

Tese de Mestrado

Cláudia Alexandra Moreira de Carvalho
Licenciatura em Biologia

Aplicação Médico-Legal da PCR em Tempo Real na Caracterização de SNPs

Orientadora: Prof. Doutora Maria de Fátima Terra Pinheiro

Tese de mestrado apresentada à Faculdade de Medicina da Universidade do
Porto, para obtenção do título de Mestre em Ciências Forenses.

Porto
Abril 2007



Se soubermos que um obstáculo é intransponível, deixa de ser um obstáculo para se tornar um ponto de partida.
(József Eotvos)

Agradecimentos

Gostaria de dirigir as primeiras palavras deste trabalho a todos os que, durante a sua realização, me colocaram desafios e disponibilizaram diversas formas de apoio, pelo que a todos expresso o meu reconhecido agradecimento. Sem a sua ajuda, este trabalho não teria sido possível.

À Comissão Coordenadora do Mestrado em Ciências Forenses, em especial à Prof. Doutora Teresa Magalhães, por me ter permitido e incentivado à frequência neste Mestrado.

À Prof. Doutora Maria de Fátima Pinheiro, pelo empenho, incentivo, pelo tempo partilhado e pela exigência constante com que orientou este estudo, reconhecendo o privilégio de ter sido sua orientanda.

À Mestre Lurdes Rebelo e ao Prof. Doutor José Manuel Cabeda, por me terem sempre recebido com toda a simpatia e pelos seus ensinamentos nas áreas da Genética Forense e da Biologia Molecular. Sem eles esta tese, sem dúvida, não seria possível. Um muito obrigado!

À Prof. Doutora Berta Martins e ao Prof. Doutor Paulo Pinho e Costa pela compreensão no tempo despendido e pela ajuda.

A todas as pessoas do Serviço de Genética e Biologia Forense da Delegação do Porto do Instituto Nacional de Medicina Legal, pela simpatia e disponibilidade, com que acolheram e participaram neste trabalho.

A todos os amigos e colegas que me ajudaram e apoiaram desde o início deste trabalho. E em especial as minhas colegas de laboratório por me aturarem tantas e tantas vezes!

À minha mãe, pai, irmã e avó, pela sua paciência e apoio em todos os momentos.

Resumo

A região não recombinante do cromossoma Y possui vários polimorfismos com interesse médico-legal, entre os quais os SNPs (*Single Nucleotide Polymorphisms*). A aplicação do seu estudo à Genética e Biologia Forense exige a utilização de técnicas suficientemente específicas, rápidas e reprodutíveis que permitam a análise de um elevado número de amostras ou de diferentes polimorfismos em simultâneo.

O objectivo principal deste trabalho consistiu na validação de protocolos distintos, usados na tipagem de SNPs, utilizando a técnica de PCR em Tempo Real.

Foram analisados 33 indivíduos não aparentados intervenientes em casos de investigação de filiação. Após a extracção de DNA, recorrendo a dois métodos: a extracção orgânica com fenol-clorofórmio-álcool isoamílico e a extracção com Chelex, procedeu-se à caracterização do M2, M170, M172, M9, M45, M207, M173 e P25, utilizando a PCR em Tempo Real, com sondas específicas TaqMan[®] em três aparelhos: o ABI PRISM 7000 (*Applied Biosystems*), o LightCycler[®] (*Roche*) e o SmartCycler (*Cepheid*). Com excepção dos polimorfismos M2 e M9, os SNPs estudados foram seleccionados de acordo com uma estratégia hierárquica baseada na árvore filogenética do cromossoma Y, proposta pelo YCC (*Y Chromosome Consortium*).

A amplificação e detecção dos polimorfismos demonstraram ser exequíveis nas 3 plataformas utilizadas. No entanto, foi necessário proceder a alguns ajustes na composição do tampão de amplificação no Lightcycler[®], incorporando albumina, para obviar a adsorção das amostras a caracterizar nos capilares de vidro. A amplificação do polimorfismo P25, também se mostrou algo problemática, devido à sua localização numa região genómica de repetição.

O trabalho efectuado permite concluir que a PCR em Tempo Real é um bom método para a caracterização de SNPs com aplicação forense. A sua elevada rapidez, especificidade, reprodutibilidade e poder de discriminação suplantam de longe a dificuldade no estabelecimento de técnicas *multiplex*. Acresce que a elevada rapidez e capacidade de se processar em paralelo um elevado número de amostras, bem como de executar em simultâneo ensaios distintos, permite ultrapassar a aparente desvantagem da impossibilidade de se analisar concomitantemente diversos polimorfismos.

Abstract

The non recombinant region of Y chromosome (NRY) is very rich in polymorphisms of forensic importance, among them the SNPs (Single Nucleotide Polymorphisms). The study of those polymorphisms is of great interest in Genetics and Forensic Biology and requires specific, fast and reproducible techniques that enable the simultaneous analysis of numerous samples or different polymorphisms.

The main goal of this thesis was the validation of a Real Time PCR methodology for the genotyping of Y chromosomal SNPs. For that purpose 33 DNA samples for paternity testing were studied. The DNA extraction was performed with two methodologies: Phenol/ chloroform and Chelex extraction. The SNPs M2, M170, M172, M9, M45, M207, M173 e P25 were genotyped with specific TaqMan[®] probes using three different Real Time PCR instruments: ABI PRISM 7000 (Applied Biosystems), LightCycler[®] (Roche) and SmartCycler (Cepheid). The selection of all SNPs, with the exception of M2 and M9, was in accordance with the hierarchy of the chromosome's Y phylogenetic tree, proposed by YCC (Y Chromosome Consortium).

Amplification and detection of all SNPs was possible in the three platforms used albeit some adjustments had to be made. In Lightcycler[®], albumin was added to the amplification buffer, to prevent the sample absorption by the glass capillaries. The localization of P25 polymorphism in a repetitive genomic region, has originated some problems in the amplification reaction.

In conclusion, we can say that Real Time PCR is a good methodology to the characterization of SNPs with forensic application. Its higher speediness, specificity and discrimination power, go beyond the complexity of multiplex techniques. Moreover the impossibility of analyse concomitantly different polymorphism is surpassed by its capacity to process, simultaneously, numerous samples and distinct assays.

Resumé

La région non recombinante du chromosome Y possède plusieurs polymorphismes ayant un intérêt médico-légal, parmi lesquels le SNPs (Single Nucleotide Polymorphisms). L'application de son étude à la génétique et à la biologie forensique exige l'utilisation de techniques qui soit suffisamment spécifiques, rapides et reproductibles pour permettre l'analyse d'un grand nombre d'échantillons ou de différents polymorphismes simultanément. L'objectif principal de ce travail consiste à valider les différents protocoles utilisés dans le typage de SNPs, utilisant la technique de PCR en temps réel.

Ont été analysées, 33 personnes non apparentées, intervenantes dans des cas de recherche de filiation. Après l'extraction de l'ADN, faisant appel à deux méthodes : extraction organique avec phénol-chloroforme-alcool isoamylique et extraction avec Chelex. Nous avons procédé à la caractérisation du M2, M170, M172, M9, M45, M207, M173 et P25, en utilisant le PCR en temps réel avec des sondes spécifiques TaqMan® à l'aide de trois appareils : l'ABI PRISM 7000 (Applied Biosystems), le LightCycler® (Roche) et le SmartCycler (Cepheid). À l'exception des polymorphismes M2 et M9, les SNPs étudiés ont été sélectionnés conformément à une stratégie hiérarchique basée sur l'arbre phylogénétique du chromosome Y, proposé par YCC (Y Chromosome Consortium).

L'amplification et la détection des polymorphismes a démontré être réalisable dans les trois cas d'études utilisés. Néanmoins, quelques ajustements ont été nécessaires pour procéder à la composition du tampon d'amplification dans le Lightcycler®, en incorporant de l'albumine pour remédier à l'adsorption des échantillons à caractériser dans les capillaires de verre. L'amplification du polymorphisme P25, a aussi mis en évidence quelques problématiques, dues à sa localisation dans une région génomique de répétition.

Le travail effectué permet donc de conclure que le PCR en temps réel est une bonne méthode pour la caractérisation de SNPs avec application forensique. Sa grande rapidité, sa spécificité, sa reproductibilité et son pouvoir de discrimination supplantent de loin la difficulté dans l'établissement de techniques multiplex. De plus la grande rapidité et capacité de traiter en parallèle un nombre élevé d'échantillons, ainsi que d'exécuter en simultanée des essais distincts, ceci permet de supprimer l'inconvénient de l'impossibilité d'analyser de façon concomitante divers polymorphismes.

Índice

RESUMO.....	4
I - INTRODUÇÃO.....	11
1.1– O DNA E SUA APLICAÇÃO FORENSE.....	12
1.2 – O CROMOSSOMA Y	15
1.3 - MARCADORES BIALÉLICOS	18
1.4 - METODOLOGIAS PARA GENOTIPAGEM DE SNPs.....	22
1.4.1 - Mini-sequenciação	23
1.4.1.1 - <i>SnaPshot</i>	24
1.4.1.2 - <i>MALDI-TOF MS</i>	25
1.4.1.3 - <i>Microarrays</i>	26
1.4.2 - PCR em Tempo Real	27
1.4.2.1 - <i>Metodologia com sondas FRET</i>	28
1.4.2.2 - <i>Metodologia com sondas TaqMan®</i>	31
1.5 - SELECÇÃO DOS SNPs A ESTUDAR.....	34
II – OBJECTIVOS.....	37
III – MATERIAL E MÉTODOS.....	39
3.1 – AMOSTRAS	40
3.2 – EXTRACÇÃO DO DNA	40
3.2.1 – Extracção com fenol-clorofórmio-álcool isoamílico	40
3.2.2 – Extracção com Chelex.....	43
3.3 – CARACTERIZAÇÃO DOS SNPs	44
3.3.1 – Selecção dos SNPs.....	44
3.3.2 – Amplificação por PCR em Tempo Real	44
IV – RESULTADOS.....	47
4.1 - ABI PRISM 7000.....	48
4.2 – SMARTCYCLER	50
4.3 - LIGHTCYCLER®	51
4.4 - ANÁLISE DO SNP P25.....	52
4.5 - RESULTADOS OBTIDOS UTILIZANDO AS TRÊS PLATAFORMAS.....	54
4.6 - ANÁLISE DE SENSIBILIDADE DA METODOLOGIA	55
4.7 - APROXIMAÇÃO A ESTUDO POPULACIONAL	57
V - DISCUSSÃO.....	60
VI - CONCLUSÃO.....	69
VII – REFERÊNCIAS BIBLIOGRÁFICAS.....	73

I - Introdução

1.1– O DNA e sua aplicação forense

Em 1985, em Inglaterra, *Jeffreys et al.* [14] referem pela primeira vez que a análise do DNA constitui uma possibilidade de estudar a diversidade humana, descrevendo um método de identificação individual que designaram de “*DNA fingerprinting*”. Este método foi tido como a solução definitiva para o estabelecimento da identificação individual, tanto na investigação biológica de parentesco, como em Criminalística.

