Ana Belén Gutiérrez Reyes

MASTER'S THESIS

MASTER'S DEGREE IN FORENSIC GENETICS, PHYSICS AND CHEMISTRY

Phenotypic prediction in forensic genetics: Evaluation of Real-Time PCR with Melting Curve Analysis for SNP detection

Professional supervisors: Dr. Cristina Gutiérrez Fornés and Clara Benavent Bofill

Academic supervisor: Dr. Ximena Terra Barbadora

In cooperation with: Laboratori Clínic ICS Camp de Tarragona, Hospital Joan XXIII

UNIVERSITAT ROVIRA i VIRGILI

TARRAGONA, JUNIO 2025

INDEX

1. ABSTRACT

Single nucleotide polymorphisms (SNPs) are the most common type of genetic variation and the main basis for individual and population-level differences. Beyond their role in disease predisposition, SNP analysis is applied in pharmacogenetics, phenotypic trait prediction, and forensic genetics.

The aim of this study was to assess the potential forensic applicability of a real-time PCR technique combined with melting curve analysis for genotyping in the context of Forensic DNA Phenotyping (FDP). This methodology was evaluated in clinical samples from patients who required Fluoropyrimidine (FL)-based chemotherapy, which can cause severe toxicity in individuals carrying pathogenic variants in the dihydropyrimidine dehydrogenase gene (*DPYD*), which encodes dihydropyrimidine dehydrogenase (DPD), the main enzyme responsible for FL catabolism. Four clinically relevant *DPYD* variants (*DPYD*2A, DPYD*13*, c.2846A>T (*p.D949V*), and *HapB3*) were evaluated.

The analysis of three simulated clinical cases using LightMix® Multi-SNiP *DPYD* kits demonstrated the reliability and consistency of the technique for identifying the selected variants, with melting temperature (Tm) values consistent with manufacturer specifications and precise genotype assignment. This enabled therapeutic recommendations aimed at reducing FL-induced toxicity risk according to pharmacogenetic guidelines.

Additionally, the forensic applicability of this technique was critically assessed, identifying limitations such as the need for DNA concentrations above 5 ng/µL, requirement for specific controls for each SNP, difficulty interpreting mixed profiles, and high per-SNP cost compared to next-generation sequencing (NGS) or microarrays. Nonetheless, the method offers advantages like short processing time, suitability for degraded DNA, and potential lyophilized kits for forensic use.

In conclusion, while this methodology proves useful in clinical pharmacogenetics, its implementation in forensic phenotypic prediction requires extensive validation, population-specific panel design, and standardized protocols to ensure result reliability, reproducibility, and evidentiary value.

Keywords: Forensic DNA Phenotyping, Single Nucleotide Polymorphisms, DPYD Genotyping, Melting Curve Analysis and Real-Time PCR.

2. ABBREVIATIONS

5-FU	5-Fluorouracil
AIM	Ancestry Informative Marker
CoDIS	Combined DNA Index System
CPIC	Clinical Pharmacogenetics Implementation Consortium
DPD	Dihydropyrimidine Dehydrogenase
DPWG	Dutch Association for the Advancement of Pharmacy
DPYD	Dihydropyrimidine Dehydrogenase Gene
EVC	Externally Visible Characteristics
FDP	Forensic DNA Phenotyping
FL	Fluoropyrimidines
FRET	Fluorescence Resonance Energy Transfer
GIANT	Genetic Investigation of ANthropometric Traits
GWAS	Genome-Wide Association Studies
HRM	High-Resolution Melting
ICMP	International Commission on Missing Persons
KASP	Kompetitive Allele Specific PCR
MPS	Massive Parallel Sequencing
NGS	Next-Generation Sequencing
qPCR	Quantitative PCR
SEFF	*Sociedad Española de Farmacogenética y Farmacogenómica*
SIL	*Sistema Informático de Laboratorio*
SNP	Single Nucleotide Polymorphism
SNS	*Sistema Nacional de Salud*
STR	Short Tandem Repeat
Tm	Melting Temperature
VISAGE	VISible Attributes through GEnomics

3. INTRODUCTION

3.1. Concept of SNPs and their relevance in human genetics

Single nucleotide polymorphisms, better known as SNPs, are genomic variants occurring at a single base position in DNA[1,2]. To date, more than 950 million human SNPs have been recorded in the dbSNP database, reflecting their remarkably high frequency as a form of genetic variation[3]. On average, one polymorphism occurs every 100 to 1000 bases within the human genome, making SNPs the most common type of genetic variation and the principal basis for individual and population-level differences[1–3].

Advances in genomic sequencing demonstrated that SNP distribution patterns vary both among individuals and across different populations, enabling the identification of specific genomic regions potentially involved in certain phenotypic traits or disease susceptibility[1,3]. Although the majority of SNPs identified in the human genome have no phenotypic effect (silent mutations), some variants were found to alter gene function or expression levels (such as missense and nonsense mutations)[3,4]. Consequently, considerable efforts have been devoted in recent decades to identifying these SNPs and investigating the mechanisms by which they contribute to phenotypic differences, particularly concerning disease susceptibility and individual drug response[1,3].

From a technical perspective, SNPs exhibit several properties that make them especially valuable in genetic studies. Their high frequency in the genome provides a large pool of potential markers for biomedical and forensic applications, allowing the selection of those with an appropriate allele frequency distribution for specific purposes[5]. Additionally, they are extremely stable markers, with mutation rates as low as 2×10^{-8} per generation, notably lower than those of other genetic markers. This stability proves particularly advantageous in kinship analyses, including those involving complex pedigrees, and contributes to interpopulation variability in allele frequencies. Furthermore, SNPs represent the simplest form of DNA polymorphism. Their detection and genotyping methods, although relatively complex, are amenable to automation and support the development of multiplex reactions capable of incorporating a high number of markers in a single genotyping assay[5]. This feature optimizes the balance between available sample quantity and the amount of genetic information obtained. Moreover, because they involve single-base changes, it is possible to design PCR reactions generating short-length amplicons, which is especially advantageous when working with degraded DNA[5].

To identify SNPs associated with specific traits or diseases, candidate gene studies were traditionally employed, focusing on variants affecting coding regions[2]. Subsequently, research also addressed the functional effects of variants located in non-coding regions, assessing their influence on molecular phenotypes[3]. Currently, the application of SNP analysis extends beyond disease predisposition studies and has been integrated into pharmacogenetics, phenotypic trait prediction, and forensic genetics, areas that will be comprehensively addressed throughout this work[1,3,5].

3.2. Clinical use of SNPs: Pharmacogenetics and the *DPYD* case

Pharmacogenetics is a branch of genetics that studies how individual variations in DNA, including SNPs, influence each person's response to specific drugs[6]. Its main objective is to identify the most effective and safest medication and dosage for each individual, anticipating adverse reactions or a lack of therapeutic efficacy[7–9]. This discipline represents one of the foundations of precision medicine, which integrates genetic, environmental, and lifestyle information to individualize treatments, improve therapeutic outcomes, and minimize risks[9]. In clinical practice, pharmacogenetics is applied through tests that detect SNPs or other variants in genes involved in drug metabolism, transport, or action. These analyses allowed clinicians to predict whether a patient would exhibit a normal, weak, excessive, or absent response to certain treatments, as well as estimate the rate of absorption, metabolism, or elimination of a drug, and the likelihood of adverse effects[9]. Therefore, both genotypic and phenotypic characterization of key enzymes in drug metabolism proved highly valuable for therapeutic selection and for improving clinical outcomes in affected patients[7].

A particularly relevant example of the clinical application of pharmacogenetics was the use of FLs, antimetabolite drugs introduced in the 1950s for the treatment of solid tumors such as colorectal, breast, pancreatic, or gastric cancer[10–12]. Their main compounds included 5-fluorouracil (5-FU), capecitabine, and tegafur[13,14], the latter two being prodrugs enzymatically converted into 5-FU. The mechanism of action of these agents consists of the inhibition of thymidylate synthase, disrupting pyrimidine synthesis and impairing DNA and RNA integrity in tumor cells[11] (**Figure 1**).

