen a temperatura ambient fins la posterior extracció.

#### 3.2. Extracció d'ADN

El següent pas ha estat l'extracció de l'ADN amb el protocol d'extracció NucleoSpin® DNA Forensic de la casa Macherey-Nagel, Inc., que es basa en l'adsorció reversible dels àcids nucleics a la membrana de sílice de les columnes NucleoSpin® DNA Forensic. Seguidament es fan una sèrie de rentats per eliminar contaminants i sals mitjançant Wash Buffers (FOW 1 i FOW 2) i finalment, s'elueix amb Elution Buffer, permetent obtenir la mostra d'ADN purificat. Tenim, per tant, 8 alíquotes que contenen l'ADN, més un control negatiu etiquetat com a CN-089.

#### 3.3. Quantificació

Per a poder mesurar la concentració inicial a la qual tenim les mostres s'han quantificat mitjançant el fluorímetre Qubit 4 amb el kit Qubit™ 1X dsDNA High – Sensitivity (HS)

Assay d'Invitrogen. Aquest aparell indica les dades de quantificació per a mostres que es troben dins el rang de la corba estàndard, realitzant una estimació de la concentració de la mostra (Taula 1).

Mostra	[DNA] Qubit (ng/μl)
EXT-001_1	TH <sup>1</sup>
EXT-001_2	TH
EXT-001_3	TH
EXT-001_4	45,5
EXT-004_1	TH
EXT-004_2	TH
EXT-004_3	TH
EXT-004_4	TH
CN-089	TL <sup>2</sup>

**Taula 1.** Valors de concentració d'ADN de les mostres de l'estudi quantificades amb el fluorímetre Qubit

<sup>1</sup>TH: Too – high: Significa que la concentració d'ADN és > 60 ng/uL

<sup>2</sup>TL: Too – low: La concentració d'ADN és ≈ 0 ng/uL. Acostuma a ser el CN o mostres que presenten un alt nivell de degradació.

A més, per saber l'estat de puresa en el qual es troben les mostres, es fa una segona quantificació amb l'espectrofotòmetre Thermo Scientific™ NanoDrop™ One, que utilitza l'absorció UV per mesurar a diferents longituds d'ona les concentracions d'ADN, ARN i la quantitat de contaminants (Taula 2).

Mostra	[DNA]NanoDrop (ng/μl)	260/280 <sup>1</sup>	260/230 <sup>2</sup>
EXT-001_1	161,5	1,91	0,93
EXT-001_2	162,0	1,92	2,06
EXT-001_3	92,8	1,89	1,94
EXT-001_4	80,8	1,87	1,89
EXT-004_1	198,8	1,89	1,96
EXT-004_2	174,1	1,93	1,93
EXT-004_3	164,0	1,90	1,99
EXT-004_4	128,5	1,90	1,88
CN-089	1,1	1,21	0,1

**Taula 2.** Valors de concentració quantificats utilitzant l'espectrofotòmetre Thermo Scientific™ NanoDrop™ One

<sup>1</sup>260/280. *Ratio* entre la concentració d'ADN (absorbeix la llum a 260 nm) i la concentració d'ARN (absorbeix a 280 nm). La relació ens indica la puresa de la mostra. Els resultats ideals haurien de sortir al voltant de 1.8

<sup>2</sup>260/230. *Ratio* entre concentració d'ADN i contaminants (absorbeixen la llum a un rang de 230 nm). Els contaminants poden ser proteïnes, carbohidrats, components del kit o de l'extracció. Els resultats òptims haurien de sortir al voltant de 2 o 2.2.

### 3.4. Experiment 1: Estudi de degradació + dilucions

Per a decidir la concentració inicial de les mostres per a l'estudi de degradació, s'han analitzat a través del Fragment Analyzer System d'Aligent, un instrument d'electroforesis capil·lar que permet determinar el nivell de fragmentació que tenien. Aquestes, anteriorment s'havien quantificat amb el fluorímetre Qubit 4 donant en tots els casos, excepte un, una concentració de TH (> 60 ng/μl). És per aquesta raó, que les 4 alíquotes de cada mostra s'ajunten en una sola i es dilueixen a 1:20 (4 μl de mostra + 38 μl d'H<sub>2</sub>O), per a que la seva concentració entri dins els nivells de detecció del Fragment Analyzer. Els resultats han estat comparats amb l'ADN de control del ForenSeq DNA Signature Prep Kit (2800 M), i ambdós mostres (EXT-001 i 004) presentaven un cert nivell de fragmentació, segurament degut al procés d'extracció a través de la columna.

Per aquest motiu, i a diferència d'altres estudis de validació del ForenSeq Prep Kit (Fattorini *et al.*, 2017; Zhang *et al.*, 2018; Sharma *et al.*, 2020; Senst *et al.*, 2022), es decideix utilitzar una concentració inicial de 10 ng/μl. Una vegada fetes les dilucions corresponents, quantifiquem per saber de quina concentració d'ADN partim (Taula 3).

Mostra	[DNA] Qubit (ng/μl)	Mitjana [DNA] (ng/μl)
EXT-001_1	9.67	9.43
EXT-001_2	9.18	
EXT-004_1	10	10
EXT-004_2	10	

**Taula 3.** Valors de concentració d'ADN de les mostres de l'estudi quantificades amb el fluorímetre Qubit a 10 ng/ul.

Una vegada tenim preparades les mostres, es preparen 12 alíquotes (6 de cada individu de 25 μl cada una, i es dissenya un programa de temperatures amb el termociclador Mastercycler X50s (eppendorf) a 98 °C amb 8 punts de temps diferents: **0, 2, 10, 20, 40, 60, 80, 120** (Zhang *et al.*, 2018; Sharma *et al.*, 2020). Aquests intervals de temps han estat seleccionats després de realitzar proves amb el Fragment Analyzer, per veure el nivell de fragmentació així com quantificacions amb el Nanodrop, per tal de tenir una

davallada de la concentració el més progressiva possible. A més, decidim també que en el punt de 10' i en el punt de 40', s'afegeix una nova variable de dilucions, a 0,1 ng/ul i a 0,02 ng/ul, així com en el punt de 60' de la mostra EXT-004, al que se li afegeix una dilució de 0,1 ng/ul.