Num primeiro momento, pensou-se que a aplicação desta técnica à prática forense iria sofrer algum atraso devido a problemas estritamente legais, mas a verdade é que em Outubro de 1986 se assiste, pela primeira vez, à sua utilização na resolução de processos penais.

Nos anos seguintes levantaram-se alguns problemas em relação a estes procedimentos, principalmente nos EUA, tendo, alguns geneticistas reconhecidos internacionalmente colocado em questão situações como a sua validade científica, a sua aplicação a casos judiciais e, mais tarde, aspectos estatísticos. Em 1992 todas estas questões se encontravam resolvidas e a identificação individual, baseada na análise de DNA, foi considerada cientificamente válida, sempre e quando fosse utilizada a tecnologia apropriada e se mantivesse um rigoroso controlo de qualidade, sendo elaboradas guias de procedimentos, baterias de testes e aconselhados procedimentos analíticos.

Os fundamentos da técnicas de análise do DNA em Genética Forense, baseiam-se em dois princípios elementares deduzidos do modelo de dupla hélice proposto, em 1953, por *Watson e Crick*, [38] para a molécula de DNA:

- ✓ Princípio da complementaridade entre as bases;
- ✓ Desnaturação e renaturação do DNA.

Esta molécula, com estas características particulares, localiza-se, na sua grande maioria, no núcleo celular, organizada no designado complemento cromossómico humano. Este é composto por 46 elementos, organizados em 23 pares, 22 dos quais são semelhantes em homens e mulheres, os cromossomas autossómicos. O par restante, os cromossomas sexuais, apresentam características diferentes dependendo do sexo.

O estudo do cromossoma Y reveste-se de extrema importância em casos criminais, uma vez que a grande maioria dos agressores são homens, desde agressões sexuais, homicídios, frutos e outros, tendo também um importante papel em situações de investigação biológica de parentesco, quando o filho é do sexo masculino.

O cromossoma Y é um cromossoma acrocêntrico e pequeno, representando apenas 2% do genoma. Devido ao facto de, no Homem, ser muito diferente do cromossoma X, a sua maior parte não recombina durante a meiose. A recombinação com o cromossoma X só se processa nas duas pequenas regiões pseudoautossómicas denominadas PAR1 e PAR2. A ausência de recombinação determina que todas as sequências, existentes na zona que não as duas regiões mencionadas, se herdem em bloco de pais para filhos. Assim, as mutações constituem as únicas fontes possíveis de variação nesta região.

O estudo destas variações possibilita:

- ✓ A identificação de autores de delitos em casos criminais;
- ✓ A determinação de vínculos de parentesco na ausência do progenitor;
- ✓ A identificação individual, por exemplo, em desastres de massa.

Actualmente, a tipagem de “*Short Tandem Repeats*” (STRs), constitui a técnica mais utilizada para a resolução destes casos, quer devido ao seu elevado polimorfismo, quer ao desenvolvimento de métodos de amplificação de múltiplos marcadores STR em simultâneo

(*multiplex*). Contudo, a caracterização de outros marcadores genéticos como os SNPs (*Single Nucleotide Polymorphisms*) poderá ser relevante na resolução de casos forenses, quer de casos de investigação biológica de parentesco, quer na área da criminalística biológica. No caso da investigação de parentesco biológico, destaca-se os casos em que o pretenso pai faleceu e, para se efectuar a perícia, tem de se recorrer ao estudo de restos cadavéricos, cujo material genético se encontra degradado.

Os SNPs são variações individuais num ponto particular do genoma, podendo consistir em inserções, deleções ou substituições nucleotídicas. São muito abundantes, com uma frequência de cerca de 1%, existindo cerca de um milhão ou mais por indivíduo. A principal motivação para a sua descoberta e detecção prende-se com o seu vasto potencial de aplicação em muitos campos, incluindo o estudo de doenças complexas, a farmacogenómica, e a evolução humana.

Estes marcadores têm vindo a despertar o interesse da comunidade forense por diversas razões. Primeiro, porque é possível efectuar, mais facilmente, a distinção entre uma mistura de amostras e uma amostra única; o processamento das amostras e a posterior análise dos resultados pode tornar-se mais automatizada (polimorfismos binários) e, por fim, os produtos de PCR resultantes da análise de SNPs podem ter um tamanho inferior a 100 pb, o que significa que são excelentes candidatos a marcadores nas situações em que o material genético está degradado. Contudo, uma vez que cada SNP apresenta apenas dois alelos possíveis, será necessária a análise de um mais elevado número de marcadores deste tipo para se atingir um poder de discriminação aceitável. Há estudos que apontam para a necessidade da análise de 25-45 SNPs para que se atinja um poder de discriminação similar ao de 13 STRs.

Assim, os SNPs nunca serão tão eficazes como os STRs na identificação genética individual, sendo que o uso simultâneo destes dois grupos de marcadores será a melhor opção.

Uma outra desvantagem apontada para a análise destes marcadores, prende-se com os métodos de genotipagem utilizados. Os métodos clássicos usados na tipagem de SNPs eram métodos demorados e que envolviam a utilização de enzimas de restrição, electroforese em gel de poliacrilamida e coloração por nitrato de prata. Outros métodos propostos para a análise de SNPs como a análise heteroduplex por dHPLC, não são suficientemente robustos para aplicação forense, para além das dificuldades na sua standardização. Têm sido utilizados outros métodos, tais como os que utilizam DNA microarrays, o MALDI-TOF MS ou a “Pyrosequencing”.

Todavia, de entre as novas tecnologias desenvolvidas para a análise de SNPs, umas das mais promissoras é a PCR em Tempo Real.

A PCR em Tempo Real é utilizada com diversas finalidades, em particular para a quantificação de ácidos nucleicos e para caracterização de determinados polimorfismos. A peculiaridade desta tecnologia prende-se com o facto do processo de amplificação ser monitorizado em tempo real, através da utilização de técnicas de fluorescência.

1.2 – O Cromossoma Y

O cromossoma Y humano é um pequeno cromossoma com cerca de 50 Mb, sendo o terceiro mais pequeno do genoma, representando apenas 2% da sua totalidade, tendo apenas maior dimensão do que os Cy 21 (47Mb) e Cy 22 (49Mb) [9]. A sua sequência foi publicada a 19 de Junho de 2003, pelo instituto Whitehead e a universidade de Washington. No entanto, apesar de apelidada de “completa”, [30] a referida sequência inclui apenas 23 Mb, dos aproximadamente 50 Mb presentes num cromossoma Y normal. A região ainda não sequenciada, aproximadamente 30 Mb, é de heterocromatina localizada no braço longo, não transcrita e composta por sequências altamente repetitivas.

A região não recombinante do cromossoma Y (NRY) compreende cerca de 95% da sua totalidade. As porções distais dos dois braços, conhecidas como regiões pseudo-autossômicas (PAR), recombinam com regiões homólogas do cromossoma X. A PAR1 localiza-se na porção distal do braço curto e tem aproximadamente 2,5 Mb, enquanto que a PAR2 constitui a porção distal do braço longo e tem menos de 1 Mb de tamanho (Figura 1).

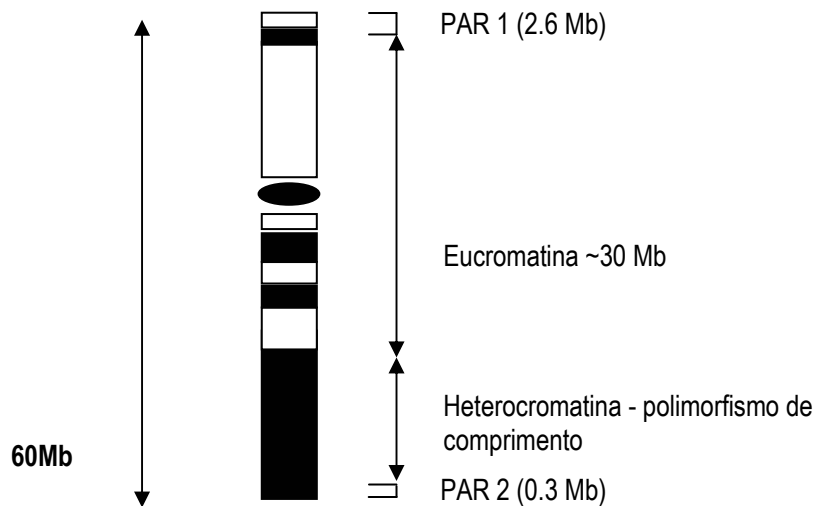


Figura 1: Cromossoma Y Humano (modificado de Jobling *et al*, 1997 [15]).

Ao contrário do que acontece com os marcadores autossômicos, que sofrem recombinação e podem estar sujeitos a modificações de geração em geração, os marcadores do cromossoma Y, assim como os do DNA mitocondrial, são chamados “marcadores de linhagem”. São transmitidos de geração em geração sem outras modificações que não as que surgem a partir de eventos mutacionais. Os marcadores do cromossoma Y permitem, deste modo, traçar linhagens paternas. A informação genética obtida a partir destes “marcadores de linhagem” é denominada de haplótipo e não de génotipo, uma vez que esta última designação corresponde a um único alelo por indivíduo. Assim sendo, os haplótipos resultantes de “marcadores de linhagem” não podem ser utilizados tão eficazmente na diferenciação de dois indivíduos, como

acontece com os genótipos obtidos na análise de marcadores autossômicos. Estes últimos constituem marcadores com segregação independente de geração em geração.

A vantagem da análise do cromossoma Y, no campo médico-legal, reside no facto de este se encontrar unicamente em indivíduos do sexo masculino. É o cromossoma Y que determina a masculinidade e o desenvolvimento dos testículos numa fase precoce da embriogénese [15], sendo a região responsável a designada SRY (*sex-determining region of the Y*). No âmbito da Genética Forense, o estudo do cromossoma Y é utilizado, fundamentalmente, na resolução de casos do âmbito criminal, nomeadamente no que respeita às agressões sexuais e na investigação de filiação.