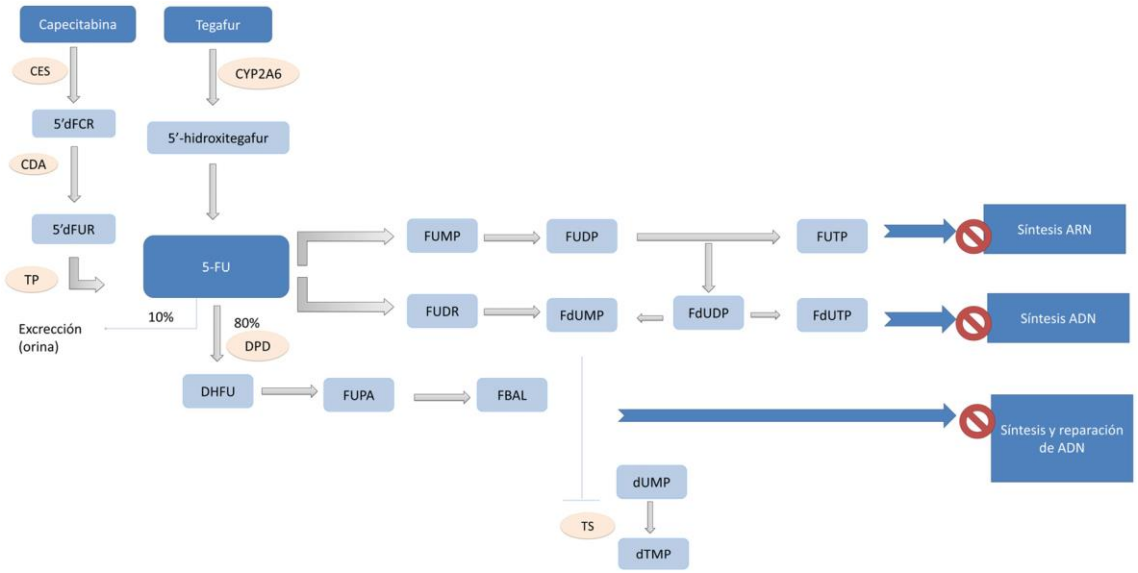


Figure 1. Schematic representation of the metabolic pathways of fluoropyrimidines and their prodrugs. Adapted from *"Fluoropirimidinas, Sociedad Española de Farmacogenética y Farmacogenómica (SEFF)"*[10].

Although their therapeutic efficacy was well established, the clinical use of FLs remained limited by a narrow therapeutic window and a toxicity profile highly dependent on the individual's metabolic capacity[14]. This toxicity commonly manifested through gastrointestinal symptoms (nausea, vomiting, diarrhea, mucositis, or stomatitis), bone marrow aplasia, or dermatological drug reactions such as hand-foot syndrome[15–17]. Managing this clinical risk posed a significant challenge in oncological practice[11,15].

The metabolism of FLs depends on the activity of the enzyme DPD, the rate-limiting step in the detoxification of FLs, catalyzing the conversion of 5-FU into inactive metabolites (**Figure 1**). This enzyme is encoded by the *DPYD* gene, located on chromosome 1p22, spanning 843 kilobases and comprising 23 coding exons[10,14,15]. Partial (observed in 3% to 8% of the Caucasian population) or complete (0.01% to 0.5%) deficiencies of this enzyme, determined by genetic variants, could lead to severe, even fatal, toxicities following FL administration[10,11,15,17]. For this reason, several international guidelines, such as the Clinical Pharmacogenetics Implementation Consortium (CPIC), the Royal Dutch Association for the Advancement of Pharmacy (DPWG), and Spain's *Sistema Nacional de Salud* (SNS), considered it essential to perform a genetic test prior to FL administration. This genotyping is focused on detecting four clinically significant SNPs in the *DPYD* gene: c.1905+1G>A (*DPYD*2A; rs3918290), c.1679T>G (DPYD*13*; rs55886062), c.2846A>T (*p.D949V*; rs67376798), and c.1129-5923C>G and c.1236G>A (*HapB3*; rs75017182 and rs56038477)[10,11,15,18]. The detection of these variants enabled

classification of patients according to their genotype and metabolic phenotype, establishing specific dose-adjustment or treatment contraindication recommendations (**Table 1**)[10,15].

Table 1. Definition of the metabolic phenotype based on the *DPYD* genotype and recommended fluoropyrimidine dosing for each case. Adapted from *"Fluoropirimidinas, Sociedad Española de Farmacogenética y Farmacogenómica"*[10] and "Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Dihydropyrimidine Dehydrogenase Genotype and Fluoropyrimidine Dosing: 2017 Update"[15].

DPD METABOLIC PHENOTYPE	*DPYD* GENE GENOTYPE	GENOTYPE (EXAMPLES)	IMPLICATIONS	DOSING RECOMMENDATIONS
Normal Metabolizer	Individuals with two normal-function alleles.	**1/*1*	Normal DPD activity and standard risk of fluoropyrimidine-related toxicity.	According to the product information.
Intermediate Metabolizer	Individuals with one normal-function allele and one reduced-function allele.	**1/HapB3*	DPD activity decreased by approximately 30% to 70%, leading to an increased risk of severe or even fatal toxicity when treated with fluoropyrimidines.	The initial dose should be reduced by 50%. Subsequent dose adjustments should be made based on toxicity or pharmacokinetic data.
	Individuals with one normal-function allele and one no-function allele, or with two reduced-function alleles.	**1/*2A* *p.D949V/HapB3*		
Slow Metabolizer	Individuals with one no-function allele and one reduced-function allele.	*HapB3/*13*	Complete DPD deficiency, resulting in a markedly increased risk of severe or potentially fatal toxicity when treated with fluoropyrimidines.	Treatment with fluoropyrimidines is contraindicated. Alternative agents should be considered.
	Individuals with two no-function alleles.	**2A/*13*		

Notably, the identification of these variants not only allows for the anticipation of toxicity risks but also facilitates the planning of personalized therapeutic strategies by adjusting FL dosing based on the patient's genetic profile. Consequently, this approach reduces the incidence of severe toxicities, optimizes clinical efficacy, and avoids empirical dose adjustment through trial and error, which increases healthcare costs and prolongs treatment durations[9,13–15,19]. Therefore, genotypic and phenotypic classification of patients using SNP analysis consolidated itself as a decisive tool for clinical decision-making. The future of precision oncology relies on the widespread implementation of pharmacogenetic studies such as those targeting *DPYD*, not only to prevent potentially fatal toxicities but also to individualize treatment regimens from the outset, thereby improving both patient safety and quality of life[10,13,15,19].

3.3. Transition to the forensic field: SNPs in phenotypic prediction

Advances in SNP detection applied to clinical practice can also be extremely useful for phenotyping physical traits in the field of forensics. These technological advances applied in forensic genetics have enabled laboratories to expand their capabilities beyond traditional methods of human identification based on short tandem repeat (STR) polymorphisms[20]. These methods, considered the gold standard to date, require comparison between unknown DNA samples and profiles registered in forensic databases (such as the Combined DNA Index System, CoDIS) or those belonging to suspects[21–23]. However, in cases lacking eyewitnesses or involving unreliable testimony, these techniques present significant limitations[20,24,25]. In this context, FDP has emerged as a complementary tool that allows for the prediction of certain externally visible characteristics (EVCs) through the analysis of DNA recovered from crime scenes or unidentified human remains[20,22,24,25]. Unlike traditional comparative methods, FDP does not rely on the existence of prior profiles in databases, thereby significantly broadening investigative possibilities in solving unsolved crimes and identifying disaster victims[20,26].

FDP predictions are based on the identification of SNPs associated with specific phenotypic traits, such as eye, hair, or skin color[25,27]. Thanks to the development of Genome-Wide Association Studies (GWAS), several genetic variants responsible for a substantial proportion of the variability observed in these traits have been identified[27]. Although current understanding of the genetics underlying human traits remains incomplete, recent progress has driven the development of prediction systems with acceptable practical reliability, particularly for pigmentary traits[24,26]. Beyond its application in case resolution, FDP acts as a "biological witness" that guides investigations and helps reduce the pool of potential suspects, especially in scenarios where no database matches are available or eyewitness accounts are unreliable[24]. In such cases, probabilistic predictions generated from DNA analysis allow for the formulation of hypotheses regarding the physical appearance of the DNA donor, with a quantifiable level of statistical confidence[22]. This process relies on validated statistical models built from extensive databases linking known genotypes to phenotypes, enabling the inference of external characteristics without requiring prior profile comparisons[22,25,27]. Additionally, SNP analysis can incorporate Ancestry Informative Markers (AIMs), providing information about the biogeographical origin of an individual; however, it must

be noted that in admixed populations, this information does not always correlate directly with physical appearance[21].

One of the main advantages of SNP analysis in forensic contexts is its suitability for working with degraded or mixed DNA samples[5,22,25,28]. SNPs exhibit low mutation rates and high molecular stability, allowing for the design of small amplicons (less than 90 bp) that facilitate amplification in degraded samples[5,22,25]. This feature enables the multiplexing of numerous markers within a single reaction, optimizing the use of limited biological material[5]. Furthermore, the incorporation of microhaplotypes (clusters of closely linked SNPs) allows for the extraction of additional information regarding individual identification and genetic ancestry, even under adverse conditions[28].

The practical utility of FDP has already been demonstrated in countries such as Canada, the United States, the United Kingdom, France, Germany, and the Netherlands. A notable example is the 2017 Candra Alston case in the United States, in which phenotypic prediction enabled the generation of a composite image from DNA recovered at the crime scene, ultimately leading to the identification and apprehension of the perpetrator[21,28].

In addition to its investigative value, FDP offers significant ethical and legal advantages[20,25]. Unlike complete genetic profiles or ancestry information, phenotypic predictions pertain to external characteristics and, in most jurisdictions, are not considered sensitive or private genetic data. This facilitates their integration into police investigations without infringing upon fundamental rights or necessitating the storage of full genetic profiles in forensic databases[20,25]. Thanks to the implementation of technologies such as Massive Parallel Sequencing (MPS) and the development of specific SNP panels, the capacity to infer physical traits from anonymous DNA has increased substantially[24]. As a result, FDP has established itself as an effective, ethical, and complementary tool in forensic genetics, with growing potential for future applications in predicting additional traits and resolving cases lacking prior database matches[5,22,28].