A partir de totes aquestes dilucions preparades, es procedeix a la preparació de la llibreria. Primer, es realitza la PCR1. Després es procedeix a l'enriquiment dels amplicons a través dels adaptadors d'índex 1, i7 (R701-R708) i d'índex 2, i5 (A505-A508). Amb aquest procediment s'etiqueta l'ADN amb una combinació única de seqüències d'índex, la qual cosa permet separar les seqüències de cada mostra durant el posterior anàlisi. Al finalitzar la PCR2, el procés segueix amb la purificació de la llibreria utilitzant les SBP, boles magnètiques que separen l'ADN amplificat d'altres components de la reacció. Seguidament, passem a la normalització de les llibreries per a que estiguin igualment representades en el *run* de seqüenciació i aconseguir una densitat de clústers consistent per a optimitzar la resolució de les mostres quan s'agrupin en un *pool*.

Un cop acabat el procés de llavats amb el suport magnètic i els reactius necessaris, es passa al procés d'agrupació de llibreries per a crear un *pool* (de 5ul de cada llibreria) que es seqüenciaran juntes en la mateixa cel·la de flux (amb un input de 8,5 ul).

### 3.5. Experiment 2: Estudi amb dilucions

Per a l'experimentació amb diferents concentracions d'ADN es decideix partir de 7 punts de dilució: **0.5 ng/ul, 0.2 ng/ul, 0.1 ng/ul, 0.02 ng/ul, 0.01 ng/ul, 0.002 ng/ul i 0.0002 ng/ul**. Aquests intervals es decideixen a través de diferents anàlisis de concentració (Tecan Infinite F Nano +) i el Fragment Analyzer, per a tenir resultats progressius i no canvis bruscos.

Partint de les concentracions inicials de les mostres originals (EXT-001 = 75.4 ng/ul i EXT-004 = 117.8 ng/ul), les diluïm a 5 ng/ul en un volum total de 50 ul. A partir dels càlculs, quantifiquem per triplicat cada mostra per saber de quina concentració partim per fer les posteriors dilucions (Taula 4)

Mostra	[DNA] Qubit (ng/μl)	Mitjana [DNA] (ng/μl)
EXT-001_1	4.86	4.88
EXT-001_2	4.84	
EXT-001_3	4.93	
EXT-004_1	4.68	4.69
EXT-004_2	4.71	
EXT-004_3	4.68	

**Taula 4.** Valors de concentració d'ADN de les 2 mostres de l'estudi quantificades amb el fluorímetre Qubit a 5 ng/ul.

D'aquí fem la dilució a 0.5 ng/ul amb un volum final de 40 ul i la dilució 0.2 ng/ul amb un volum final de 80 ul i la quantifiquem (Taula 5 i 6)

Mostra	[DNA] Qubit (ng/μl)
EXT-001	0.474
EXT-004	0.466

**Taula 5.** Concentració d'ADN de les 2 mostres quantificades amb el fluorímetre Qubit a 0.5 ng/ul

Mostra	[DNA] Qubit (ng/μl)	Mitjana [DNA] (ng/μl)
EXT-001_1	0.196	0.191
EXT-001_2	0.184	
EXT-001_3	0.194	
EXT-004_1	0.192	0.187
EXT-004_2	0.184	
EXT-004_3	0.186	

**Taula 6.** Concentració d'ADN de les 2 mostres quantificades amb el fluorímetre Qubit a 0.2 ng/ul.

A partir de la dilució de 0.2 ng/ul, farem les dilucions seriades següents (0.002 ng/ul i 0.0002 ng/ul) amb un volum final de 80 ul.

A partir d'aquestes mostres, i com en l'experimentació anterior, es realitza la PCR1 i la PCR2, la purificació, normalització, pool, i finalment es carrega el seqüenciador amb totes les mostres (a excepció de la EXT-001, 2.5 ng i 0.5 ng), amb un input de 7 ul.

### 3.6. Seqüenciació i anàlisi de dades.

Els *runs* de seqüenciació han estat realitzats amb el sistema Miseq™ FGx d'Illumina, Abans de poder carregar l'input de mostra, es necessita fer un rentat post-run i un rentat de manteniment per a eliminar possibles restes de *runs* anteriors. Una vegada preparada la mostra amb el Buffer corresponent, s'injecta en el cartutx de carrega perforant-lo i es col·loca en el seqüenciador.

Una vegada ha acabat el *run*, es genera un document amb les mètriques de qualitat que es converteixen a un full de dades de Microsoft Office Excel, on consten elements com la densitat de clúster, els clústers que han passat el filtre, el percentatge de “phasing” i “pre-phasing” així com el número total de lectures del Forenseq Universal Analysis Software. A més, també es decideix fer un anàlisi amb les mètriques de control de qualitat (QC), per ajustar el número de lectures i el percentatge d'aquestes que passen el filtre del programa.

Una vegada tenim generat els resultats del *run* en el Software del Forenseq, es crea un nou anàlisi dins el programa per tal de modificar els paràmetres que venen per defecte segons els protocols d'interpretació del propi laboratori. Seguint aquestes instruccions s'omple un altre full de dades de Microsoft Office Excel, per anar fent un seguiment d'aquelles dades que es modifiquen de cada rèplica, per després, aplicar les normes dels consensos i arribar a tenir un perfil genètic complet.

## **4. RESULTATS I INTEPRETACIÓ**

### 4.1. Experiment 1

El *run* de l'experiment 1 passa totes les mètriques de qualitat (*clúster density*, *clúster passing filter*, *phasing* i *prephasing*), així com l'índex i les lectures de qualitat (5164620 lectures en total). El control positiu de l'experiment no ha tingut cap drop-out, és a dir, s'han pogut seqüenciar tots els STRs i els SNPs. Del control negatiu, en canvi, per contaminació durant el procés d'extracció o del propi seqüenciador, s'han seqüenciat dos STRs. Amb un nombre total de 137 lectures, s'ha detectat l'al·lel homozigot 18, 18 a D2S1338, i l'al·lel homozigot 25, 25 a FGA. Tot i això, l'experiment passa els filtres de qualitat perquè no sobrepassa els límits de detecció de 3 loci.