Uma observação das estatísticas existentes permite verificar que a maioria dos crimes violentos é cometida por homens (>90%), e que a análise de DNA, para uma grande parte destes crimes se torna muito útil. Assim sendo, o estudo de marcadores genéticos, que existem apenas na fracção masculina de vestígios criminais, pode ser bastante informativo. Em casos de crimes sexuais, por exemplo, a caracterização de marcadores do cromossoma Y é muito vantajosa, principalmente na presença de grandes quantidades de DNA feminino comparativamente com a de DNA masculino. O estudo dos referidos marcadores proporciona informação específica acerca do agressor, e em situações que envolvam mais do que um agressor, a sua caracterização permite determinar o seu número.

Na resolução de investigação de filiação, no caso do filho ser do sexo masculino, a comparação da informação obtida a partir dos marcadores genéticos do cromossoma Y do suposto pai e do filho, permite excluir facilmente aquele a quem é atribuída a paternidade, no caso de se tratar de uma exclusão de paternidade.

As mesmas características que conferem ao cromossoma Y várias vantagens no âmbito da Genética e Biologia Forense, constituem, por outro lado, algumas das suas limitações. Não havendo recombinação, e sendo a ocorrência de mutações a única fonte de variação ao longo

do tempo entre indivíduos masculinos da mesma família, a maior parte da informação contida neste cromossoma é transmitida directamente de pai para filho. Assim, enquanto um resultado de exclusão pode ser muito informativo numa investigação (criminal/filiação), a concordância entre o perfil genético de um suspeito/pai e o perfil genético de uma evidência ou de um suposto filho, apenas significa partilha do haplótipo em questão, podendo este ser proveniente de qualquer indivíduo do sexo masculino familiar do suposto suspeito/pai, por linhagem paterna.

Contudo, esta última limitação, ou seja, a existência de familiares que partilham o mesmo haplótipo do cromossoma Y, pode ajudar na identificação de vítimas em desastres de massas, aumentando o número de amostras que podem ser estudadas em investigações de indivíduos desaparecidos, ou ainda, em investigações de filiação, em que não exista a possibilidade de se estudar o pretenso pai.

Do balanço deste conjunto de vantagens e desvantagens, relativo ao estudo de marcadores genéticos do cromossoma Y em Genética Forense, pode-se concluir que a análise de marcadores autossómicos é mais vantajosa, uma vez que proporciona um maior poder de discriminação, devendo, por este facto, a informação obtida a partir do estudo do cromossoma Y ser utilizada como complemento na conclusão de perícias médico-legais.

1.3 - Marcadores bialélicos

Os marcadores genéticos tradicionais para análise forense, como os 13 *loci* STR incluídos na base de dados nacional dos Estados Unidos, designada CODIS (*Combined DNA Index System*) implementada pelo FBI (*Federal Bureau of Investigation*), em 1998, têm sido muito utilizados [9]. Contudo, são usados no quotidiano dos laboratórios forenses, para além destes STRs autossómicos, outros incluídos em *kits* comerciais, podendo-se, na actualidade, estudar mais de 15 marcadores e a amelogenina. No entanto, especialistas forenses têm,

nalgunas situações específicas, que recorrer à análise de um maior número de polimorfismo para obter informação adicional acerca de uma amostra em particular.

A determinação do sexo, por exemplo, consegue-se através do estudo da amelogenina, incluída nos referidos kits comerciais utilizados na caracterização de STRs autossómicos; mas, em situações concretas de amostras muito degradadas, em que o resultado pretendido não é conseguido através da tipagem destes STRs, recorre-se, muitas vezes à análise do DNA mitocondrial, uma vez que este se encontra em maior número de cópias por célula do que o DNA nuclear sendo, por outro lado, mais resistente à degradação. É neste mesmo contexto que a utilização dos marcadores do cromossoma Y tem vindo a aumentar. Estes são cada vez mais utilizados, para além dos casos de DNA degradado, para obter a informação pretendida da fracção masculina de amostras em que há mistura de células, (por exemplo, em casos de violação).

Assim, é admissível pensar-se que, em poucos anos, uma grande variedade de marcadores e tecnologias estarão disponíveis para melhorar a investigação forense, na perspectiva de ajudar a resolução de casos a partir do estudo de vestígios biológicos. Uma classe de marcadores genéticos, que tem vindo a receber muita atenção nos últimos anos, em parte devido à sua abundância no genoma humano, que poderá ser de importância fundamental na evolução da Genética Forense, são os “*Single Nucleotide Polymorphisms*” (SNPs).

Um “*Single Nucleotide Polymorphisms*” ou SNP, é definido pela existência da variação de uma única base entre indivíduos, num ponto particular do genoma, desde que a frequência da variante menos frequente seja de 1%, ou superior [19]. Os SNPs representam a mais abundante fonte de variabilidade do genoma humano. Estima-se que, quando se comparam dois genomas, os SNPs ocorram, aproximadamente, em cada 1000-2000 pares de bases. Torna-se difícil fazer uma estimativa do número total de SNPs no genoma humano; no entanto, encontram-se já descritos cerca de 7 milhões, [19] em diferentes bases de dados públicas e privadas, a grande

maioria dos quais já validados; ou seja, confirmados como polimórficos num ou mais grupos populacionais.

No cromossoma Y e até ao ano de 2000, encontravam-se descritos 40 marcadores bialélicos [15,32] possíveis de ser analisados por PCR. Este número tem vindo a aumentar progressivamente, durante os últimos anos [29], acompanhado de estudos populacionais publicados [34], sendo que, de momento, o *SNP Consortium Ltd.* apresenta uma listagem de 841 SNPs do cromossoma Y, muitos ainda não validados e caracterizados.

A motivação inicial para a descoberta e detecção de Y-SNPs foi o elevado potencial de aplicação destes marcadores em muitos campos, incluindo o estudo de doenças complexas. Podem ser utilizados como marcadores para a identificação de genes responsáveis por este tipo de doenças, na farmacogenómica, na análise da variabilidade de resposta a diferentes drogas, e em estudos de evolução humana [31].

Na área forense, o interesse por este tipo de marcadores genéticos tem vindo a aumentar continuamente, e isto porque os SNPs possuem algumas características que tornam a sua utilização muito apropriada em estudos forenses. A primeira, e talvez a principal, prende-se com o facto de os produtos de PCR poderem ter menos de 100 pares de bases (bp), o que significa que a análise destes marcadores permite obter melhores resultados em amostras de DNA degradado do que com os STRs (produtos entre os 300-400 bp). Para além disso possuem baixas taxas de mutação, factor muito importante em estudos de filiação, [31] sendo também possível o uso da PCR *multiplex* para a sua análise. O processamento da amostra e a análise dos dados obtidos pode ser automatizada, uma vez que não se torna necessária a separação dos fragmentos através do seu tamanho. Não existem artefactos associados à caracterização dos alelos, o que simplifica a interpretação dos resultados. Por fim, é possível prever a origem étnica e determinar características fisionómicas, o que pressupõe a selecção de determinados SNPs [9].

A utilização destes marcadores apresenta, no entanto, algumas limitações.

A primeira, frequentemente referida, é o elevado número de SNPs a estudar. Inicialmente foi referido que o número necessário de SNPs deveria ser quatro vezes superior ao número de STRs utilizados [31]. Assim, o estudo de cerca de 60 SNPs seria o número conveniente para atingir um poder de discriminação semelhante ao obtido, para a realização da mesma análise, recorrendo aos STRs em *multiplex*, utilizados na área da Genética Forense. Contudo, actualmente defende-se que, no caso dos Y-SNPs, este número pode ser muito reduzido através da sua adequada selecção, devendo esta ser baseada na sua especificidade populacional, seguindo uma estratégia hierárquica relacionada com a árvore filogenética proposta pelo YCC (*Y Chromosoma Consortium*). As vantagens relativamente a custos envolvidos, não são evidentes. Apesar do custo actual do estudo de SNPs ser baixo, isto é, substancialmente inferior ao dos kits comerciais usados para caracterização dos *loci* STR, a verdade é que, por exemplo, em casos de investigação de filiação biológica, a utilização dos referidos kits torna todo o processo bastante simplificado, compensando, deste modo, os custos envolvidos. Para além das razões apontadas anteriormente, a experiência acumulada dos últimos dez anos ao nível dos STRs, faz toda a diferença. Se existirem mutações ou polimorfismos em regiões flanqueadoras de *loci* STR, estas são conhecidas e facilmente detectadas, no caso dos SNPs são situações que poderão causar problemas. No entanto, podem ser ultrapassadas efectuando uma extensa validação dos SNPs em diferentes grupos populacionais [4].

Assim, actualmente não é razoável pensar que os SNPs irão substituir os STRs como marcadores genéticos de eleição na área forense. Não restam, no entanto, dúvidas de que a utilização destes marcadores genéticos, em algumas aplicação específicas, como sejam a definição de haplogrupos do cromossoma Y e de DNA mitocondrial para inferir acerca da origem populacional de determinada amostra vestigial, pode ser muito útil. Torna-se óbvio que a

Genética Forense pode e deve tirar partido da grande quantidade de informação que tem sido publicada nos últimos anos, não apenas em relação ao crescente número de marcadores bialélicos disponíveis nas bases de dados, como, também, no que concerne ao desenvolvimento das tecnologias para a sua caracterização.

A standardização e a validação interlaboratorial das estratégias de tipagem a utilizar, será fulcral para um aumento do uso dos SNPs na área forense. Não existe um método ideal para a tipagem de SNPs, e a selecção da técnica mais adequada irá depender dos objectivos do utilizador. Para a resolução de perícias forenses estes objectivos passam pela necessidade de utilizar técnicas com uma boa possibilidade de recorrer à realização de PCR *multiplex*, técnicas robustas, precisas e de elevada especificidade, em que os custos reduzidos também não devem ser esquecidos.

1.4 - Metodologias para Genotipagem de SNPs

As tecnologias para genotipagem de SNPs têm sofrido um rápido desenvolvimento nos últimos anos. Esta é uma situação que pode ser explicada pelo facto destes marcadores genéticos serem utilizados para o mapeamento de doenças complexas. Como resultado desta conjugação de interesses, tem sido descrita uma grande variedade de protocolos para a sua caracterização.