3.3.1. Projects, Consortia, and International Initiatives

As previously mentioned, the analysis of SNPs associated with FDP has been established as a highly valuable tool in forensic genetics[22,24]. However, to ensure the reliability and universal applicability of these predictive systems, it is essential to minimize population biases and validate their accuracy across diverse ethnic and geographic groups[22]. To this end, various international initiatives and consortia have been developed in recent years,

aiming to enhance the prediction of EVCs through SNP analysis, while expanding the range of traits assessed and refining the accuracy of the predictive models employed[22,26].

3.3.1.1. HIrisPlex-S

HIrisPlex-S is a system designed to predict eye, hair, and skin color from DNA, based on two prior developments: IrisPlex (focused on eye color) and HIrisPlex (predicting both eye and hair color)[24]. IrisPlex utilized six SNPs in pigmentation-related genes, achieving over 90% accuracy for distinguishing between blue and brown eyes[21]. In 2013, HIrisPlex incorporated eighteen additional SNPs to predict hair color, reaching an accuracy range between 75% and 92%[21,25]. Subsequently, HIrisPlex-S integrated seventeen SNPs for skin color prediction, bringing the total to forty-one markers analyzed[24,29]. The system operates through two multiplex assays and multivariate statistical models, capable of producing reliable results from minimal quantities of DNA, including degraded samples[24]. It has been validated through international collaborative studies, demonstrating high precision and reliability[21]. Additionally, an open-access interactive tool (https://hirisplex.erasmusmc.nl/) allows users to input genotype data for the forty-one SNPs and obtain probabilistic predictions for these three phenotypic traits, establishing HIrisPlex-S as a reference standard in forensic phenotyping[20–22,24,29].

3.3.1.2. VISAGE (VISible Attributes through Genomics)

VISAGE is a European consortium launched in 2017 under the Horizon 2020 framework, dedicated to advancing FDP through SNP analysis[22,28]. Its objective is to develop robust tools for inferring physical appearance, age, and biogeographical ancestry from DNA[22,24,28]. Among its major contributions is the implementation of MPS for the analysis of multiple markers in limited or degraded forensic samples, along with the inclusion of novel predictive SNPs[21,24]. VISAGE has also promoted the development of dedicated software solutions for data interpretation, supporting their future integration into routine forensic and investigative practice[24,28]. Moreover, VISAGE is establishing an international SNP database validated in diverse population groups to mitigate bias and extend the range of reliably predictable phenotypic traits. This initiative represents a significant advancement in forensic identification, particularly in cases lacking suspects or comparative genetic profiles[24,25].

3.3.1.3. Other relevant initiatives

In addition to the established projects mentioned above, several other forensic initiatives are exploring the application of SNP-based phenotyping to improve genetic identification practices[22,24]. Noteworthy among these is the work of the *Instituto de Ciencias Forenses* at the Universidad de Santiago de Compostela, in collaboration with the International Commission on Missing Persons (ICMP) in The Hague, which has developed a panel of 1,400 SNPs for the identification of missing persons, specifically tailored to international contexts lacking direct reference samples[28]. Furthermore, some private companies, such as Parabon Nanolabs (a U.S.-based firm specialized in FDP and genetic genealogy) have begun applying genetic phenotyping to real forensic cases through DNA-based composite sketches. However, these methodologies currently lack peer-reviewed scientific publications and formal validation studies[21]. Additionally, international consortia such as GIANT (Genetic Investigation of ANthropometric Traits), which focus on identifying genetic variants associated with human physical traits, may contribute novel markers of forensic interest in the near future[25].

3.3.2. Types of phenotypic predictions based on SNPs

Currently, the analysis of SNPs allows for the inference of certain EVCs. However, the reliability of these predictions varies depending on the specific trait considered[24]. Pigmentation-related traits, such as eye, hair, and skin color, achieve the highest levels of accuracy, owing to the identification of genetic variants directly involved in these biological processes[24,28]. In contrast, other phenotypes, including hair shape and texture, the presence of freckles, predisposition to baldness, or stature, exhibit more limited predictive capacity and still require extensive validation. At present, the most reliable predictions focus on simple traits, such as blue eyes, brown hair, or fair skin, while research continues to expand these analyses to other physical features[21,24]. This disparity is largely attributable to the fact that pigmentation traits are predominantly determined by a relatively small number of genes that concentrate most of the relevant genetic information. In contrast, other phenotypes are genetically more complex and remain in earlier stages of investigation[26].

3.3.2.1. Prediction of skin, eye, and hair color

Skin color is a complex trait influenced by multiple genes involved in melanin production and has been evolutionarily shaped by ultraviolet radiation exposure in different

geographical regions[24,25,27]. Several SNPs have been identified as key predictors of skin color, particularly in admixed populations and in South Asian groups[21,24,25]. However, predictions do not yet achieve complete reliability and are often combined with biogeographical ancestry data to improve their forensic value[22,25]. Regarding eye color, the HERC2 and OCA2 genes play a central role, enabling highly accurate prediction of blue and brown tones through systems such as HIrisPlex-S[24]. Nonetheless, intermediate colors (e.g., green, hazel) remain difficult to predict due to genetic complexity and limited differentiation between the associated genetic profiles, motivating ongoing searches for new biomarkers[21,22,27]. This trait stands out for its high forensic reliability, even when working with limited or degraded DNA samples[22,24,25]. As for hair color, its prediction primarily depends on the MC1R gene, which regulates the balance between eumelanin (dark pigment) and pheomelanin (light pigment), complemented by other genes[21,24,27]. Current systems reliably predict black, red, and brown hair, while blonde hair prediction remains less precise due to factors such as natural color changes with age, sex, and aging (variables still insufficiently integrated into predictive models)[22,24,27]. Although red hair prediction is notably robust, owing to its strong genetic association with MC1R, there is a recognized need to refine prediction models for other shades and to account for individual and temporal phenotypic variability[21,22,25].

3.3.2.2. Prediction of emerging traits

At present, phenotypic prediction aims to extend beyond traditional pigmentation traits, toward emerging characteristics such as hair shape and texture, freckling, baldness, stature, and other complex phenotypes. However, the forensic reliability and applicability of these predictions vary considerably according to the trait[21,24,25,27]. Some predictions, such as hair shape, texture, or baldness, are highly complex because they depend on multiple genetic variants[21,24,25,27]. Others, like freckling, show a strong association with a single gene[22,25]. In contrast, stature (despite being linked to multiple SNPs) remains limited by the influence of environmental and hormonal factors[27]. A common feature across all these traits is the need for further research, validation in diverse populations, and the development of multivariable panels integrating multiple SNPs to improve both the accuracy and the real-world applicability of these predictions in forensic genetics[21,22,24,25,27,28].

3.3.3. Current perspective, limitations, and ethical-legal challenges in forensic phenotypic prediction

From an ethical and legal standpoint, this type of analysis raises concerns regarding potential violations of privacy, the risk of discrimination, and the misuse of sensitive genetic information[21,24,26,28]. Consequently, it is recommended that phenotypic predictions be employed solely as auxiliary tools within investigative processes, avoiding definitive interpretations[22,24,26]. Furthermore, the frequent conflation between phenotypic prediction and biogeographical ancestry inference presents an additional risk of racial bias[22,28]. Despite these limitations, the future of FDP appears promising. Emerging technologies such as massively parallel sequencing and artificial intelligence are expected to enhance the accuracy of phenotype predictions, even from highly degraded DNA samples[24,27,28,30]. In this regard, certain Spanish projects are currently exploring the prediction of facial morphological traits from DNA, although significant technical challenges persist[25,31]. To ensure the ethical and responsible use of these techniques, it is essential to establish clear principles of transparency, provide specialized training for involved professionals, implement a robust legal framework, and adopt social protection measures aimed at preventing prejudice and unfounded suspicion[21,22,26].

3.4. SNP detection techniques

Currently, several methodologies are available for the detection and genotyping of SNPs, each with specific characteristics, capabilities, and limitations that determine their suitability according to the type of study and available resources. These methodologies can be classified into three main categories: PCR-based techniques, DNA microarrays, and NGS[32].

- **PCR-Based SNP Genotyping:**

PCR-based techniques typically employ specific primers to amplify DNA regions containing the variants of interest and detect their presence through different strategies, such as fluorescence or melting curve analysis, commonly using quantitative PCR (qPCR). This category includes methods such as Kompetitive Allele Specific PCR (KASP), TaqMan assays, OpenArray systems, and High-Resolution Melting (HRM) analysis[32].

- **Microarray SNP Genotyping:**

DNA microarrays operate through chips containing thousands of specific probes that hybridize with the sample DNA when the variant is present. Detection is performed using specialized scanners that capture the fluorescence signal, and the resulting data are processed by dedicated software that normalizes and filters the information, facilitating genotype interpretation[32]

- **Next-Generation Sequencing of SNP Genotyping:**

NGS enables the highly accurate analysis of medium to large numbers of SNPs in variable sample sizes, owing to reduced costs and excellent precision. The standard procedure involves an initial multiplex PCR to simultaneously amplify the SNPs of interest, followed by amplicon indexing and massively parallel sequencing[32].