Les mostres EXT-001 i EXT-004 d'ADN de bona qualitat han estat degradades a través del termociclador a 98°C a diferents intervals de temps i s'han analitzat a través del Fragment Analyzer per a observar el diferents nivell de fragmentació resultant. A partir d'aquí, i realitzant un *ratio* dels fragments més curts (de 0 a 100 parells de bases) vers al fragments més llargs (de 100 al límit màxim de detecció de parells de bases), s'ha calculat el percentatge de degradació de les mostres (Taula 7). Quan més exposat l'ADN al calor, major la degradació i major el nombre de drop-outs (ja que també la concentració disminueix).

Com es pot observar a la Taula 7, a partir dels 2 min ja tenim una degradació del 22,58 %, que va augmentant de manera progressiva a mesura que s'augmenta el temps d'exposició. És entre el minut 20 i el minut 40 quan veiem un salt més abrupte, passant d'una mostra degradada al 41% a una mostra amb una degradació del 70%, en el cas de la mostra EXT-001; i d'un 30% a un 65% en el cas de la mostra EXT-004, factor que es veurà també recolzat amb la pèrdua de més al·lels i locus en l'anàlisi posterior. En aquest cas, la degradació màxima a la qual s'arriba també varia entre mostres, d'aproximadament un 86%, entre els 80 i 120 minuts, en la mostra EXT-001; i entre un 80 i 94% en l'EXT-004. Aquesta discordança de degradació entre els temps no és d'estranyar, ja que els fragments son tan petits que son difícils de detectar i consegüentment hi ha un percentatge d'error en el càlcul.

Mostres	0-100	100-UM	Suma 100-UM	%Degradació		Correcció %
EXT-001 0 (1)	4,56	87,74	92,3	4,94	5,42	<b>0,00</b>
EXT-001 0 (2)	7,55	120,43	127,98	5,90		<b>0,00</b>
EXT-001_0.5 2'	10,1	25,97	36,07	28,00		<b>22,58</b>
EXT-001_0.5 10'	13,91	23,13	37,04	37,55		<b>32,13</b>
EXT-001_0.5 20'	23,51	26,14	49,65	47,35		<b>41,93</b>
EXT-001_0.5 40'	24,52	7,68	32,2	76,15		<b>70,73</b>
EXT-001_0.5 60'	41,33	11,31	52,64	78,51		<b>73,09</b>
EXT-001_0.5 80'	46,65	3,48	50,13	93,06		<b>87,64</b>
EXT-001_0.5 120'	29,11	3,01	32,12	90,63		<b>85,21</b>

EXT-004 0 (1)	3,08	103,7	106,78	2,88	3,76	<b>0,00</b>
EXT-004 0 (2)	7,11	145,94	153,05	4,65		<b>0,00</b>
EXT-004_0.5 2'	12,31	33,38	45,69	26,94		<b>23,18</b>
EXT-004_0.5 10'	18,54	28,73	47,27	39,22		<b>35,46</b>
EXT-004_0.5 20'	17,7	28,31	46,01	38,47		<b>34,70</b>
EXT-004_0.5 40'	33,91	16,1	50,01	67,81		<b>64,04</b>
EXT-004_0.5 60'	6,43	2,13	8,56	75,12		<b>71,35</b>
EXT-004_0.5 80'	36,06	0,77	36,83	97,91		<b>94,14</b>
EXT-004_0.5 120'	21,04	4,22	25,26	83,29		<b>79,53</b>

**Taula 7.** Càlcul del percentatge de degradació de les mostres a partir de la *ratio* entre els fragments de 0 a 100 bases i de 100 a UM (lindar màxim). Dades extretes del Fragment Analyzer.

Els drop-outs no es detecten a 2 min, però si que comencen a aparèixer al minut 10, tant al·lèlics (Penta E) com de locus (DYS19 i DYS522) a la mostra EXT-004, mentre que per a la mostra EXT-001 ens hem d'esperar fins a la dilució 1:5 del mateix temps de fragmentació per veure'n un al·lèlic al Penta E (Taula 8). En el cas de drop-outs referents a iSNPs, no es detecten fins al minut 10 amb una dilució 1:5, en el cas de l'EXT-004 (a rs12997453 rs1355366 rs10488710), però no es fins als 40 minuts sense dilució que en detectem tant a l'EXT-001 com a l'EXT-004, podent interpretar així, que la dilució i la disminució de quantitat d'ADN és un factor important per a l'aparició d'aquests drop-outs (a menys ADN, menys lectures i menys detecció). A l'analitzar la taula resum (Taula 8), s'observa que a mesura que augmenta el temps de fragmentació amb el termociclador, sorgeixen més drop-outs a causa de la degradació de l'ADN, arribant a la seva totalitat en els STR-s autosòmics, sexuals i iSNPs a intervals de temps de 120 min.

Els iSNPs han estat més sensibles a la degradació tot i les seves mesures d'amplicons d'entre 63 fins a 231 bp (Sharma *et al.*, 2020). Tot i que la cobertura mitjana d'aquests ha estat aproximadament de 6 vegades menor comparat amb els STRs, encara tenen un ús limitat en casos de genètica forense (Sharma *et al.*, 2020).

Mostra	27 aSTRs i AMEL		24 Y-STRs		7 X-STRs		94 iSNPs	
	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004
0 min	0 <sup>1</sup>	0	0	0	0	0	0	0
2 min	0	0	0	0	0	0	0	0
10 min	1 INC <sup>4</sup>	1 ADO <sup>2</sup>	0	2 LDO <sup>5</sup>	0	1INC	0	0
0.5_10 min	1 ADO	2 INC 2 ADO	0	4 INC 2 LDO	0	1INC 1LDO	2 INC	2 LDO 4 INC 1 ADO
0.1_10 min	2 LDO 2 INC 1 ADO	8 INC 3 LDO	0	7 LDO 9 INC 1 ADO	1 LDO 1 INC 1 ADO	2 INC 3 LDO	9 LDO 1 ADO 26 INC	26 INC 17 LDO 2 ADO
20 min	1 ADO	1 INC	0	2 LDO 3 INC	1 LDO	1 LDO	0	0
40 min	3 ADO 2 INC 4 LDO	1 ADO 1 LDO	0	9 LDO	3 LDO 2 INC	2 LDO	19 INC 1 ADO 10 LDO	9 INC 4 LDO 1 ADO
0.5_40 min	11 LDO 2 INC 4 ADO <b>1 ADI<sup>3</sup></b>	3 ADO 2 INC 3 LDO <b>1 ADI</b>	0	2 INC 11 LDO <b>1 ADI</b>	4 LDO 1 INC	2 LDO	38 LDO 30 INC 3 ADO	31 INC 11 LDO 2 ADO
0.1_40 min	20 LDO 5 ADO	10 ADO 10 LDO 2 INC <b>1 ADI</b>	0	18 LDO	7 LDO	6 LDO	70 LDO 20 INC	35 INC 44 LDO
60 min	1 ADO 3 LDO 1 INC	15 LDO 2 INC <b>1 ADI</b>	0	17 LDO 2 INC 2 ADO	1 LDO	6 LDO	2 LDO 9 INC	28 INC 47 LDO
0.5_60 min	/	3 ADO 21 LDO 3 INC	/	22 LDO 2 INC	/	7 LDO	/	17 INC 72 LDO
80 min	3 ADO 9 LDO 4 INC	6 ADO 18 LDO 4 INC	0	1 ADO 1 INC 22 LDO	<b>1 ADI</b> 2 LDO 1 ADO	7 LDO	21 INC 30 LDO 2 ADO	11 INC 78 LDO
120 min	28 LDO	27 LDO 1 INC	0	24 LDO	7 LDO	7 LDO	8 INC 84 LDO	2 INC 92 LDO