A escolha da melhor tecnologia a utilizar para diferentes aplicações específicas, irá depender dos objectivos finais, assim como do número de marcadores a usar e do tamanho da amostra. No que diz respeito à área forense, poderão ser usadas diferentes tecnologias dependendo das situações (casos criminais, em que a quantidade e a qualidade do DNA é um ponto crítico; resolução de casos de parentesco biológico ou construção de bases de dados

genéticas com interesse criminal). As melhores metodologias para a genotipagem de SNPs autossómicos, do cromossoma Y ou mitocondriais, serão também diferentes, tendo em consideração o número de marcadores a analisar por amostra e a estratégia desta mesma análise. A escolha passará ainda pela necessidade de realização da PCR *multiplex* e, também, dos próprios SNPs a estudar. De um modo geral, todas as tecnologias com limitações na realização de *multiplex* devem ser excluídas como candidatas para o estudo de amostras forenses podendo, no entanto, ser muito úteis na determinação das frequências alélicas dos diferentes SNPs e na criação das bases de dados genéticas com interesse criminal.

As diferentes tecnologias que têm surgido, baseiam-se em diferentes métodos de discriminação alélica e na utilização de várias plataformas de detecção. Para melhor compreender cada uma destas tecnologias é importante distinguir entre reacções de discriminação alélica, metodologias empregues para esta discriminação e métodos de detecção. Os produtos resultantes de uma reacção de discriminação alélica podem ser detectados usando mais do que um método, e o mesmo método de detecção pode analisar produtos resultantes de diferentes reacções e metodologias.

De seguida, serão descritas algumas das tecnologias de que se tem vindo a falar, de forma sucinta.

1.4.1 - Mini-sequenciação

Neste método de caracterização de SNPs, um iniciador ou *primer* específico liga-se à cadeia de DNA complementar, numa região imediatamente adjacente ao local da existência do SNP [31]. A extensão da cadeia ocorre por acção da DNA polimerase, adicionando um único nucleótido complementar à base existente no local polimórfico. Este método é baseado na grande precisão com que o nucleótido é incorporado pela DNA polimerase. A detecção de

SNPs, no genoma humano, exige uma amplificação prévia da região flanqueadora do local de variação.

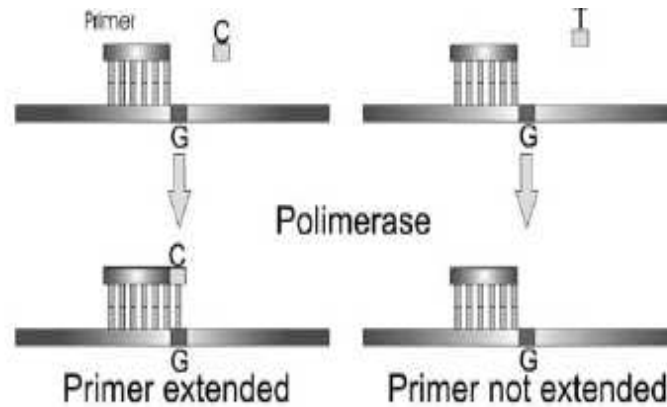


Figura 2: Mini-sequencição – emparelhamento de um *primer* no DNA alvo, imediatamente antes da localização do polimorfismo e sua extensão com um único nucleótido complementar à base polimórfica.

Após a amplificação, e antes de se proceder à reacção de mini-sequencição, torna-se necessária a remoção dos reagentes em excesso resultantes da amplificação por PCR, como *primers* e dNTPs. A análise posterior dos produtos resultantes da extensão do primer, pode ser efectuada utilizando diferentes metodologias. Assim, podem-se usar nucleótidos marcados ou não marcados, a combinação de ddNTP com dNTP, ou somente ddNTP. Os mais utilizados para a detecção dos referidos produtos de extensão, são a electroforese capilar e detecção por fluorescência (SnaPshot), a espectrometria de massa (MS), MALDI-TOF (*Matrix Assisted Laser Desorption/Ionization Time-Of-Flight*) e microarrays.

1.4.1.1 - SnaPshot

Este método baseia-se na reacção de produtos de mini-sequencição seguida de electroforese e detecção da fluorescência. Trata-se do uso do *kit SnaPshot™* da *Applied*

Biosystems, que permite a realização simultânea de várias reacções de extensão com diferentes *primers*, utilizando ddNTPs marcados com fluorescência.

Um primer não marcado emparelha na extremidade 3', na base imediatamente adjacente ao local do SNP que, por sua vez, se liga a um único ddNTP marcado com um fluoróforo, sendo este diferente para cada ddNTP. Este método permite a realização de PCR *multiplex*, ou seja, utilizando vários *primers*, pode-se efectuar a análise simultânea de diferentes SNPs. Para tal, são adicionados nucleótidos na extremidade 5' de cada *primer*, o que irá permitir a sua posterior separação por tamanho dos fragmentos obtidos. Utiliza-se, frequentemente, para o efeito ou uma cauda poly(T) com 3 a 5 pares de bases, ou sequências que não são complementares a nenhuma região do genoma humano.

A cor do pico, presente na análise por mini-sequenciação, representa o nucleótido específico do local de variação; enquanto que o local ocupado por este mesmo pico corresponde ao seu tamanho, que está relacionado com o *locus* ou posição do SNP que, por sua vez, varia consoante a cauda adicionada aos diferentes *primers* utilizados. As amostras homozigóticas para um determinado alelo surgirão como um único pico e as heterozigóticas, com dois picos adjacentes.

A grande vantagem desta tecnologia reside no facto dos produtos resultantes da extensão dos *primers* poderem ser analisados por electroforese e detecção da fluorescência, em instrumentos já existentes na maioria dos laboratórios de Genética Forense. Esta é uma técnica sensível, permite a análise sinultânea de vários polimorfismo e é suficientemente robusta para a tipagem de amostras criminais.

1.4.1.2 - MALDI-TOF MS

Esta é uma tecnologia que se baseia na determinação do peso molecular dos produtos resultantes de uma reacção de mini-sequenciação. Pode dizer-se que é um método de detecção

muito mais directo, quando comparado com as outras metodologias, através das quais se infere a identidade dos produtos a partir da monitorização da fluorescência resultante de moléculas marcadas [31]. O MALDI-TOF proporciona um elevado poder de resolução, pois permite distinguir qual o ddNTP incorporado ao *primer* de extensão.

Os produtos resultantes da reacção de mini-sequenciação são depositados numa matriz na superfície de uma placa ou chip. A matriz e os produtos de DNA são volatilizados por um raio laser, num processo designado de volatilização. A energia do laser é transferida para a matriz, resultando na sua vaporização e envio dos pequenos fragmentos de DNA, produtos da extensão do *primer*, para o “*flight tube*”. Os fragmentos de DNA deslocam-se, posteriormente, desde o “*flight tube*” até um detector por aplicação de um campo eléctrico. O tempo que medeia entre a aplicação do impulso eléctrico e a colisão dos produtos de DNA no detector, denomina-se “*time of flight*” e representa uma medida precisa do peso molecular dos produtos. Assim, moléculas mais pequenas fazem o percurso em menos tempo, chegando em primeiro lugar ao detector do que as moléculas mais pesadas. Existem também programas informáticos apropriados que convertem o “*time of flight*” na massa exacta de cada molécula.

Uma das limitações desta tecnologia é a necessidade da utilização de uma amostra muito pura. Esta técnica permite, também, a realização de *multiplex* desde que os produtos resultantes da mini-sequenciação não sejam sobreponíveis em termos de tamanho. Neste caso, segue-se o mesmo procedimento utilizado no *SnaPshot*, ou seja, a adição de caudas na extremidade 5' do *primer*.

1.4.1.3 - Microarrays

O uso de microarrays é também compatível com a genotipagem de SNPs por mini-sequenciação, a qual pode ser realizada na superfície de *chips* ou em solução.

Na realização da reacção de mini-sequenciação em chips, os *primers* são fixados na sua superfície, podendo também ser designada de APEX (*Arrayed Primer Extension*). De seguida, estes sofrem extensão por acção da DNA polimerase, com ddNTPs marcados, sendo posteriormente medida a fluorescência do microarray por *scanner*.

Quando a mini-sequenciação é realizada em solução, é efectuada uma reacção em que se utilizam *primers* com uma sequência única e específica na sua extremidade 5'. Assim, cada SNP possuirá uma cauda de identificação específica. Procede-se, de seguida, a uma hibridização dos produtos da mini-sequenciação com sequências complementares reversas às caudas de cada *primer*, sequências estas que se encontram nos “spots” dos microarrays.

Esta é uma metodologia que tem a vantagem de permitir a análise em grande escala, podendo-se estudar centenas ou milhares de SNPs em poucas amostras; apresenta, no entanto, a limitação dos chips não poderem ser construídos ou modificados em pequenos laboratórios, de modo a possibilitar a selecção de marcadores de interesse.

1.4.2 - PCR em Tempo Real

A PCR em Tempo Real é utilizada com diversas finalidades, particularmente para a quantificação de ácidos nucleicos e para genotipagem. Desde o seu conhecimento em 1996, o número de publicações com a utilização da PCR em Tempo Real tem vindo a crescer de uma forma quase exponencial. A peculiaridade desta tecnologia prende-se com o facto do processo de amplificação ser monitorizado em tempo real, através da utilização de técnicas de fluorescência.

São usadas sondas de hibridização marcadas com fluorocromos e termocicladores de “tempo real”, que acumulam duas funções, a de termocicladores e a de medidores de fluorescência. São monitorizadas, em “tempo real”, modificações nas características da

fluorescência das sondas que se ligam aos produtos da PCR de sequências alvo, durante a amplificação. As diferentes afinidades dessas sondas com os produtos de PCR de genes normais ou mutantes ou, ainda, de sequências de DNA de interesse que possam apresentar dois alelos alternativos, resultará em diferentes genótipos. O método anteriormente referido é designado ASO (*Allele Specific Hybridization*), e baseia-se na distinção entre dois fragmentos de DNA que apenas diferem num nucleótido, por hibridização [31]. Para o efeito, são desenhadas duas sondas específicas para cada alelo, normalmente com o local polimórfico no centro da sequência e em que, após optimizações do método, apenas ocorre o emparelhamento perfeito entre a sonda e o DNA alvo, sendo esta ligação estável.

Uma das possíveis abordagens para a análise dos resultados obtidos, utilizando este método, é a análise das curvas de fusão ou de “melting”, isto é, as sequências amplificadas podem ser caracterizadas em função da sua temperatura de “melting”. Estas representam a temperatura dependente da fluorescência e são função do tamanho e composição em bases do produto amplificado. A sua detecção pode ser efectuada, quer utilizando fluoróforos específicos para DNA de cadeia dupla (ds) como o SYBR Green I, quer com sondas específicas para cada sequência como os “*molecular beacons*” e as sondas de hibridização. Tem sido descrito que a análise de curvas de “melting”, através da utilização do SYBR Green, pode ser usada para a caracterização de polimorfismos de inserção/delecção e de SNPs.