Each methodology offers specific advantages and limitations that determine its appropriateness based on the project type, budget, time constraints, and technical resources. Some of these benefits and limitations are detailed in **Table 2**[32].

Table 2. Advantages and limitations of the different methodologies for SNP detection and genotyping. Adapted from *Agrigenomic Diversity Unleashed: Current Single Nucleotide Polymorphism Genotyping Methods for the Agricultural Sciences.*

Methodology	Advantages	Limitations
PCR-based Techniques	Simple, rapid, and cost-effective methods.	Low multiplexing capacity (ranging from 1 to 200 SNPs per assay, depending on the technique).
	Require accessible equipment and personnel with basic PCR expertise.	Limited ability to analyze rare SNPs or variants located in poorly characterized regions.
	Ideal for high throughput genotyping of a limited number of well-characterized SNPs.	Require prior characterization of SNPs and specific primers for each target.
DNA Microarrays	Enable simultaneous analysis of up to 2.5 million SNPs per sample.	High initial cost for equipment and reagents.
	Automated workflows and simplified data analysis using dedicated software.	Dependence on custom-designed chips and proprietary scanners from manufacturers such as Illumina or Affymetrix.
	Suitable for large-scale studies in medical, agricultural, or forensic genetics.	Less flexibility to modify SNP panels once manufactured.
Next-Generation Sequencing	High accuracy and flexibility in the number of SNPs analyzed (ranging from hundreds to tens of thousands).	Greater technical and bioinformatic complexity.
	Possibility of customizing SNP panels according to project needs.	Requires sequencing equipment and expertise in data analysis.
	Increasingly competitive costs, especially for medium-scale studies.	Longer processing times compared to PCR-based techniques.

3.4.1. PCR with Melting Curve Analysis: The *DPYD* Case

Real-time PCR with melting curve analysis is a versatile technique that can be applied to any situation requiring the identification of SNPs within DNA samples. Its capacity to differentiate allelic variants based on specific hybridization and Tm properties makes it suitable not only for clinical pharmacogenetics but also for forensic applications and population studies. This section will address its application for the genotyping of the DPYD gene.

In *DPYD* genotyping, one of the techniques employed is real-time PCR using HybProbe probes, which belongs to the group of PCR-based techniques and represents a variant of Fluorescence Resonance Energy Transfer (FRET)-based technologies. This technique allows for the rapid and precise detection of SNPs from DNA samples. For *DPYD*, this is one of the routine methods used to detect the four clinically relevant variants. In this approach, a *DPYD* fragment is amplified using specific primers, and detection is carried out via a mutation-specific probe. These probes are characterized by emitting fluorescence only when hybridized to complementary DNA. The probe binds to a region of the amplified fragment that includes the mutation site. Following amplification, a melting curve analysis is performed by gradually increasing the temperature until the Tm is reached, at which point the probe dissociates from the complementary DNA, causing a decrease in fluorescence signal. Any mismatch within the probe binding region (caused by the presence of a mutation) destabilizes the DNA-probe hybrid, resulting in probe dissociation at a lower temperature, which is reflected as a leftward shift in the Tm on the melting curve. The probes selected for this system are optimized to provide maximum discrimination between wild-type and mutant sequences. Genotypic results are interpreted by comparing the obtained Tm values with those of the reference standards supplied with the kit, enabling the identification of the different genotypes or phenotypes (normal homozygous, heterozygous, and mutant homozygous). This technique represents a rapid tool for estimating the patient's likely phenotype and preventing toxicity events during FL-based chemotherapy administration.

Note: The specific kits and protocols employed for this technique, including the LightMix® In-Vitro Diagnostics Kit Multi-SNiP *DPYD* and LightCycler® FastStart DNA Master HybProbe, are detailed in the Methodology and Materials section.

4. HYPOTHESIS

Given the growing interest and development of FDP in recent years, the investigation and validation of new methodologies for the detection of SNPs associated with EVCs represents a promising approach to expand the capabilities of forensic genetics. In this context, the hypothesis proposes that real-time PCR with melting curve analysis, routinely applied in clinical pharmacogenetics for the genotyping of *DPYD* variants, constitutes a robust, sensitive, and relatively simple tool for the detection of SNPs related to phenotypic traits in forensic DNA samples, including those of limited quantity or degraded quality.

5. OBJECTIVES

This study aims to describe and critically assess the methodological application of real-time PCR with melting curve analysis for the genotyping of SNPs in the *DPYD* gene, and to explore its potential utility in forensic genetics for phenotypic prediction based on SNP analysis.

Specific Objectives

1. To document the protocol currently used for *DPYD* variant genotyping in a hospital setting, detailing sample preparation, amplification, and melting curve interpretation procedures.
2. To review the key technical features of this method that may be relevant for forensic applications, including its suitability for degraded DNA, capacity for generating short amplicons, and potential for multiplexing.
3. To discuss the theoretical feasibility of applying this technique to forensic prediction of phenotypic traits, considering its prospective advantages and current limitations.

6. METHODOLOGY AND MATERIALS

This section describes the protocol routinely applied in a clinical pharmacogenetics setting for the genotyping of *DPYD* variants using real-time PCR and melting curve analysis. Given the technical characteristics of this methodology, it holds potential applicability in forensic genetics, particularly for the detection of SNPs associated with externally visible traits in DNA samples of compromised quality. The detailed description of this procedure aims to assess its technical features and discuss its possible extrapolation to forensic phenotypic prediction.

6.1. Experimental design

All oncologic patients diagnosed at Hospital Joan XXIII in Tarragona who require FL-based chemotherapy routinely underwent *DPYD* genotyping for the detection of the four clinically relevant variants (*DPYD*2A*, *DPYD*13*, *p.D949V*, and *HapB3*), in order to prevent severe toxicities associated to this treatment.

6.2. DNA purification

Peripheral blood samples anticoagulated with EDTA were used to purify genomic DNA using the QIAamp DNA Blood Mini Kit (250) (Qiagen, Barcelona, Spain), as shown in **Figure 2**. The leukocyte fraction, or buffy coat, was obtained by centrifugation of whole blood at 3500 rpm for 20 minutes at 4 °C. After centrifugation, three fractions were identified: plasma (upper), buffy coat (middle), and erythrocytes (lower). The middle fraction was carefully collected for subsequent processing.

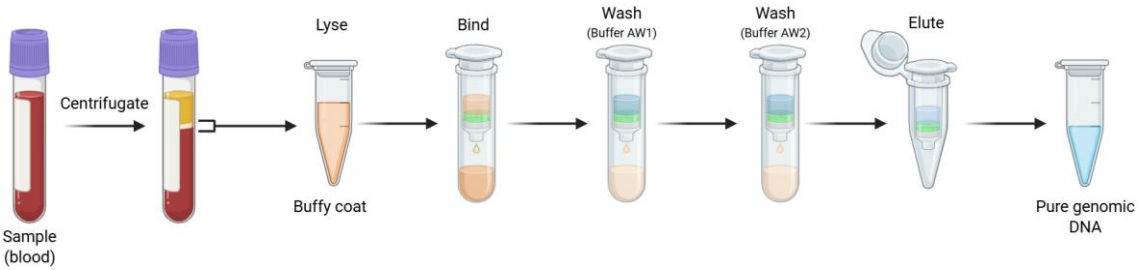


Figure 2. Schematic overview of the purification process of genomic DNA. Adapted from "DNA Purification from Blood or Body Fluids (Spin Protocol)", Qiagen, Barcelona, Spain. Created with *BioRender.com*.

DNA purification was performed automatically using the QIAcube® instrument (Qiagen, Barcelona, Spain), following the manufacturer's DNA Purification from Blood or Body Fluids: Spin Protocol (Qiagen, Barcelona, Spain), adapted for automated execution. This

procedure included an addition of buffer and ethanol to facilitate DNA binding to the silica membrane of the QIAamp Mini spin columns (Qiagen, Barcelona, Spain). The lysed sample was loaded onto the column, followed by successive wash steps to remove proteins, nucleases, and potential PCR inhibitors. Finally, purified DNA was eluted in 100 µL of buffer (≈30ng/µL) at room temperature after a 5-minute incubation on the membrane to optimize yield. The resulting DNA was free of contaminants and suitable for subsequent applications.

6.3. SNP detection in the *DPYD* gene using Real-Time PCR and Melting Curve Analysis

Simultaneous detection of the four clinically relevant variants was carried out using the commercial LightMix® in-vitro diagnostics kit Multi-SNiP *DPYD* (Roche®, Barcelona, Spain) and LightCycler® FastStart DNA Master HybProbe (Roche®, Barcelona, Spain), which allows identification of clinically relevant polymorphisms through melting curve analysis.