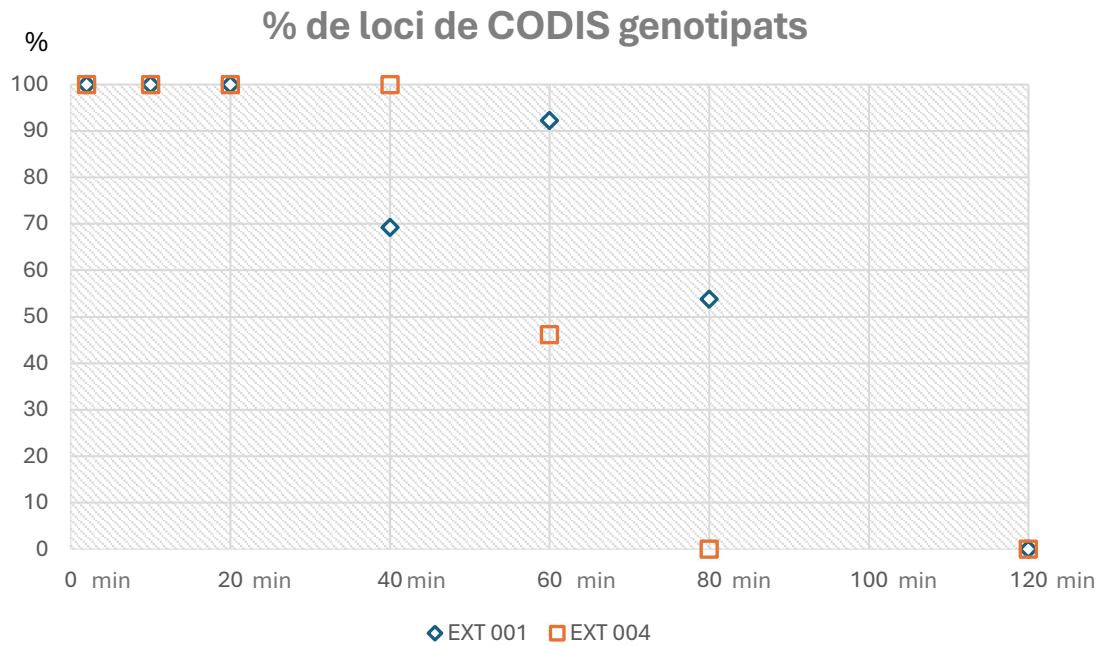
**Taula 8.** Anàlisi d'efectes estocàstics de l'experiment 1 després de que les mostres hagin estat exposades a calor durant diferents intervals de temps. <sup>1</sup>0: No s'han detectat drop-outs / <sup>2</sup>ADO: drop-out al·lèlic / <sup>3</sup>ADI: drop-in al·lèlic / <sup>4</sup>INC: Inconclusiu. També per als Y-STRs de la mostra EXT-001. / <sup>5</sup>LDO: Drop-out del Locus (quan no es detecta senyal).

En la totalitat de l'experiment es detecten 6 drop-ins al·lèlics. Per una banda, la mostra EXT-001, en presenta un als 40 min de degradació, amb una dilució 1:5. En concret, apareix l'al·lel 11, 11 a D22S104. A més, torna a detectar-se un altre drop-in a 80 min, amb l'al·lel 14 a DXS7132. En el cas de la mostra EXT-004, n'apareixen 4 més, en locus diferents. Per una banda, a 40 min de degradació i amb una dilució 1:5, es troba l'al·lel 13 a D9S1122, i al mateix temps un al·lel 18 a DYS576. D'altra banda al minut 40 amb una dilució 1:25, es detecta un altre drop-in a D20S482, amb l'al·lel 13; així com al minut 60, amb l'al·lel 10 a D17S130. En tots els casos ha estat possible distingir quin era el vertader al·lel i quin l'artefacte degut a les mostres de referència. A mesura, per tant que augmenta el temps de degradació i es fragmenta més l'ADN, així com el factor de dilució, que disminueix de manera significativa la quantitat de mostra genòmica, apareixen més artefactes per errors de lectura en el seqüenciador.

S'han comparat també l'experiment amb els 13 locus del CODIS<sup>6</sup> més importants que s'han detectat en les diferents mostres exposades a diferents temps (Figura 4). Aquelles mostres menys degradades, les de 2, 10 i 20 min, es detecten en la seva totalitat els 13 loci STRs del CODIS. A 40 min, observem una davallada en la mostra EXT-001, de 4 loci. En aquest cas, però, 2 d'ells son per pèrdua al·lèlica (D3S1358 i D16S539), i el vWA es per un inconclusiu. L'únic que mostra un drop-out total es el D21S11. Mentre que en la mostra EXT-004, s'han pogut detectar tots. En canvi, a 60 min, en la mostra EXT-001 només es troba un drop-out inconclusiu a D21S11, mentre que en la EXT-004, ja en trobem fins a 7 (Taula 9). És per tant, a partir d'aquest interval de temps, on trobem una degradació mitjana, que els marcadors comencen a fallar per falta de lectures i fragmentació, i comencen a aparèixer drop-outs que dificultarien l'anàlisi forense. Com s'observa a la Taula 9, a 80 min en l'EXT-001 encara es poden recuperar alguns marcadors, mentre que en l'EXT-004 ja no; i a 120 min, en ambdues mostres ja perdem la totalitat de loci ja sigui parcialment o totalment.