As metodologias mais comuns com base neste procedimento, usadas em diferentes áreas, são as sondas FRET e as TaqMan®.

1.4.2.1 - Metodologia com sondas FRET

A detecção de SNPs, através desta metodologia, foi desenvolvida inicialmente para uma plataforma específica de PCR em Tempo Real, o LightCycler® (Roche). Este é um termociclador

ultra-rápido, que utiliza o ar para promover o aquecimento e arrefecimento dos capilares onde se realiza a reacção de PCR. Isto permite uma redução efectiva do tempo de cada ciclo da PCR.

Para a sua execução são desenhadas dois oligonucleotidos marcados com fluorescência. Uma das sondas é marcada com fluoresceína na extremidade 3', e a outra sonda é marcada com outro fluoróforo na extremidade 5'. As sequências das duas sondas são escolhidas de forma a que estas hibridizem em locais adjacentes do DNA alvo (*target*). Quando estes oligonucleótidos estão hibridizados, os dois fluoróforos estão posicionados na proximidade um do outro. Quando a fluoresceína é excitada emite fluorescência, que vai excitar o fluoróforo da outra sonda que, por sua vez, emite fluorescência. Esta sonda receptora só irá absorver a fluorescência emitida pela primeira, quando as duas sondas se encontrarem ligadas a uma curta distância na cadeia de DNA .

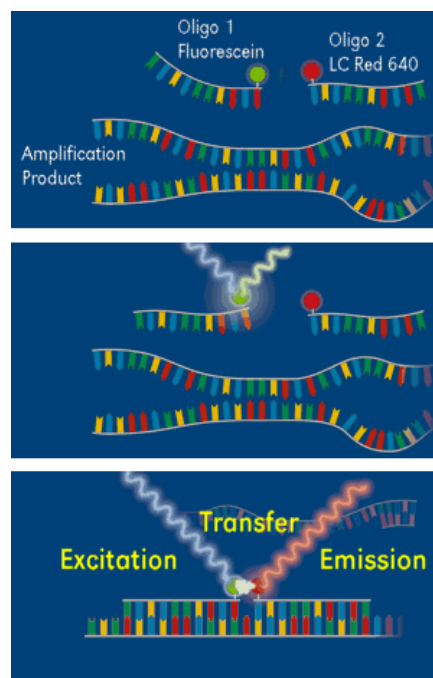


Figura 3: Mecanismo de funcionamento da metodologia de detecção de SNPs com utilização de sondas FRET.

Este processo de transferência de energia de um marcador fluorescente para um segundo é denominado “*Fluorescence Resonance Energy Transfer*” (FRET).

A utilização deste método com sondas de hibridização tem as seguintes vantagens:

- ✓ A fluorescência é o resultado directo da hibridização de duas sondas independentes, o que resulta na obtenção de resultados com alta especificidade.
- ✓ A fluorescência resultante de sondas de hibridização não depende da acção da actividade exonucleotídica da polimerase, ao contrário dos ensaios TaqMan®. Neste método, pode-se acompanhar a temperatura de “melting” das sondas.
- ✓ As sondas de hibridização são facilmente desenhadas, sintetizadas e optimizadas.

As sondas, enquanto se encontram em solução, não emitem qualquer fluorescência detectável a 640nm, tal só acontece depois destas hibridizarem com os respectivos locais complementares na cadeia de DNA, e quando se encontram suficientemente próximas. Quando a variante alélica está presente, o emparelhamento com a sonda é menos estável, implicando uma transição para valores de fluorescência mais altos, a temperaturas mais baixas. Quando não ocorre mutação, a sonda é mais estável, sendo necessária uma temperatura mais elevada para que ocorra o desemparelhamento.

Com esta metodologia podem ser genotipados, simultaneamente, mais do que um SNP, através da combinação do uso de diferentes marcadores de fluorescência com o desenho de sondas com distintas temperaturas de “melting” (T_m). Uma primeira versão do LightCycler® permitia a detecção de apenas dois marcadores de fluorescência, um modelo mais recente possibilita já a distinção entre 4 fluoróforos. Encontra-se já descrito um *multiplex* para a detecção de 4 diferentes SNPs do cromossoma Y, com a utilização da metodologia anteriormente descrita [17]. A principal vantagem desta tecnologia é, sem dúvida, a sua grande sensibilidade.

1.4.2.2 - Metodologia com sondas TaqMan®

Esta metodologia tem por base a actividade exonucleotídica 5' da Taq DNA polimerase, que promove a rotura da ligação da sonda que hibridiza com a cadeia complementar, criando um sinal de fluorescência. Para este efeito são necessárias duas sondas TaqMan® que irão diferir no local polimórfico, uma será complementar do alelo ancestral e a outra da sua variante. Estas duas sondas são marcadas com dois fluoróforos diferentes, um na extremidade 5', o "reporter", e outro na extremidade 3', o "quencher". Enquanto as sondas se mantêm intactas o "quencher" absorve a fluorescência emitida pela sonda "reporter". Durante a PCR e na fase de emparelhamento, as sondas TaqMan® hibridam com a sua cadeia complementar. Durante a fase de extensão a sonda "reporter" é clivada pela actividade 5' da Taq DNA polimerase, e o fluoróforo é libertado, ocorrendo assim um aumento do sinal de fluorescência. Com o acumular do número de ciclos da PCR vai haver o incremento da intensidade da fluorescência, sendo esta proporcional ao aumento do número de produtos de amplificação.

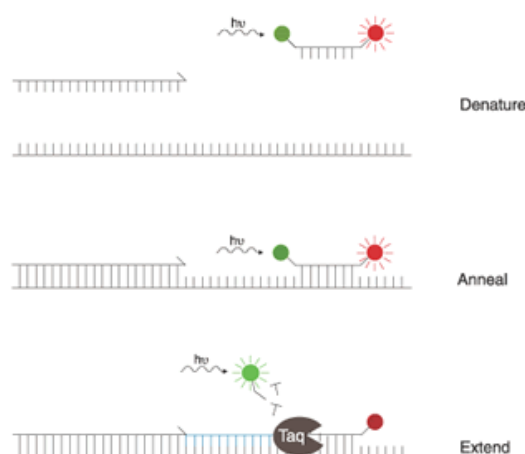


Figura 4: Metodologia de detecção de SNPs com sondas TaqMan.

A genotipagem de cada amostra é efectuada através da determinação da intensidade da fluorescência emitida. Um aumento do sinal de fluorescência de um dos fluoróforos, indica homozigotia para um dos alelos, conforme a cor da fluorescência utilizada na marcação do respectivo alelo. O aumento das duas cores será indicativo da presença de uma heterozigotia.

Estas metodologias designadas de homogéneas ou em “tubos fechados” [19], em que todos os reagentes necessários para a amplificação e genotipagem são adicionados simultaneamente, são rápidas, fáceis de executar e pouco sujeitas a contaminações.

Após esta visão resumida das diferentes tecnologias disponíveis para a genotipagem de SNPs, são vários os aspectos a ter em conta para a sua aplicação na área forense, tais como a sensibilidade do método, a reprodutibilidade, a precisão, a capacidade para a realização de PCR *multiplex* e a sua possibilidade de utilização tendo em conta o número de amostras e/ou o número de SNPs a caracterizar. É ainda importante não esquecer a flexibilidade proporcionada por cada tecnologia, o tempo a despendar e os custos envolvidos, podendo ser utilizados diferentes plataformas.

Relativamente às possibilidades de cada método, no que diz respeito ao número de amostras e SNPs a genotipar, tudo dependerá da aplicação específica pretendida. Em algumas situações é necessário genotipar poucos SNPs, mas num grande número de amostras. Noutros casos, o contrário, um grande número de SNPs em poucas amostras. Para os propósitos forenses, habitualmente, trabalha-se com número reduzido de amostras, por exemplo, em casos de filiação biológica e de criminalística, enquanto para outros fins é necessário analisar um grande número de amostras em simultâneo, como seja na construção de bases de dados genéticas.

Um importante factor limitativo, no campo da Genética Forense, e que afectará qualquer método e tecnologia, é a quantidade e a qualidade do DNA disponível. Para ultrapassar esta dificuldade é mais vantajoso o uso de tecnologias que empreguem uma PCR, antes da execução

da discriminação alélica, em vez de tecnologias em que o SNP é detectado directamente no DNA genómico.

As tecnologias designadas de homogéneas e baseadas em reacções de hibridização FRET, como as do LightCycler[®], as TaqMan[®] e os “Molecular Beacon”, realizam a PCR e a discriminação alélica numa mesma reacção. Esta vantagem permite a eliminação de alguns passos intermédios, favorecendo a automatização e a capacidade da análise em simultâneo de um grande número de amostras, especialmente utilizando plataformas como as que pressupõem o uso de sondas TaqMan[®]. A principal limitação destes métodos é, sem dúvida, a reduzida capacidade para a realização de PCR *multiplex*. São, por isso, boas tecnologias para a selecção de SNPs de interesse e para a construção de bases de dados genéticas, não sendo tão apropriadas para resolução de casos criminais na rotina laboratorial.

Na actualidade, em laboratórios de Genética Forense, as tecnologias mais utilizadas são as de mini-sequenciação, em particular o SnaPshot [8,13,26], porque o método de detecção é realizado em instrumentos de electroforese capilar, os mesmos que são usados para a análise dos STRs. A capacidade de realização de *multiplex* com esta tecnologia é aceitável, mas é necessário algum trabalho na optimização do desenho dos *primers* e na reacção de mini-sequenciação, para a detecção de cada SNP. Os outros métodos de detecção, como a MALDI-TOF e os “microarrays”, podem ter aplicação forense, [21] apresentando, no entanto, a desvantagem de necessitarem de equipamento específico. A PCR capacidade de se efectuar *multiplex* do MALDI-TOF é baixa, e em relação aos “microarrays”, pode-se apontar a sua baixa reprodutibilidade e dificuldade de validação.

Assim sendo, pode-se concluir que as tecnologias que parecem acumular as características necessárias para a caracterização de SNPs com fins forenses, devem ser exploradas de forma a melhor serem conhecidas as suas limitações e vantagens. Em face do exposto, será pouco provável que apenas uma dessas tecnologias reúna todos os requisitos

necessários e não será tarefa fácil definir qual a melhor tecnologia para a genotipagem de SNPs nesta área.