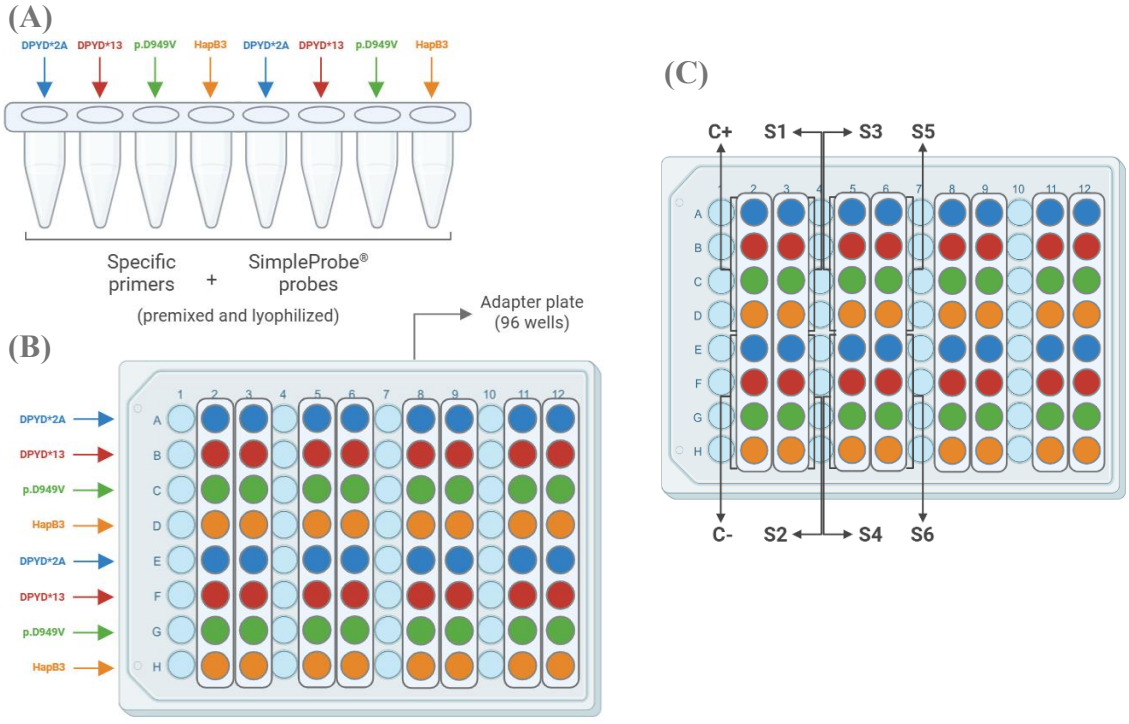


Figure 3. Schematic representation of: **(A)** the contents of each well and the location of the different variants; **(B)** the arrangement of the 8-well strips on the 96-well adapter plate; **(C)** the distribution of controls (C) and samples (S) on the plate. Adapted from "*LightMix® in-vitro diagnostics kit Multi-SNiP DPYD*", Roche®, Barcelona, Spain. Created with *BioRender.com*.

The LightMix® kit consists of 8-well strips, each containing a lyophilized mixture of specific primers and SimpleProbe® probes for a single reaction, along with a multiallelic

positive control for the four clinically relevant variants and the wild-type variant. Each strip allowed the analysis of two samples for the four variants (one target per well), as detailed in **Figure 3**. For each reaction, 2 µL of DNA sample were added per well. As each sample was distributed in 4 wells (one per variant), a total of 8 µL of DNA per sample was required for the complete analysis. Other reagents necessary for the PCR reaction, such as the MasterMix, H_2O, and magnesium (which facilitates DNA polymerase activity and stabilizes DNA hybrids during PCR), were provided by the LightCycler® kit.

PCR reactions were performed in the 8-well strips placed on a 96-well adapter plate, using the LightCycler® 480 Instrument II (Roche®, Barcelona, Spain). Each analytical run included a positive and negative control for each SNP, and samples were loaded into the wells according to the layout shown in **Figure 3**. The thermal cycling program consisted of four steps, as indicated in **Table 3**. The first two steps corresponded to the denaturation and amplification of genetic material, while the third step was used to obtain the melting curves.

Table 3: Thermal cycling program using the LightCycler® 480 Instrument II: 1) Sample denaturation and enzyme activation; 2) PCR amplification cycling of the target DNA; 3) Melting curve analysis of PCR-amplified DNA; 4) Instrument cooling. Adapted from "*LightMix® in-vitro diagnostics kit Multi-SNiP DPYD*", Roche®, Barcelona, Spain.

STEP	**1**	**2**			**3**			**4**
Analysis Mode	None	Quantification mode			Melting curves mode			None
Cycles	1	45			1			1
Target (°C)	95	95	60	72	95	43	75	40
Duration (s)	600	5	10	15	30	120	1	30
Ramp (°C/s)	4.4	4.4	2.2	4.4	4.4	1.5	0.20	1.5

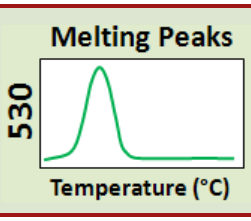
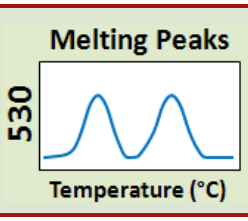
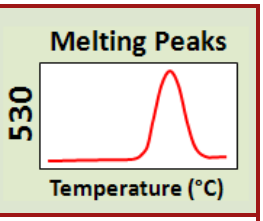
After PCR, a gradual temperature increase was performed to induce dissociation of the SimpleProbe® probes from the amplified DNA, resulting in a decrease in fluorescence that enabled the recording of melting curves with the LightCycler® 480 Instrument II. Melting curve generation and analysis were conducted using LightCycler® 480 SW 1.5.1 software (Roche®, Barcelona, Spain), monitoring fluorescence changes in relation to temperature variation and determining the characteristic Tm of each variant.

6.4. Melting Curve Analysis

Manual analysis of the melting curves was performed using the "Tm Calling" mode of the software, verifying that the positive control exhibited two melting peaks per SNP

(wild-type and clinically relevant alleles) and that the negative control showed no peaks. Each sample was interpreted by comparing its Tm values with the reference values indicated in **Table 4**, provided by the manufacturer for the equipment and reagents used, as well as with the Tm obtained from their respective positive controls.

Table 4: Typical analysis results. **Note:** The values of the melting temperatures may vary ±2.5°C between different runs. The ΔT between the melting peaks for heterozygous genotypes may vary ±1.5°C. Adapted from "*LightMix® in-vitro diagnostics kit Multi-SNiP DPYD*", Roche®, Barcelona, Spain.

For simplicity **WildType** **Heterozygous** **Mutant** are color coded					
			Left Peak	**Two peaks**	**Right Peak**
	1	*DPYD*2A*	GG 50ºC	Δ 7ºC	AA 57ºC
	2	*DPYD*13*	CC 58ºC	Δ 7ºC	AA 65ºC
	3	*p.D949V*	TT 60ºC	Δ 4ºC	AA 64ºC
	4	*HapB3*	CC 55ºC	Δ 8ºC	GG 64ºC
	5	*DPYD*2A*	GG 50ºC	Δ 7ºC	AA 57ºC
	6	*DPYD*13*	CC 58ºC	Δ 7ºC	AA 65ºC
	7	*p.D949V*	TT 60ºC	Δ 4ºC	AA 64ºC
	8	*HapB3*	CC 55ºC	Δ 8ºC	GG 64ºC

6.5. Validation of Results for Subsequent Treatment Administration

Once the genetic variants in the analyzed samples were detected, the results were recorded through the hospital's *Sistema Informático de Laboratorio (SIL),* Modulab (Werfen®, Barcelona, Spain), linked to the physician's request. This record specified the presence or absence of clinically relevant variants in the analyzed sample and identified the detected variants. Simultaneously, a new report entitled "*Estudio DPYD*" was created in the patient's electronic medical record. This section included the automatically generated results report produced by the system upon indicating the identified variant and the classification of the detected variants according to current clinical guidelines and recommendations, such as those published by the CPIC, the SNS, and the *SEFF*. Once the report was incorporated into the medical record and the results were validated, a notification was sent to the Hospital Pharmacy Department, which proceeded to adjust the pharmacological treatment and/or dosing according to the patient's genetic profile categorization.

7. RESULTS AND DISCUSSION

Note: In accordance with "*Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*" and to safeguard patient confidentiality, the results presented herein have been elaborated as recreations based on real cases, ensuring that no individual patient can be identified.

7.1. Melting curve analysis

7.1.1. Sample 1

The melting curves obtained for the four SNPs of the *DPYD* gene for Sample 1 are shown in **Figure 4**.