---

<sup>6</sup> Proyecto Biológico, Universitat d'Arizona (2000). Trad. de la Universitat d'Alcalá de Henares (2004). Disponible a: [https://biomodel.uah.es/epb/human\\_bio/activities/blackett2/str\\_codis.html](https://biomodel.uah.es/epb/human_bio/activities/blackett2/str_codis.html)



**Figura 4.** Percentatge de loci genotipats relatius al CODIS de l'experiment 1.

	2'		10'		20'		40'		60'		80'		120'	
	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004
D3S1358							ADO			LDO	LDO	LDO	LDO	LDO
vWA							INC			INC	INC	LDO	LDO	LDO
FGA											INC	INC	LDO	LDO
D8S1179												ADO	LDO	LDO
D21S11							LDO		INC	LDO	INC	LDO	LDO	LDO
D18S51												INC	LDO	LDO
D5S818										LDO		ADO	LDO	LDO
D13S317										INC		LDO	LDO	LDO
D7S820										LDO	LDO	LDO	LDO	LDO
D16S539							ADO			LDO	LDO	LDO	LDO	LDO
THO1												ADO	LDO	INC
TPOX												INC	LDO	LDO
CSF1PO												ADO	LDO	LDO
AMEL										LDO		LDO	LDO	LDO

**Taula 9.** Efectes estocàstics de l'experiment 1 produïts en els marcadors específics del CODIS.

## 4.2. Experiment 2

El *run* de l'experiment 2 passa totes les mètriques de qualitat (*clúster density*, *clúster passing filter*, *phasing i prephasing*), així com l'índex i les lectures de qualitat (4573479 lectures en total). El control positiu de l'experiment no ha tingut cap drop-out, és a dir, s'han pogut seqüenciar tots els STRs i els SNPs; i en el cas del control negatiu, no s'ha detectat cap senyal, indicant que no hi ha cap tipus de contaminació.

Com s'observa a la Taula 10, els drop-outs dels STRs es detecten en la mostra EXT-004 per primera vegada, a partir de la dilució 1:25, és a dir, amb un input de material genètic de 0.1 ng. En aquest cas, s'observen dos de tipus al·lèlic a D5S818, on es perd l'al·lel 11, i a Penta E, on es perd l'al·lel 12; mentre que la resta de desaparicions son degudes a resultats inconclusius després de realitzar els consensos. Quan passem a un input de 0.05 ng, la mostra EXT-001 segueix sense tenir problemes amb drop-outs, mentre que a la mostra EXT-004 es tornen a detectar dos al·lèlics. En aquest cas, al locus D1S1656 amb la pèrdua de l'al·lel 17.3, i a D7S820, on es perd l'al·lel 12.

És a partir de l'input de 0.01 ng on ja es veu un augment considerable dels drop-outs en ambdues mostres. En el cas de l'EXT-001, se'n detecten 9 al·lèlics i 4 drop-outs de locus a D1S1656, D5S818, D8S1179 i DXS10103; mentre que a la mostra EXT-004 el número puja a 7 al·lèlics i 15 drop-outs de locus, detectant així una pujada considerable de pèrdua d'informació. Finalment, en l'últim grau de dilució amb un input de 0.001 ng, l'augment de drop-outs es brusc. En el cas de l'EXT-001, dels a-STRs només se'n recuperen 2, D7S820 i el D9S1122, de la mateixa manera que amb els X-STRs, detectant-se el DXS7132 i el HPRTB. En canvi, en l'EXT-004 desapareix la informació per complet a través de drop-outs de locus i al·lèlics a excepció del DYS549, que es l'únic que s'ha pogut detectar.

Experiment Dilucions	ForenSeq™ DNA Signature Prep Kit Primer Mix A							
	27 aSTRs i AMEL		24 Y-STRs		7 X-STRs		94 iSNPs	
Mostra	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004
2.5 ng	/	0	/	0	/	0	/	0
1 ng	0	0	0	0	0	0	0	0
0.5 ng	/	0	/	0	/	0	/	0
0.1 ng	1 INC	1 INC 2 ADO	0	4 INC	0	0	5 INC	3 ADO 6 INC 1 LDO
0.05 ng	0	2 ADO 2 INC <b>1 ADI</b>	0	3 INC	0	2 INC	15 INC 3 LDO	2 ADO 27 INC 1 LDO
0.01 ng	8 ADO 5 INC 3 LDO <b>1 ADI</b>	6 INC 6 LDO 7 ADO	0	6 INC 8 LDO	1 ADO 2 INC 1 LDO	1 LDO 1 INC	2 ADO 34 LDO 41 INC	35 INC 33 LDO 5 ADO
0.001 ng	16 LDO 7 INC 3 ADO	25 LDO 1 INC 2 ADO	0	22 LDO 1 ADO	5 LDO	7 LDO	80 LDO 14 INC	89 LDO 5 INC

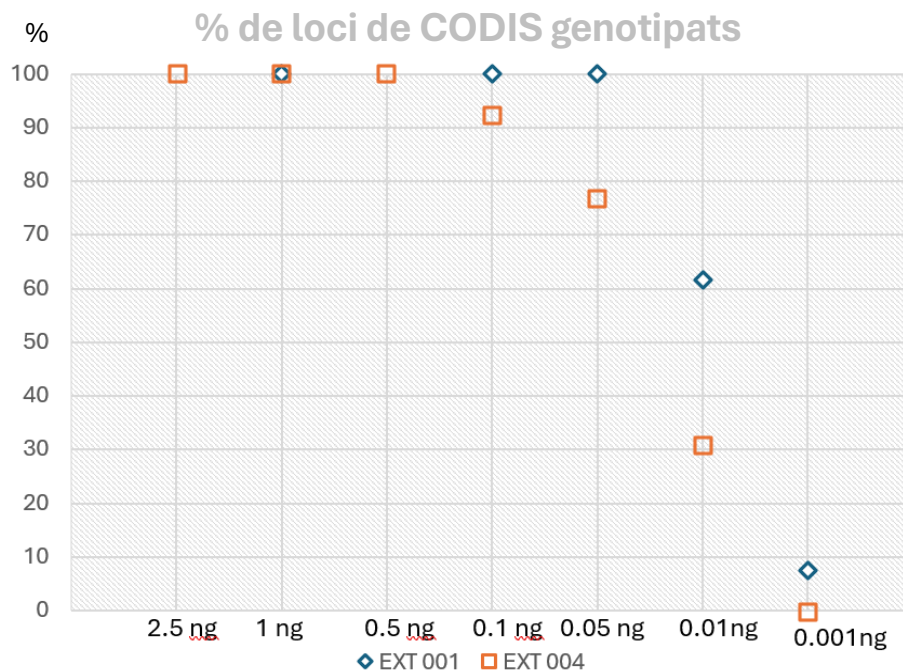
**Taula 10.** Anàlisi d'efectes estocàstics de l'experiment 2 amb diferents inputs d'ADN i per tant, diferents graus de dilució