1.5 - Selecção dos SNPs a estudar

Nos últimos anos, tem-se assistido a um aumento da informação disponível relativamente à região não recombinante do cromossoma Y (NRY). Este aumento deve-se, em parte, à descoberta dos polimorfismos bialélicos nesta região. Nos finais de 1996 eram conhecidos pouco mais do que 60 SNPs localizados na região NRY.

Por esta altura, a informação, que podia ser retirada do estudo destes marcadores genéticos era utilizada para estabelecer linhagens paternas. Com o passar dos anos começou a perceber-se a utilidade que estes SNPs poderiam ter em áreas como a reconstrução genealógica, em Genética Médica, em estudos de evolução humana e, mais recentemente, em Genética Forense.

Em 2001 conheciam-se já cerca de 200 SNPs do cromossoma Y, sendo estes detectados pelo método de DHPLC (*Denaturing High-Performance Liquid Chromatography*) [12,29,34].

Com este aumento do número de SNPs conhecidos, houve um incremento do número de publicações em que estes eram referenciados, tendo surgido um problema relacionado com a nomenclatura mais apropriada para a classificação dos haplogrupos. Até 2002 eram conhecidas 7 diferentes nomenclaturas, o que dificultava a comparação de resultados entre as várias publicações [15,18,29,32]. Nesta perspectiva, foi feito um esforço para estabelecer uma equivalência entre todos os sistemas existentes e, em 2002, o “Y Chromosome Consortium”

chegou a um consenso, publicando uma proposta de nomenclatura sistemática e sem ambiguidades para os haplogrupos definidos pelos Y-SNPs [1].

Nesta nomenclatura, os haplogrupos são associadas em “clades”, denominadas por letras maiúsculas de A-R, sendo estas divididas em “subclades”. As “subclades”, dentro de cada haplogrupo, denominam-se acrescentando um sistema alternativo alfanumérico a seguir à letra maiúscula correspondente ao haplogrupo principal. Dentro das “subclades” pode-se, ainda, distinguir os “nested clades”, sendo estes classificados usando letras minúsculas. Esta é, portanto, uma nomenclatura que alterna números e letras. Acresce ainda que é suficientemente flexível, de modo a permitir alterações ocasionadas pela descoberta de novas mutações e linhagens. Como resultado final, permite a construção de uma árvore filogenética, partindo da caracterização dos polimorfismos bialélicos conhecidos.

Para este estudo foram utilizados 8 marcadores bialélicos do cromossoma Y (M2, M170, M172, M9, M45, M207, M173, P25), seleccionados com base numa estratégia hierárquica, segundo a conhecida e estável filogenia do cromossoma Y. Uma estratégia hierárquica significa que se seleccionam SNPs capazes de explorar as “clades” mais basais da árvore, e também os braços mais relevantes para a definição dos haplogrupos pertinentes; neste caso, os haplogrupos mais representativos em populações europeias.

Foram então várias as razões consideradas na altura da selecção dos SNPs estudados. Os primeiros a ser genotipados foram o M2 e o M9, e a principal razão da sua escolha foi a existência prévia do estudo destes marcadores por uma técnica convencional: a PCR-RFLP no grupo de amostras a usar. Deste modo, estas poderiam funcionar como controlos numa fase inicial do trabalho. Todos os outros foram escolhidos por serem informativos em populações europeias previamente estudadas, pela possibilidade da sua detecção por PCR, pela existência confirmada, por estudos já publicados, de duas variantes alélicas e, finalmente, por permitirem a definição de uma filogenia de haplogrupos.

SNP	Polimorfismo	Haplogrupo YCC definido	Referência
M2	A -> G	E3a	Seielstad <i>et al</i> 1994
M170	A -> C	I	Shen <i>et al</i> 2000
M172	T -> G	J2	Shen <i>et al</i> 2000
M9	C -> G	K - R	Underhill <i>et al</i> 1997
M45	G -> A	P - R	Shen <i>et al</i> 2000
M207	A -> G	R	Shen <i>et al</i> 2000
M173	A -> C	R	Shen <i>et al</i> 2000
P25	C -> A	R1b	Hammer <i>et al</i> 2001

Tabela 1: Características dos polimorfismos bialélicos [28,29,32].

II – Objetivos

No presente estudo procurou-se contribuir para a elaboração de protocolos realizáveis em instrumentos de PCR em Tempo Real, para a caracterização de SNPs com interesse na área da Genética Forense. Estes protocolos permitirão a utilização desta metodologia na tipagem destes marcadores genéticos que serão, possivelmente, utilizados, num futuro próximo, nesta área da Medicina Legal.

Pretende-se assim:

- 1 Validar protocolos para a tipagem de SNPs, utilizando a técnica de PCR em Tempo Real, em três plataformas.
- 2 Analisar as vantagens e limitações do uso desta técnica na área forense.

A realização destes objectivos de carácter geral, permitirão atingir ainda outros mais específicos. Os SNPs utilizados neste estudos são SNPs do cromossoma Y, com uma possível utilidade na resolução de perícias médico-legais do âmbito da Genética e Biologia Forense.

A validação dos protocolos permitirá avaliar algumas das características das metodologias utilizadas, como a sua robustez, especificidade e sensibilidade. Possibilitará, ainda, conhecer alguns prós e contras da utilização dos SNPs do cromossoma Y, bem como permitir efectuar a sua melhor selecção para as diferentes aplicações forense.

Será efectuada uma aproximação a um estudo populacional, proporcionando uma ideia aproximada das frequências dos SNPs estudados numa amostra da população do Norte de Portugal.

III – Material e Métodos

3.1 – Amostras

As amostras utilizadas neste trabalho são provenientes de indivíduos fenotipicamente saudáveis, não relacionados, envolvidos em casos de investigação biológica de parentesco, colhidas mediante obtenção de consentimento informado. Para o estudo em questão, as referidas amostras eram desprovidas da respectiva identificação, para assegurar o anonimato. As amostras eram de sangue venoso colhido em tubos com o anticoagulante EDTA K3. Estas, foram armazenadas a -20°C até se proceder à extracção de DNA, a partir do sangue total.

Para a análise dos SNPs, estudaram-se 33 indivíduos do sexo masculino.

3.2 – Extracção do DNA

Procedeu-se à extracção do DNA recorrendo, fundamentalmente, a dois métodos. A partir de sangue líquido, por um método modificado [20], que utiliza a extracção orgânica com fenol-clorofórmio-álcool isoamílico (casos de filiação); e a partir de manchas de sangue, pelo método descrito por Singer-Sam *et al.* (1989) [37] que utiliza uma resina quelante, o Chelex (casos de filiação e criminalística biológica).

3.2.1 – Extracção com fenol-clorofórmio-álcool isoamílico

Lise/digestão celular

A 200 μl de sangue periférico:

- ✓ Juntou-se 500 μl de Tampão de Lise.
- ✓ Incubou-se em banho-maria com agitação, 2 horas a 56°C (ou O.N. a 37°C).

Extracção

Após incubação, procedeu-se à extracção:

- ✓ Juntou-se 200 µl de fenol-clorofórmio-álcool isoamílico (25:24:1). Agitou-se no “vortex” para homogeneizar as fases.
- ✓ Centrifugou-se a 12000 rpm, 5 minutos.
- ✓ Recuperou-se a fase aquosa (~400 ml) para um tubo, devidamente etiquetado.

Precipitação e lavagem

Ao tubo com a fase aquosa após a extracção:

- ✓ Adicionou-se 1/10 do seu volume (40µl) de Acetato de Sódio a 2M, 1 µl de Glicogénio (20mg/ml) e dois volumes (800µl) de Etanol Absoluto frio (-20°C). Inverteram-se os tubos suavemente para misturar todos os reagentes.
- ✓ Os tubos que continham o DNA visível, centrifugaram-se 5 minutos a 13000 rpm.
- ✓ Se o DNA não fosse visível (a maior parte das vezes), colocavam-se os tubos 1 hora a -20°C e centrifugavam-se 20 minutos a 13000 rpm.
- ✓ Em qualquer dos casos anteriores, decantou-se o sobrenadante de cada um dos tubos.
- ✓ Lavaram-se os “pellets” com 1 ml de Etanol a 80% frio (1000µl).
- ✓ Centrifugaram-se 1 minuto a 12000 rpm e retiraram-se novamente os sobrenadantes.
- ✓ Deixou-se secar o “pellet” de DNA no concentrador de DNA (DNA SpeedVac, DNA 110-Savant).

- ✓ Ressuspendeu-se em volume adequado de H₂O MiliQ ou tampão TE (~20-40μl).
- ✓ Quantificou-se o DNA presente em cada um dos tubos.

Reagentes

Cloreto de Sódio (NaCl) Merk (6404)

Acetato de Sódio (Na₂CH₃CO₂) Merk (6267)

Proteínase K Sigma (PO390)

SDS Sigma (L-4390)

Fenol-clorofórmio-álcool isoamílico USB (20084)

Etanol Absoluto Merk (0983)

Tris Sigma (T-8524)

EDTA-dissódico Sigma (E-5134)

Ditiotreitol Sigma (D-0632)

Glicogénio Sigma (G-8876)

Soluções

1. Tampão de Lise: Tris 10mM pH8, EDTA 10mM, NaCl 0,1M, SDS a 2% (w/v), DTT 40mM, Proteinase K a 0,2 mg/ml.

Nota: O DTT e a Proteinase K só se juntaram no momento em que se adicionaram às amostras.

2. Acetato de Sódio 2M.

3. Glicogénio a 20 mg/ml.

4. Etanol a 80% (v/v)

5. Tampão TE-Tris 10 mM, EDTA 1 mM, pH 7,5.

3.2.2 – Extracção com Chelex

1. Pipetou-se 1ml de H₂O MiliQ para um microtubo estéril de 1,5ml.
2. Juntou-se 3mm² da mancha ou 4µl de sangue total e agitou-se suavemente.
3. Incubou-se à temperatura ambiente de 15 a 30 minutos.
4. Agitou-se em “vortex” 5-10 segundos (cuidadosamente ou inverteu-se periodicamente).
5. Centrifugou-se 2 a 3 minutos a 10000 – 15000 rpm.
6. Com uma micropipeta, eliminou-se todo o sobrenadante, à excepção de 20-30µl, deixando o tecido ou papel no tubo, quando se tratava de uma mancha.
7. Juntou-se Chelex a 5% (w/v) (170-180µl) até atingir um volume final de 200µl, com uma ponta de 1ml, ou uma ponta amarela com a extremidade cortada, para ter maior diâmetro.
8. Incubou-se a 56°C durante 15 a 30 minutos em banho-maria.
9. Agitou-se vigorosamente em “vortex” 5-10 segundos.
10. Mergulhou-se em água a ferver durante 8 minutos, em flutuadores. Nota: com uma agulha estéril, perfuraram-se as tampas dos tubos, antes de os levar à ebulição.
11. Agitou-se vigorosamente em “vortex” 5-10 segundos.
12. Centrifugou-se 2 a 3 minutos a 10000 – 15000 rpm.
13. Recolheu-se o sobrenadante e colocou-se noutro tubo, tendo o cuidado de não arrastar a resina.
14. Juntou-se uma parte do sobrenadante aos tubos que já continham os reagentes necessários para efectuar a PCR (até 20% do volume total da mistura).