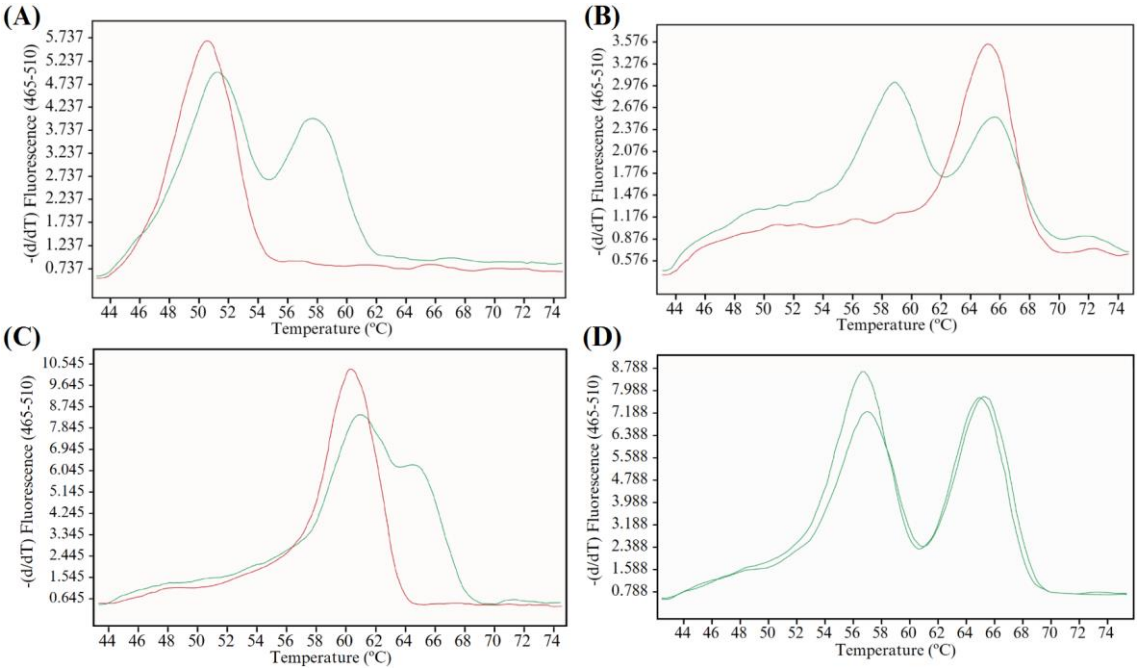


Figure 4. Melting curves obtained for the different variants in the genotyping of Sample 1: (A) *DPYD*2A*; (B) *DPYD*13*; (C) *p.D949V*; (D) *HapB3*, relative to the positive control. The X-axis represents temperature (°C) and the Y-axis shows the negative derivative of fluorescence (between 465 and 510 nm). Data obtained from LightCycler® 480 SW 1.5.1.

In **Figures 4A**, **4B**, and **4C**, the two green melting peaks corresponding to the positive control were observed, with Tm values of 51°C and 58°C, 59°C and 66°C, and 61°C and 65°C, respectively. In **Figure 4D**, the two positive control peaks presented the lowest fluorescence intensities, with Tm values of 57°C and 66°C. These values are consistent with those indicated in **Table 4**, in accordance with the specifications of the LightMix® in-vitro diagnostics kit Multi-SNiP *DPYD*.

Subsequently, the melting peaks of Sample 1 were analyzed, comparing their Tm values with those indicated in **Table 4**:

- In **Figure 4A**, a single peak at 51.0°C was observed, corresponding to a wild-type genotype for *DPYD*2A*.
- In **Figure 4B**, a single peak at 66°C was recorded, indicating a wild-type genotype for *DPYD*13*.
- In **Figure 4C**, a single peak at 60°C was detected, likewise compatible with a wild-type genotype for *p.D949V*.
- In **Figure 4D**, two peaks (those presenting the highest fluorescence) at 57°C and 66°C were identified, evidencing the presence of a wild-type allele and a *HapB3* allele, thus confirming a heterozygous genotype for this SNP.

Consequently, it was concluded that Sample 1 exhibited a heterozygous **1/HapB3* genotype, associated with the Intermediate Metabolizer phenotype, as indicated in **Table 1**. According to the literature, as *HapB3* is a reduced-function allele, this genotype could result in decreased DPD enzymatic activity, thereby increasing the risk of toxicity with FL-based treatments. For this reason, a 50% reduction in the initial dose of FLs, together with close clinical monitoring during treatment, particularly throughout the first two cycles, was recommended to assess the need for additional dose adjustments depending on the occurrence of adverse reactions.

7.1.2. Sample 2

The melting curves corresponding to the four SNPs of Sample 2 are shown in **Figure 5**.

In **Figures 5B**, **5C**, and **5D**, the two green melting peaks corresponding to the positive control were observed, with Tm values of 58°C and 65°C, 60°C and 64°C, and 56°C and 64°C, respectively. In **Figure 5A**, the two positive control peaks presented the lowest fluorescence intensities, with Tm values of 50°C and 57°C. All these values coincided with the theoretical values specified in **Table 4**.

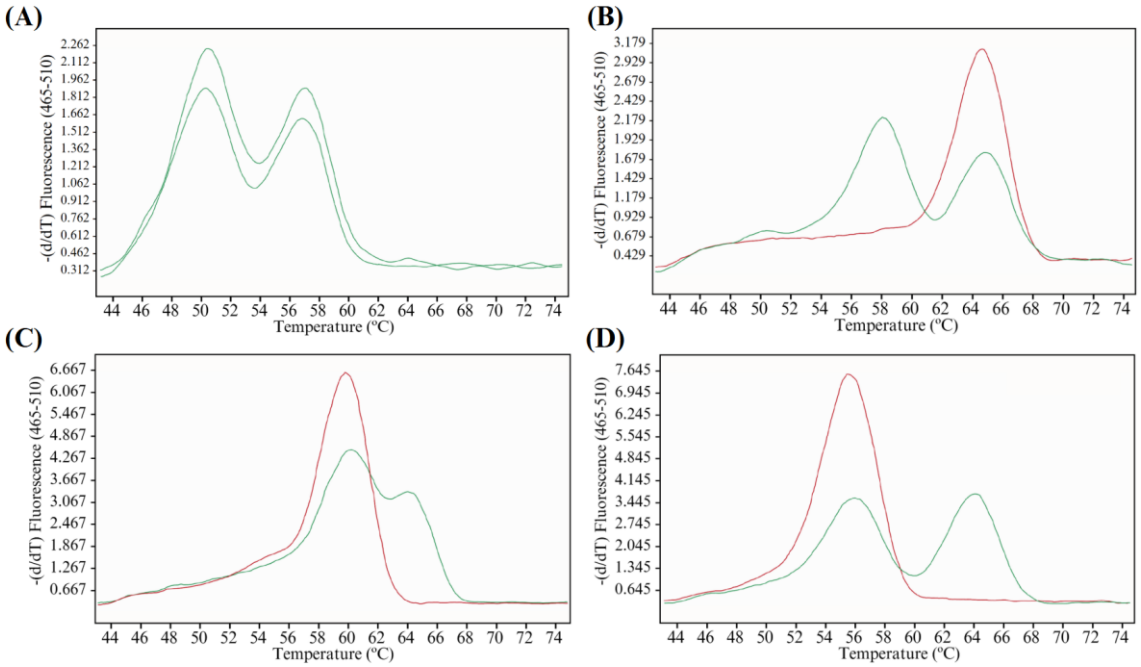


Figure 5. Melting curves obtained for the different variants in the genotyping of Sample 2: (A) *DPYD**2A; (B) *DPYD**13; (C) *p.D949V*; (D) *HapB3*, relative to the positive control. The X-axis represents temperature (°C) and the Y-axis shows the negative derivative of fluorescence (between 465 and 510 nm). Data obtained from LightCycler® 480 SW 1.5.1.

The analysis of the melting peaks of Sample 2 revealed:

- In **Figure 5A**, two peaks at 50°C and 57°C, indicating a heterozygous genotype for *DPYD*2A*.
- In **Figure 5B**, a single peak at 65°C, corresponding to a wild-type genotype for *DPYD*13*.
- In **Figure 5C**, a single peak at 60°C, corresponding to a wild-type genotype for *p.D949V*.
- In **Figure 5D**, a single peak at 56°C, indicating a wild-type genotype for *HapB3*.

Thus, it was concluded that Sample 2 exhibited a heterozygous **1/*2A* genotype, also associated with the Intermediate Metabolizer phenotype. As reported in the literature, since *DPYD*2A* is a loss-of-function allele, this genotype (like **1/HapB3*) entails a reduction in DPD activity, increasing the risk of FL-related toxicity. Consequently, a 50% reduction in the initial dose and close clinical monitoring during the initial phases of treatment were likewise recommended.

7.1.3. Sample 3

The melting curves obtained for the four SNPs of Sample 3 are shown in **Figure 6**.