En el cas de drop-outs referents a iSNPs, no es detecten fins a l'input de 0.1 ng, en concret a la mostra EXT-004 amb 3 al·lèlics i un de locus a rs2920816. A partir de l'input de 0.05 ng ja detectem un augment dels drop-outs a la mostra EXT-001 rs12997453 rs2342747 rs1031825, mentre que l'EXT-004 es manté estable. En el següent grau de dilució (0.01ng) ja perdem un gran nombre de marcadors, només podent-se recuperar un 18% en l'EXT-001 i aproximadament un 22% a l'EXT-004. Finalment, a l'últim input de 0.001 ng, no es detecta cap iSNP i per tant hi ha una pèrdua total d'informació.

En la totalitat de l'experiment només es detecten 2 drop-ins al·lèlics. Per una banda a 0.05 ng en la mostra EXT-004 s'observa un al·lel 11 al locus D16S539, mentre que a 0.01 ng es detecta el mateix al·lel número 11 a D17S1301. En aquest cas podria ser degut a algun tipus de contaminació creuada entre les dues mostres, ja que per a aquest mateix locus l'EXT-004 presenta un marcador homozigot 11, 11. Així i tot, en tots els

casos ha estat possible distingir quin era el vertader al·lel i quin l'artefacte degut a les mostres de referència.

S'ha tornat a comparar l'experiment amb els 13 locus del CODIS<sup>7</sup> més importants que s'han detectat en les diferents mostres exposades a diferents graus de dilució (Figura 5). Amb un input de 2.5 ng, 1 ng i 0.5 ng, s'observa que es detecten la totalitat dels STRs, mentre que a partir de l'input 0.1 ng, i en el cas de la mostra EXT-004, ja hi ha una davallada de 3 loci, un inconclusiu a D5S818, i dues pèrdues al·lèliques a D7S820 i D16S539 (Taula 11). En canvi, en l'EXT-001 es detecten la totalitat de loci del CODIS fins a l'input de mostra de 0.01 ng, quan ja hi ha només una recuperació d'aproximadament un 60% enfront un 30% de l'EXT-004. Ja en l'últim punt, tot i que en la mostra EXT-004 ja no es recupera cap tipus d'informació, en l'EXT-001 encara hi ha un 10% de marcadors que s'han pogut genotipar, en concret el D7S820. Tot i que amb un sol marcador no serviria per arribar a conclusions en casos de context forense, s'observa un gran potencial de la tècnica de seqüenciació paral·lelament massiva, arribant a detectar algun marcadors en mostres de reduïda concentració.



**Figura 5.** Percentatge de loci genotipats relatius al CODIS de l'experiment 2.

<sup>7</sup> Proyecto Biológico, Universitat d'Arizona (2000). Trad. de la Universitat d'Alcalá de Henares (2004). Disponible a: [https://biomodel.uah.es/epb/human\\_bio/activities/blackett2/str\\_codis.html](https://biomodel.uah.es/epb/human_bio/activities/blackett2/str_codis.html)

	2.5 ng		1 ng		0.5 ng		0.1 ng		0.05 ng		0.01 ng		0.001 ng	
	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004	EXT-001	EXT-004
D3S1358	/				/							INC	LDO	LDO
vWA	/				/						AD O	INC	LDO	LDO
FGA	/				/							ADO	LDO	ADO
D8S1179	/				/		INC				LDO	ADO	LDO	INC
D21S11	/				/								INC	LDO
D18S51	/				/							ADO	LDO	LDO
D5S818	/				/			AD O		INC	LDO	LDO	INC	LDO
D13S317	/				/								INC	LDO
D7S820	/				/					AD O		LDO		LDO
D16S539	/				/					AD O			LDO	LDO
TH01	/				/								ADO	LDO
TPOX	/				/						AD O	ADO	INC	LDO
CSF1PO	/				/						INC	ADO	LDO	LDO
AMEL	/				/						INC	LDO	LDO	LDO

**Taula 11.** Efectes estocàstics de l'experiment 2 produïts en els marcadors específics del CODIS

## **5. CONCLUSIONS**

Aquest estudi ha avaluat el rendiment del seqüenciador MiSeq FGx en mostres de dilucions variables, juntament amb mostres fragmentades de manera gradual a través del termociclador a 98°C, podent estudiar els percentatges de degradació. Les dades aportades mostren que quan l'input d'ADN de referència era de fins a 0.05 ng o el percentatge de degradació no supera el 65-70%, es generen resultats de seqüenciació acceptables. A partir d'un input de 0.01 ng i amb degradacions superiors a 70%, és a dir, amb temps de degradació superiors a 40 min, la qualitat dels resultats es redueix significativament.

En comparació amb altres estudis (Zhang *et al.*, 2018; Fattorini *et al.*, 2017; Jäger *et al.*, 2017), que analitzen els loci més vulnerables amb el sistema FGx, s'observa que tot i que si que hi ha coincidències en SNPs com l'rs1736442 o l'STR Penta D que si que es veuen afectats per efectes de degradació i de dilució, també se'n detecten d'altres com el Penta E, rs719366, DXS8378, DYS19, DYS392, DYS522, que es comencen a veure afectats en temps de degradació de 10 min, produint-se drop-outs ràpidament. En canvi hi ha d'altres, com el D8S1179 i el D9S1122, que es poden detectar fins a minuts

d'afectació de 80; el TH01 i vWA que també donen resultats acceptables a alts períodes de degradació o inclús el DYS549 detectant-se fins i tot en l'últim període de dilució de 0.001 ng. En el cas d'alguns SNPs, com el rs576261 i el rs2040411, s'observen de manera correcta al màxim punt d'exposició de calor de 120 minuts i amb dilucions elevades de 0.01 ng.