Reagentes

Chelex Sigma (7981)

Soluções

Solução stock de Chelex 100 a 5% (w/v) em H₂O MiliQ (ponta de pipeta de 1ml).

3.3 – Caracterização dos SNPs

A metodologia experimental utilizada para a caracterização dos SNPs a estudar, baseou-se na aplicação da técnica de PCR em Tempo Real, recorrendo ao uso de sondas TaqMan[®].

3.3.1 – Selecção dos SNPs

A selecção dos SNPs, baseou-se quer nas frequências publicadas para cada um dos SNPs em diferentes estudos populacionais geograficamente próximos, quer na sua importância para a definição dos diferentes haplogrupos. Os SNPs seleccionados definem os haplogrupos que têm sido descritos como sendo os mais frequentes na Europa [35], estando a maioria localizados nos ramos principais da árvore filogenética do cromossoma Y, proposta pelo YCC (Y Chromosome Consortium) [1].

3.3.2 – Amplificação por PCR em Tempo Real

Após a selecção dos SNPs a estudar, seguindo a estratégia atrás referida, foi localizada e confirmada a sequência de cada um deles através de um BLAST, utilizando para tal os *primers* usados para a sua amplificação e referidos em publicações anteriores, [3,35].

Procedeu-se, de seguida, ao desenho dos pares de *primers* (*forward* e *reverse*) e das sondas (Tabela 2) correspondentes, para a identificação de cada SNP seleccionado, recorrendo-se ao programa “*Primer Express Software Version 2.0*” da *Applied Biosystems*, que utiliza critérios específicos para as sondas TaqMan®. As sondas utilizadas são Sondas TaqMan® MGB (*Minor Groove Binder*). Estas sondas TaqMan® apresentam duas características particulares que as distinguem das outras sondas TaqMan®. O “quencher” na extremidade 3’ destas sondas não apresenta qualquer fluorescência, o que permite uma medida mais precisa da contribuição do fluoróforo da extremidade 5’. Estas sondas MGB possuem também na mesma extremidade 3’ uma molécula própria, que permite uma ligação mais forte da sonda à cadeia complementar, aumentando, deste modo, a respectiva temperaturas de “melting” (Tm).

SNP	Polimorfismo	GenBank/bdSNP	Primer 5'-3' Fw/Rv	Sonda FAM	Sonda NED
M2	A→G		CCTTGGGCTCCCTGTTTAA TGGAAAATACAGCTCCCCCTTTA	AAAGTCCgTGAGATCT	AAGTCCaTGAGATCTG
M9	G→C	rs3900	CTGCAAAGAAACGGCCTAAGA TGTAAGACATTGAACGTTTGAACATG	TGGTTGAATCCTCTTTA	ATGGTTGAATTCTCTTTAT
M45	G→A	rs2032631	GGATATCAAAAATTGGCAGTGAAA GGCCTGGACCTCAGAAGGA	TAGATAaGCAAAAAGC	ATAGATAgGCAAAAAG
M207	A→G		CAACTATGGGGCAAATGTAAGTCA ACTTCTTTTGCCAATTAGGTCAC TTC	TCTCTTTcTAAAT TTC	TCTCTTTtTAAATTTCTTG
P25	C→CA	rs150173	GGGTAAAGTGAATTATCTGCCTGAA CATACATGATGGACCGAGATACGA	CTGCCTGaAAATAGAA	TGCCTGcAAATAGA
M170	A→C	rs2032597	GCTCTTATTAAGTTATGTTTTCATATTCTGTGC AACAGGTCCTCATTTTACAGTGAGACAC	CATTGTTCaTTTTTTTC	ATTGTTCcTTTTTTTCA
M172	T→G	rs2032604	GATTTTTTCATTTTATCCCCCAAAC AAATATCAGCCAGGTACAGAGAAAGTT	TTTGATGCTTTACTTAAAA	TGATGCTTGACTTAAA
M173	A→C	rs2032624	CAGTACTCACTTTAGGTTTGCCATATAAA CACTGGCTTATCATTTCTGAATATTAACA	CATTTAGAACACTTTGTCA	TTAGAACCCCTTTGTCATC

Tabela 2: Sequências dos *primers* e sondas utilizados para amplificação e detecção dos polimorfismo.

A PCR em Tempo Real foi realizada em três diferentes plataformas de PCR, para possível comparação de resultados, sendo estas, ABI PRISM 7000 (*Applied Biosystems*), LightCycler (*Roche*) e Smart Cycler (*Cepheid*).

Em todas elas a detecção dos SNPs foi realizada através da utilização de Sondas *TaqMan*, tendo-se usado um protocolo base de amplificação. Este inclui, para um volume final de 25µl de Master Mix (*Applied Biosystems*) 1x, *primer* forward e reverse a 900nM, sonda1 a 200nM e sonda2 a 100nM, e uma concentração de DNA de aproximadamente 10ng.

Este protocolo foi utilizado com bons resultados para a amplificação e detecção dos SNPs, tanto no ABI PRISM 7000 (*Applied Biosystems*), como no Smart Cycler (*Cepheid*), tendo sido necessário proceder a alguns ajustes na composição do tampão de amplificação para o Lightcycler®. Em relação a esta plataforma foi também efectuada a incorporação de albumina (concentração = 0,25mg/µl), de forma a obviar a adsorção das amostras a caracterizar pelos capilares de vidro. Esta modificação de protocolo mostrou ser eficiente.

Os parâmetros de amplificação encontram-se resumidos de seguida.

{	1 ciclo	50° durante 2 min
	1 ciclo	95° durante 10 min
	40 ciclos	95° durante 15 seg
		60° durante 1 min

IV – Resultados

Neste capítulo relatamos os resultados obtidos para as 33 amostras dos SNPs seleccionados, em três plataformas que possibilitam efectuar PCR em Tempo Real: ABI PRISM 7000, SmartCycler e o LightCycler®.

4.1 - ABI PRISM 7000

O primeiro protocolo de PCR a ser optimizado foi o realizado no ABI PRISM 7000, uma vez que esta é a plataforma disponível na Delegação do Porto do Instituto Nacional de Medicina Legal, e também, porque os *primers* e sondas utilizados, assim como a respectiva Mix de amplificação, são produtos da mesma empresa que comercializa o equipamento (*Applied Biosystems*).

Para a execução da amplificação foi utilizada o seguinte protocolo: TaqMan Universal PCR Master Mix (*Applied Biosystems*) 1x; 900nM de *primer* F; 900nM de *primer* R; 200nM sonda1; 100nM sonda2 e 10ng de DNA, para um volume final de 25µl. A Master Mix contém todos os restantes reagentes necessários para a ocorrência da PCR, AmpliTaq Gold DNA Polimerase, AmpErase UNG, dNTPs com dUTP, fluoróforo de referência (ROX) e tampão.

As condições da PCR foram as seguintes:

Passo 1 – 1 ciclo - 50°C durante 2 min

Passo 2 – 1 ciclo - 95°C durante 10 min

Passo 3 – 40 ciclos - 95°C durante 15 seg

- 60°C durante 1 min

Antes e após a realização deste protocolo de amplificação, no ABI PRISM 7000, é necessária a leitura dos níveis de fluorescência nos tubos utilizados para a reacção, utilizando-se um ciclo de 1 min a 60°C.

Os resultados obtidos podem ser visualizados sob a forma gráfica (Figura 5), podendo-se observar cada um dos alelos no eixo do x e do y, respectivamente. Assim, dividindo o gráfico em 4 quadrantes, e conhecendo o nosso controlo negativo, sabe-se quais as amostra correspondentes a cada alelo e quais as heterozigóticas (situação que não ocorre no nosso estudo, devido ao facto de se estar a trabalhar com SNPs do cromossoma Y).

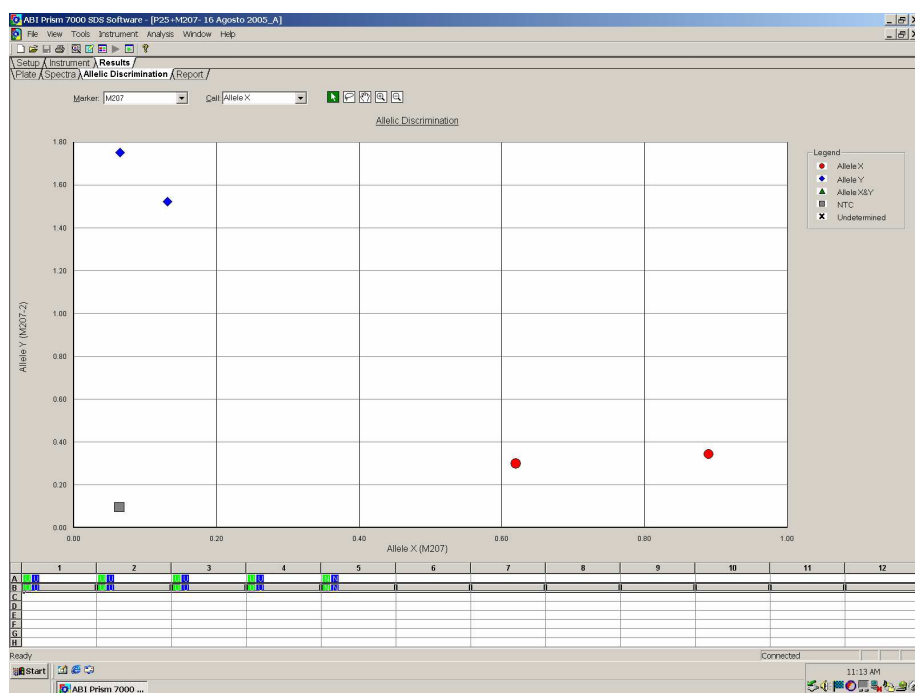


Figura 5: Visualização gráfica dos resultados obtidos no ABI PRISM 7000, podendo-se observar o controlo negativo e a distinção entre amostras homozigóticas para dois diferentes alelos (eixo dos X e dos Y).