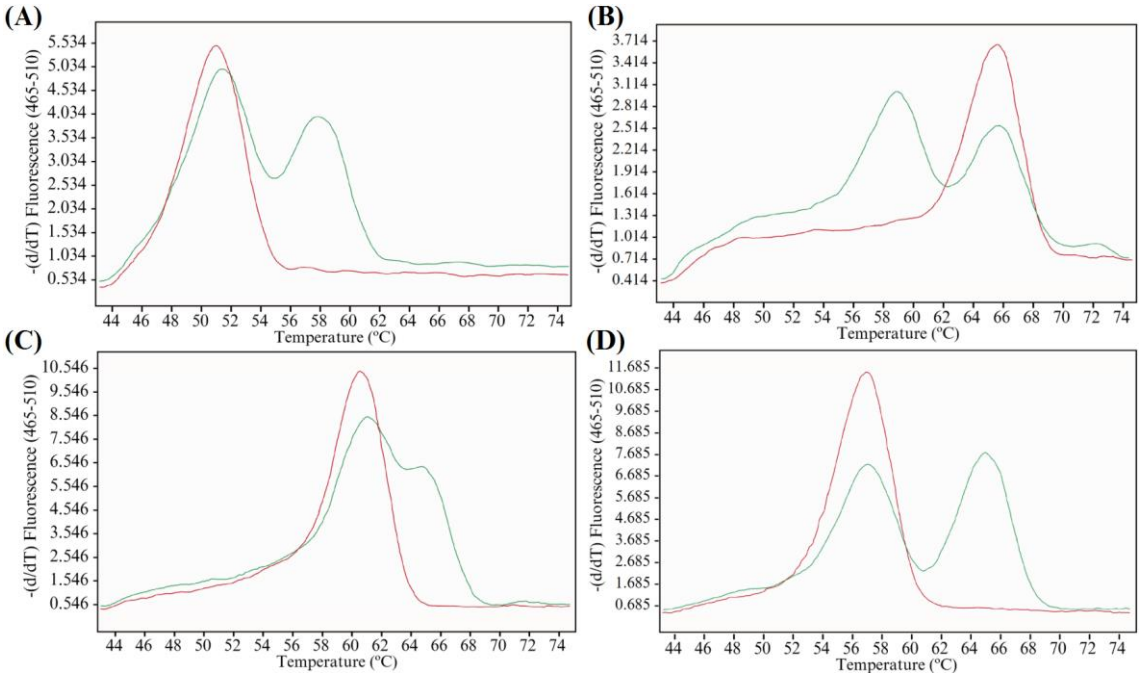


Figure 6. Melting curves obtained for the different variants in the genotyping of Sample 3: (A) *DPYD*2A*; (B) *DPYD*13*; (C) *p.D949V*; (D) *HapB3*, relative to the positive control. The X-axis represents temperature (°C) and the Y-axis shows the negative derivative of fluorescence (between 465 and 510 nm). Data obtained from LightCycler® 480 SW 1.5.1.

In **Figures 6A**, **6B**, **6C**, and **6D**, the two green melting peaks corresponding to the positive control were identified, with Tm values of 51°C and 58°C, 59°C and 66°C, 61°C and 65°C, and 57°C and 65°C, respectively. These values coincided with the theoretical values provided in **Table 4**.

The analysis of the melting peaks of Sample 3 revealed:

- In **Figure 6A**, a single peak at 51°C, corresponding to a wild-type genotype for *DPYD*2A*.
- In **Figure 6B**, a single peak at 66°C, indicating a wild-type genotype for *DPYD*13*.
- In **Figure 6C**, a single peak at 61°C, compatible with a wild-type genotype for *p.D949V*.
- In **Figure 6D**, a single peak at 57°C, corresponding to a wild-type genotype for *HapB3*.

Accordingly, it was concluded that Sample 3 exhibited a homozygous wild-type genotype for all genotyped SNPs. This corresponds to the Normal Metabolizer phenotype, as stated in **Table 1**. Based on the literature consulted, as this individual harbors normal-function alleles for all genotyped variants, normal DPD enzymatic activity and a standard risk of FL-related toxicity would be expected. Therefore, the recommended dose is the standard dose specified in the product information for the drug. Nevertheless, as in previous cases, close clinical monitoring of the patient would be advisable to assess the need for additional dose adjustments according to the occurrence of potential adverse reactions, since the absence of clinically relevant SNPs in this panel does not rule out the possibility of toxicity. Other variants not included in this genotyping assay may still lead to adverse reactions during FL-based treatment.

7.2. Forensic applicability

The results obtained in this study confirmed that the real-time PCR technique combined with melting curve analysis is a reliable, rapid, and cost-effective tool for the detection of specific genetic variants, as demonstrated with the selected SNPs within the *DPYD* gene. This approach could also be applied to the forensic field if primers targeting SNPs associated with EVCs were developed. To achieve a prediction of the best-studied phenotypic traits, such as eye, hair, and skin color, it would be necessary to design primers for, for example, the forty-one SNPs included in the HIrisPlex-S system. However, its extrapolation to the forensic field, particularly to FDP, presents a series of technical, operational, and economic limitations and considerations that require careful evaluation prior to potential implementation.

Firstly, one of the main inherent limitations for its forensic applicability was the quantity and quality of DNA typically available from crime scene samples. Most forensic samples displayed concentrations below 5 ng/µL, the minimum threshold established to ensure reliable results with this methodology. Additionally, the fragmented and degraded nature of forensic DNA, together with the possible presence of inhibitors, often compromised both the amplification efficiency and the reliability of the melting profiles obtained, frequently generating irregular peaks or atypical Tm values that hindered correct allelic assignment. This challenge was compounded by the forensic prioritization of STR profiling, both by regulation and its higher individualizing power. Consequently, in cases of limited DNA quantity, a substantial portion of the extracted material was allocated for

STR analysis, leaving a reduced amount available for SNP detection. Given that this technique requires at least 5 ng/µL of DNA per reaction, and considering, for instance, the forty-one SNPs included in the HIrisPlex-S system (currently the forensic standard for predicting some EVCs), the total DNA volume needed far exceeded what is typically recovered from most forensic samples, particularly in trace DNA scenarios. Another significant limitation was the difficulty in detecting and interpreting DNA mixtures. Unlike STRs, for which mixture interpretation is well-standardized, the superimposed melting curves generated by SNP profiles from multiple contributors proved virtually impossible to deconvolute, restricting the technique's application to single-source samples or those where differential DNA separation could be achieved beforehand (e.g., sexual assault cases with sufficient semen for isolating the perpetrator's DNA).

From an economic perspective, the analysis of multiple SNPs using real-time PCR and melting curve analysis resulted in a considerably higher cost per SNP than that of NGS systems when a comprehensive phenotypic prediction was intended. For example, in the protocol developed for the *DPYD* gene, the cost of sixty-four reactions (four SNPs across sixteen samples) approached €1000, which, when extrapolated to the forty-one SNPs of HIrisPlex-S, would have proved excessive. This issue could have been partially mitigated by modular dissection of reaction plates, grouping SNPs into trait-specific plates (for instance, one plate for the six SNPs related to eye color, another for the eighteen SNPs for hair, and another for the seventeen SNPs associated with skin color), thereby reducing the number of required positive and negative controls per plate and optimizing resource utilization. Moreover, this technique inherently requires a specific positive and negative control for each SNP on every plate, increasing operational costs and limiting available space for sample testing within each run. However, this drawback could be addressed through the development of thematic plates with lyophilized primers and probes, which also promotes improved conservation, standardization, and availability for accredited forensic laboratories. From an operational standpoint, another disadvantage was that allelic assignment based on melting curve profiles required manual visual review by qualified personnel, as atypical profiles, intermediate peaks, or overlapping signals could compromise data interpretation. While feasible, this review process introduced a subjective component and demanded prior expertise, thereby limiting the full automation of the workflow. Additionally, it was necessary to consider human population genetic variability, which required prior validation of any SNP panel in the reference population

where it would be applied. Otherwise, population-specific biases might have been introduced, potentially affecting the predictive reliability of inferred phenotypes. For example, although the HIrisPlex-S system has been validated primarily in European populations, its applicability in other population groups might be reduced or require the incorporation of additional SNPs. Regarding the predictive power of selected SNPs, it is important to note that the number of variants currently associated with comprehensive phenotypic prediction remains limited, and not all traits can be predicted with sufficient reliability to hold forensic value. Furthermore, many of these phenotypic traits may be influenced by environmental, hormonal, or aging-related factors, which introduces uncertainty into the evidentiary value of inferred phenotypes.

Despite these limitations, the technique presented advantages that could justify its use in specific scenarios. Notably, it offered rapid execution, with preliminary results obtainable in approximately less than two hours, without requiring sequencing, thereby reducing processing times and costs in certain investigations. Moreover, its capability to work with short DNA fragments (amplicons typically below 170 bp) favored its application in degraded samples, a frequent condition in forensic investigations. Another potential advantage lay in the feasibility of developing commercial kits with lyophilized primers and reagents, specifically optimized for forensic environments. This would simplify its implementation in accredited forensic laboratories and enable rapid response in criminal investigations. Furthermore, while the use of different fluorophores theoretically permits the multiplexing of multiple SNPs per well, this option significantly increases both reagent and instrumentation costs and presents limitations in terms of the number of distinguishable fluorophores within a single assay. Finally, although other methodologies such as NGS or microarrays offer more comprehensive SNP detection, they also present drawbacks, including high costs, the requirement for specialized equipment and personnel, and extended processing times. In this context, real-time PCR combined with melting curve analysis positioned itself as a complementary tool, with potential applicability in specific cases such as sexual assaults involving abundant biological material or dental remains. Nonetheless, its broader implementation for FDP would have required comprehensive forensic validation, the design of population-specific panels, and the development of standardized operating protocols ensuring the reliability, reproducibility, and traceability of the results obtained.

8. CONCLUSION

This study demonstrated that real-time PCR combined with melting curve analysis constitutes a reliable, rapid, and reproducible technique for the detection of specific SNPs in clinical settings, as evidenced by its routine use for *DPYD* variant genotyping in FL pharmacogenetics. The method proved highly effective in identifying the four clinically relevant *DPYD* variants, enabling accurate genotype determination and supporting individualized therapeutic recommendations aimed at reducing the risk of severe toxicity in oncologic patients.