En el cas de drop-ins, s'han pogut detectar en diferents locus. En l'experiment 1, es detecten per primera vegada quan es combina la degradació a temps de 40 min amb un input de 0.5 ng, en el cas de DYS576 i D22S1045 (que coincideix amb el documentat a l'estudi de Zhang *et al.*, 2018), i el D9S1122. En el mateix temps, i reduint l'input a 0.1 ng, també se'n detecta un altre, a D20S482. Finalment a 60 i a 80 minuts n'apareixen dos més, a DXS7132 i a D17S130 (tant en l'experiment 1 com en el 2, aquest últim amb input de 0.01 ng). En el cas de l'experiment de dilucions també n'apareix un altre amb un input de 0.05 ng a D16S539. A partir de l'observació dels diferents casos de drop-in es pot arribar a concloure que és la combinació de la degradació, juntament amb la reducció de quantitat d'ADN, el factor que produeix més errors en la seqüenciació de les mostres.

Comparant els percentatges de degradació en els diferents intervals de temps, així com els percentatges de STRs i SNPs tipificats en cada punt, es poden correlacionar els resultats amb el fet de que a mesura que s'incrementa el temps de degradació i dilució, hi ha una major quantitat de drop-outs, tant al·lèlics com de loci, tal i com s'ha pogut veure en altres estudis com Zhang *et al.*, 2018; Sharma *et al.*, 2020. A partir d'aquí, es poden crear 3 grups en funció del tipus de degradació de les mostres. Per una banda, la degradació moderada, on s'han pogut tipificar entre el 100 i 80% dels marcadors (minuts 0 a 20), i que presenta un percentatge de degradació d'entre 0 a 45%. Després trobem la degradació severa, entre un 45% i 70% d'afectació, en el que es poden identificar entre un 80 i un 60% dels loci i que correspondria a temps d'exposició al calor situats entre els 20 i els 60 minuts. I finalment trobem la degradació avançada, del minut 60 fins al 120, que presenten una degradació de més del 70% i on només es poden seqüenciar fins al 59% dels marcadors, tant de STRs com de SNPs.

En relació a aquests tres grups de casos en mostres diligents, s'han comparat amb l'anàlisi dels marcadors que formen el CODIS utilitzant el mateix kit ForenSeq™ DNA Signature Prep. Es pot afirmar que el 100% dels marcadors han estat tipificats en el rang de degradació moderada, és a dir fins a 20 minuts, tot i que en el cas de la mostra EXT-

004, també s'ha pogut obtenir el perfil complet en el minut 40, amb una degradació per tant, més severa. Així i tot, en el temps de 60 minuts i en el cas de la mostra EXT-001 a 80 minuts, amb una degradació de més del 70%, s'han pogut identificar aproximadament més del 50% dels marcadors del CODIS, la qual cosa demostra, que tot i que hi hagi una gran quantitat de drop-outs, encara hi ha informació que es recupera de qualitat.

Segons Sharma *et al.* (2020), degut a la mida dels amplicons de les col·leccions de SNPs o altres marcadors no tradicionals, com per exemple insercions o delecions, podrien ser una bona manera d'estudiar les mostres degradades severes, tot i que no poden ser utilitzats actualment en les bases de dades i proporcionen menys informació que els STRs (Zhang *et al.*, 2018). Tot i que els iSNPs presenten una taxa de drop-in inferior a la dels STRs (Zhang *et al.*, 2018), com s'ha vist en aquest estudi i en altres anteriors (Zhang *et al.*, 2018; Jäger *et al.*, 2017; Fattorini *et al.*, 2017; Senst *et al.*, 2022), han estat afectats per els drop-outs abans, així que, es podria valorar l'opció de que poguessin ser exclosos per poder oferir més espai en la cel·la de flux per als STRs o millorar les reaccions d'aquests per beneficiar-se de la mida dels seus amplicons (Sharma *et al.*, 2020). Aquests resultats indiquen la importància de triar diferents tipus de marcadors segons la qualitat i les condicions inicials de la mostra per poder tenir el grau d'efectivitat més elevat possible.

En relació amb les restes humanes que han estat enterrades al sol durant llargs períodes de temps, i afectades per factors químics, físics i biològics, es possible que inhibidors com l'àcid hùmic o la força iònica juguin un paper important en la reducció del nombre de loci que es poden tipificar amb èxit (Zhang *et al.*, 2018). És per aquesta raó que tot i que l'escalfament artificial de mostres de referència a través d'un termociclador dona resultats fiables per a simular la degradació natural, s'ha de tenir en compte que pot haver diferències en les longituds dels fragments d'ADN en els models generats experimentalment al laboratori, ja que les mostres de camp sovint estan exposades a unes condicions específiques.

És molt important, per tant, l'estudi a través de l'ús de rèpliques, per poder millorar la fiabilitat i reduir la interferència associada a efectes estocàstics (Zhang *et al.*, 2018). Tenint en compte que la quantitat de mostres d'ADN es sempre limitada en contextos forenses, podria ser recomanable aplicar fins a tres rèpliques, i tot i que en l'estudi actual no s'ha pogut realitzar, almenys 2 repeticions per anàlisi per a confirmar els

al·lèls veritables també augmentaria el grau de qualitat de l'estudi (Zhang *et al.*, 2018). De la mateixa manera, les dades que s'han derivat de la present experimentació suggereixen que tant l'índex o el percentatge de degradació, així com la quantitat inicial d'input de material genètic, són precondicions importants que afecten a la tipificació. És per això que aquest estudi inicial podria ser utilitzat com a referència comparativa per a posteriors validacions i comparacions amb mostres de degradació natural.