The evaluation of this technique's potential forensic applicability for FDP revealed several limitations that currently restrict its use in routine forensic practice. Among the primary challenges identified were the limited quantity and quality of DNA typically available from forensic samples, the need for specific positive and negative controls for each SNP in every reaction plate, and the difficulty of interpreting DNA mixtures using melting curve profiles. Additionally, the comparatively high per-SNP cost of real-time PCR assays rendered the technique less economically viable than other methodologies, such as NGS, particularly when comprehensive phenotypic prediction panels involving large numbers of SNPs were required. The analysis also emphasized the importance of validating SNP panels in specific population groups to avoid population stratification bias and ensure the reliability of phenotypic predictions. Furthermore, it was noted that many phenotypic traits remain genetically complex and insufficiently characterized, limiting the forensic predictive power of currently available SNP markers. Environmental, hormonal, and age-related factors further complicate the interpretation of inferred phenotypes. Nonetheless, several operational advantages of the technique were identified, including its short processing time, suitability for degraded DNA through the use of short amplicons, and the potential for commercial development of lyophilized reaction kits optimized for forensic laboratories. These features may justify its selective application in specific forensic scenarios, such as cases involving abundant or well-preserved biological material, or when rapid preliminary results are required.

In summary, while real-time PCR with melting curve analysis demonstrated clear clinical utility and technical feasibility for SNP genotyping, its extrapolation to forensic FDP remains limited. Nevertheless, the application of SNP detection techniques for FDP represented a complementary tool with potential value in cases where already established

and validated methods, such as STR analysis, were inconclusive. Furthermore, advances in artificial intelligence currently offer significant opportunities to enhance the practical utility of SNP-based FDP, particularly through the generation of predictive composite sketches. For all of these reasons, the investigation of novel SNPs associated with EVCs should be increasingly promoted.

9. ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to the Institut Català de la Salut Camp de Tarragona, and especially to Hospital Universitari Joan XXIII, for giving me the opportunity to carry out my master's internship at their facilities. I am particularly grateful to the Molecular Biology Department, where I was fortunate to complete this formative stage.

My deepest thanks go to Cristina Gutiérrez Fornés and Clara Benavent Bofill for their dedication, patience, and constant willingness to guide, teach, and support me throughout my stay. Their knowledge and encouragement have been essential to both my learning and the completion of this project. I would also like to thank all the Consultants, Residents, and Technicians of the department, who have generously offered their help and provided the resources necessary to adapt a clinical procedure to a forensic-oriented approach.

I am especially grateful to Ximena Terra Barbadora for her valuable guidance, suggestions, and continuous support during the writing of this thesis.

Finally, I would like to thank my family for their unconditional support throughout these years. They have been, and continue to be, a fundamental pillar in my life, and their trust and affection have been key to reaching each of the goals I have set during this stage.

10. REFERENCES

1. Wu, K. *et al.* Recent Progress in Single-Nucleotide Polymorphism Biosensors. *Biosensors (Basel)* 13, 864 (2023).
2. Single Nucleotide Polymorphisms (SNPs). https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms-SNPs.
3. Antontseva, E. V., Degtyareva, A. O., Korbolina, E. E., Damarov, I. S. & Merkulova, T. I. Human-genome single nucleotide polymorphisms affecting transcription factor binding and their role in pathogenesis. *Vavilov Journal of Genetics and Breeding* 27, 662–675 (2023).
4. Durland, J. & Ahmadian-Moghadam, H. *Genetics, Mutagenesis*. (2025).
5. Fondevila, M., Phillips, C. & Lareu, M. V. Aplicaciones no convencionales de marcadores bialélicos (SNPS) en genética forense. *Boletín Galego de Medicina Legal e Forense* 20, (2014).
6. NCI Dictionary of Cancer Terms. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/pharmacogenetics.
7. Arturo Prior-González, O. *et al.* Farmacogenética y su importancia clínica: hacia una terapia personalizada segura y eficiente. *Medicina Universitaria* 13, 41–49 (2010).
8. Definition of pharmacogenetics - NCI Dictionary of Cancer Terms - NCI. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/pharmacogenetics.
9. Pharmacogenetics : MedlinePlus Medical Encyclopedia. https://medlineplus.gov/ency/article/007794.htm.
10. 'Fluoropirimidinas' Sociedad Española de Farmacogenética y Farmacogenómica (SEFF). Acceso: 8 Abril 2025. Disponible: https://seff.es/fluoropirimidinas/.
11. De Mattia, E., Milan, N., Assaraf, Y. G., Toffoli, G. & Cecchin, E. Clinical Implementation of Rare and Novel DPYD Variants for Personalizing Fluoropyrimidine Treatment: Challenges and Opportunities. *Int J Biol Sci* 20, 3742–3759 (2024).
12. Sharma, B. B. *et al.* Pathogenic DPYD Variants and Treatment-Related Mortality in Patients Receiving Fluoropyrimidine Chemotherapy: A Systematic Review and Meta-Analysis. *Oncologist* 26, 1008–1016 (2021).
13. Pratt, V. M. *et al.* DPYD Genotyping Recommendations. *The Journal of Molecular Diagnostics* 26, 851–863 (2024).

14. Zubiaur Precioso, P. & Abad Santos, F. *Manual de Aplicación de La Farmacogenética Clínica*.

15. Amstutz, U. *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Dihydropyrimidine Dehydrogenase Genotype and Fluoropyrimidine Dosing: 2017 Update. *Clin Pharmacol Ther* 103, 210–216 (2018).

16. Loganayagam, A. *et al.* Pharmacogenetic variants in the DPYD, TYMS, CDA and MTHFR genes are clinically significant predictors of fluoropyrimidine toxicity. *Br J Cancer* 108, 2505–2515 (2013).

17. 'Fluorouracilo, capecitabina, tegafur y flucitosina en pacientes con déficit de dihidropirimidina deshidrogenasa.' Agencia Española de Medicamentos y Productos Sanitarios (AEMPS). Acceso: 4 Abril 2025. Disponible: https://www.aemps.gob.es/informa/fluorouracilo-capecitabina-tegafur-y-flucitosina-en-pacientes-con-deficit-de-dihidropirimidina-deshidrogenasa/.

18. Aprobación del acuerdo sobre el Catálogo de pruebas genéticas de la Cartera común de servicios del Sistema Nacional de Salud. https://www.sanidad.gob.es/organizacion/consejoInterterri/docs/1553.pdf.

19. Henricks, L. M., Opdam, F. L., Beijnen, J. H., Cats, A. & Schellens, J. H. M. DPYD genotype-guided dose individualization to improve patient safety of fluoropyrimidine therapy: call for a drug label update. *Annals of Oncology* 28, 2915–2922 (2017).

20. Sguazzi, G. *et al.* 'Biological identikit': Development of a SNPs-panel for the analysis of forensic DNA phenotyping and ancestry. *Forensic Sci Int Genet Suppl Ser* 8, 199–201 (2022).

21. Marano, L. A. & Fridman, C. DNA phenotyping: current application in forensic science. *Research and Reports in Forensic Medical Science* Volume 9, 1–8 (2019).

22. Schneider, P. M., Prainsack, B. & Kayser, M. The Use of Forensic DNA Phenotyping in Predicting Appearance and Biogeographic Ancestry. *Dtsch Arztebl Int* (2019) doi:10.3238/arztebl.2019.0873.

23. CODIS and NDIS Fact Sheet — FBI. https://www.fbi.gov/how-we-can-help-you/dna-fingerprint-act-of-2005-expungement-policy/codis-and-ndis-fact-sheet.

24. Canales Serrano, A. Forensic DNA phenotyping: A promising tool to aid forensic investigation. Current situation. *Spanish Journal of Legal Medicine* 46, 183–190 (2020).

25. Kayser, M. & Schneider, P. M. DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. *Forensic Sci Int Genet* 3, 154–161 (2009).

26. Kayser, M. Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Sci Int Genet* 18, 33–48 (2015).

27. Dabas, P., Jain, S., Khajuria, H. & Nayak, B. P. Forensic DNA phenotyping: Inferring phenotypic traits from crime scene DNA. *J Forensic Leg Med* 88, 102351 (2022).

28. Hunter, P. Cold cases and ancient trade routes. *EMBO Rep* 22, (2021).

29. HIrisPlex-S Eye, Hair and Skin Colour DNA Phenotyping Webtool. https://hirisplex.erasmusmc.nl/.

30. Ketsekioulafis, I. *et al.* Artificial Intelligence in Forensic Sciences: A Systematic Review of Past and Current Applications and Future Perspectives. *Cureus* (2024) doi:10.7759/cureus.70363.

31. El IIS Aragón participa en un proyecto para crear retratos robot a partir de pruebas de ADN | ConSalud. https://www.consalud.es/autonomias/aragon/iis-aragon-proyecto-crear-retratos-robot-pruebas-adn_126185_102.html.

32. Lawrie, R. D. & Massey, S. E. Agrigenomic Diversity Unleashed: Current Single Nucleotide Polymorphism Genotyping Methods for the Agricultural Sciences. *Applied Biosciences* 2, 565–585 (2023).

33. Some corrections and translations from Spanish to English for this work were performed using the ChatGPT language model. OpenAI, 2025. [Online]. Available: https://www.openai.com/chatgpt