Es pot concloure per tant, que la plataforma de MiSeq FGx així com el seu kit de seqüenciació en mostres a alts nivells de degradació i diferents graus de dilució, demostren fiabilitat i representen una alta sensibilitat dins el grup de tècniques de seqüenciació de nova generació. Tot i que és difícil poder simular els contextos específics forenses, i més concretament en aquells on el pas del temps i els danys postdeposicionals són presents en gran mesura, com per exemple en les Fosses de la Guerra Civil, s'ha vist que aquest estudi preliminar ha pogut establir els primers límits d'aquesta tècnica i detectar quins marcadors són més propensos a estar afectats per drop-outs al·lèlics o de loci, així com per drop-ins i altres artefactes. Aquest enfocament d'estudi així com la validació de la metodologia pot oferir resultats molt prometedors en una tecnologia que s'està implementant cada vegada més en els laboratoris de genètica forense i està millorant substancialment els resultats en termes de qualitat, efectivitat i sensibilitat dels anàlisis en relació als mètodes convencionals.

## **6. BIBLIOGRAFIA**

Alonso, A. (2004). “Conceptos básicos de ADN forense. Nuevas técnicas de investigación del delito”. *Intervenciones Corporales y ADN*, 1860-1871.

Brinkmann, B.; Klintschar, M.; Neuhuber, F.; Hühne, J. i Rolf, B. (1998). “Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat”. *Am. J. Hum. Genet* 62, 1408–1415.

Budowle, B.; Moretti, T.; Niezgoda, S. i Brown, B. (1998). “CODIS and PCR-based short tandem repeat loci: Law enforcement tools.” *Second European Symposium on Human Identification*, 1–16.

Calafell, F. *et al.* (2016). “An assessment of a massively parallel sequencing approach for the identification of individuals from mass graves of the Spanish Civil War (1936–1939)”. *Electrophoresis Journal* 37, 2841-2847.

Clayton, A.; Whitaker, T.; Urquhart, S.; Millican, J.; Downes, P. *et al.* (1995). “Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples”. *Biotechniques* 18 670–677.

Davenport, L.; Devesse, L.; Syndercombe, D. i Ballard, D. (2023). “Forensic identity SNPs: Characterisation of flanking region variation using massively parallel sequencing”. *Forensic Science International: Genetics* 64, 102847.

Edwards, A.; Civitello, A.; Hammond, H. i Caskey, C. (1991). “DNA typing and genetic mapping with trimeric and tetrameric tandem repeats”. *American Journal of Human Genetics* 49, 746–756.

Fattorini, P.; Previderé, C.; Carboni, I.; Marrubini, G; Sorçaburu-Cigliero, S.; Grignani, P.; Bertoglio, B.; Vatta, P. i Ricci, U. (2017). “Performance of the Forenseq™ DNA Signature Prep kit on highly degraded samples”. *Electrophoresis Journal* 38, 1163-1174.

Freire-Aradas, A.; Huidobro, V. i Carracedo, A. (2019). “Marcadores fenotípicos de interés forense”. A Crespillo M.C. i Barrio, P. A. (eds.), *Genética Forense, del Laboratorio a los Tribunales*. Diaz de Santos.

- García, V. i Sánchez, A. (2018). “Del perfil al retrato genético: dependencia de trayectoria en las tecnologías del ADN forense”. *Revista Colombiana de Filosofía de la Ciencia* 18 (37), 261-276.
- Gopinath, S. *et al.* (2016). “Developmental validation of the Yfiler1 Plus PCR Amplification Kit: An enhanced Y-STR multiplex for casework and database Applications”. *Forensic Science International: Genetics* 24, 164-175.
- Gross, T.; Fleckhaus, J. i Schneider, P. (2021). “Progress in the implementation of massively parallel sequencing for forensic genetics: Results of a European-wide survey among professional users”. *International Journal of Legal Medicine* 135, 1425–1432.
- Jäger, A. *et al.* (2017). “Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories”. *Forensic Science International: Genetics* 28, 52-70
- Kayser, M. (2017). “Forensic use of Y-chromosome DNA: a general overview”. *Human Genetics* 136, 621-635.
- Latham, K. i Miller, J. (2019). “DNA recovery and analysis from skeletal material in modern forensic contexts”. *Forensic Sciences Research* 4 (1), 51-59.
- Magaña, J.; Arenas-Sordo, M. i Gómez, R. (2009). “La electroforesis capilar como una nueva estrategia en la medicina y el diagnóstico clínico”. *Revista Médica de Chile* 137, 946-956.
- Nachman, M. i Crowell, S. (2000). “Estimate of the mutation rate per nucleotide in humans”. *Genetics* 156 (1), 297–304.
- Novroski, N. i Cihlar, J. (2022). “Evolution of single-nucleotide polymorphism use in forensic genetics”. *WIREs Forensic Science* 4, 1459: 1476.
- Pifarré, A. (2019). “Marcadores STR autosómicos de interés forense: amplificación y artefactos”. A Crespillo M.C. i Barrio, P. A. (eds.), *Genética Forense, del Laboratorio a los Tribunales*. Diaz de Santos

- Senst, A.; Caliebe, A.; Scheurer, E. i Schluz, I. (2022). “Validation and beyond: Next generation sequencing of forensic casework samples including challenging tissue samples from altered human corpses using the MiSeq FGx System”. *Journal of Forensic Sciences* 67, 1382–1398.
- Sharma, V.; van der Plaat, D.; Liu, Y. i Wurmbach, E. (2020). “Analyzing degraded DNA and challenging samples using the ForenSeq™ DNA Signature Prep kit”. *Science & Justice* 60, 243-252.
- Syndercombe, D. (2021a). “Mitochondrial DNA in forensic use”. *Emerging Topics in Life Sciences* 5, 415-426.
- Syndercombe, D. (2021b). “The Y chromosome and its use in forensic DNA analysis”. *Emerging Topics in Life Sciences* 5, 427-441.
- Udogadi, N.S., *et al.*(2020). “Forensic DNA Profiling: Autosomal Short Tandem Repeat as a Prominent Marker in Crime Investigation”. *Malays J Med Science* 27 (4), 22-35.
- Watson J., *et al.* (2016). *Biología Molecular del Gen*. Médica Panamericana
- Zhang, Q.; Zhou, Z.; Liu, Q.; Liu, L.; Shao, L. *et al.* (2018). “Evaluation of the performance of Illumina’s ForenSeq™ System on serially degraded samples”. *Electrophoresis Jorunal* 39, 2674-2684.