A tipagem pode ainda ser confirmada recorrendo à observação das curvas de amplificação (Figura 6). A observação destas é de extrema importância, fornecendo grande ajuda em situações de dúvida. Neste caso, o que se observa é a variação da fluorescência para cada amostra em relação aos três fluoróforos: o FAM, o NED e o fluoróforo de referência, o ROX.

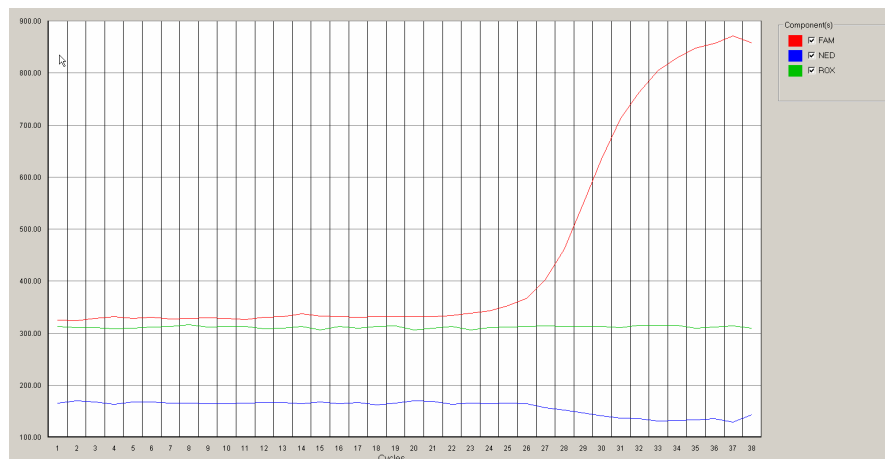


Figura 6: Representação gráfica das curvas de amplificação de uma amostra específica. Pode verificar-se que a referida amostra possuirá o alelo específico detectado pela sonda marcada com o fluoróforo FAM.

4.2 – SmartCycler

Para a amplificação no SmartCycler seguiu-se exactamente as mesmas condições, quer em relação às concentrações utilizadas, quer às condições da PCR. Os resultados obtidos são, também, apresentados de uma forma gráfica elucidativa (Figura 7), que se restringe, neste caso, a curvas de variação da fluorescência.

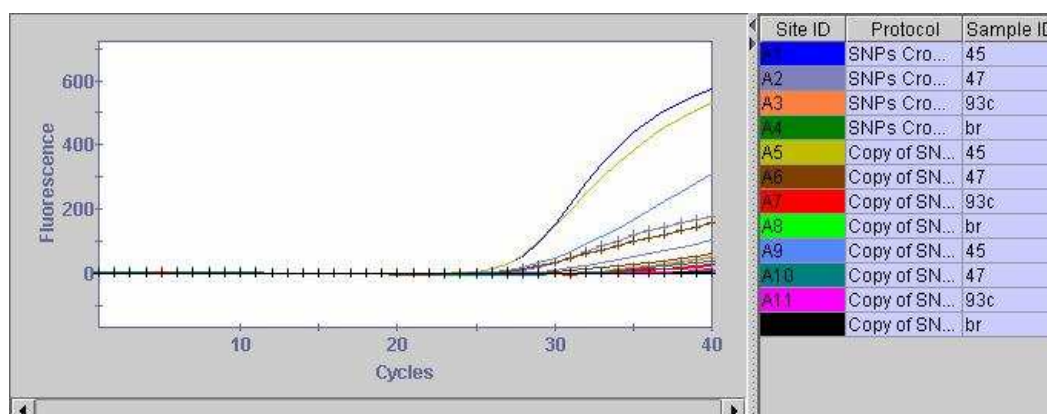


Figura 7: Apresentação gráfica dos resultados obtidos na plataforma SmartCycler, onde se pode observar a variação da fluorescência durante a reacção de PCR.

4.3 - LightCycler®

A realização da PCR no LightCycler® apresentou alguns problemas e não foi suficiente a transposição directa das condições anteriormente utilizadas. As primeiras experiências não mostravam a ocorrência de qualquer amplificação (Figura 8).

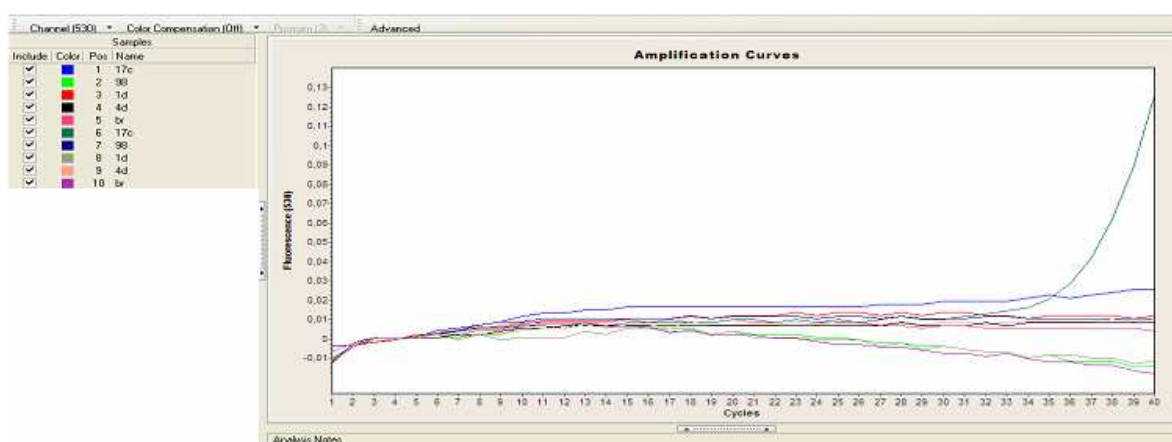


Figura 8: Ausência de amplificação durante a reacção de PCR na LightCycler, com excepção da amostra 1.

Depois de realizadas várias tentativas, com diferentes condições de PCR, verificou-se que o que estava a acontecer era o fenómeno já anteriormente referido em alguma bibliografia [11,22], e que irá ser sucintamente descrito.

Este aparelho funciona com a utilização de capilares de vidro, o que proporciona variações mais rápidas de temperatura, ideais para permitir a diminuição do tempo de cada PCR. No entanto, o vidro dos capilares adsorve, por vezes, a Taq DNA polimerase existente na mistura de amplificação (mix), não possibilitando que esta ocorra. Este contratempo foi ultrapassado, através da adição de albumina à mix, deste modo, os capilares a absorviam preferencialmente, concedendo a utilização da Taq presente na reacção. Por isso, passou-se a adicionar à mix de amplificação, numa concentração de 0,25mg/μl.

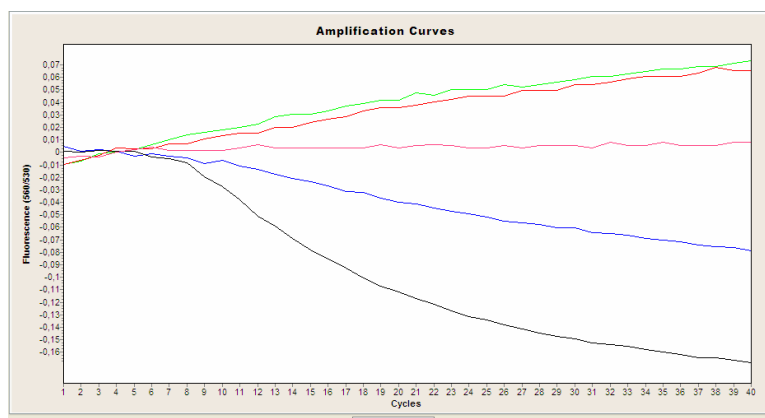


Figura 9: Amplificação durante a reacção de PCR no LightCycler, após a adição de albumina na mix de amplificação.

4.4 - Análise do SNP P25

Dos SNPs caracterizados, é importante evidenciar um deles, o P25, pois este SNP localiza-se numa região repetitiva do cromossoma, com 3 repetições, sendo que a variação nucleotídica só ocorre numa delas. Assim, aquando da amplificação iremos sempre obter sinal do alelo ancestral, C, mesmo quando existe também sinal do alelo A. A melhor forma de ultrapassar esta dificuldade de diferenciação entre as duas variantes alélicas, no ABI PRISM 7000, é através da observação das curvas de amplificação, conseguindo-se, desta forma, efectuar uma classificação inequívoca destas duas variantes.

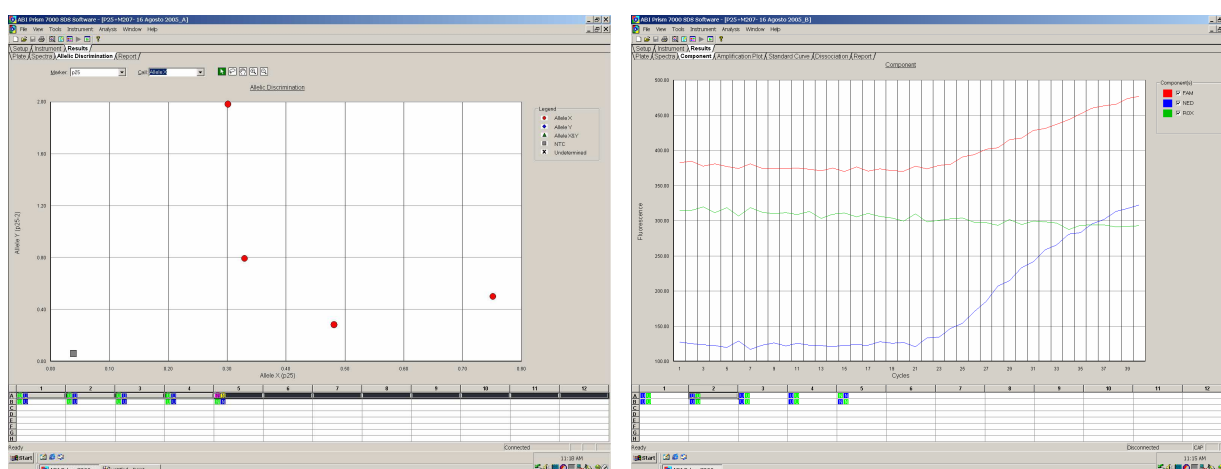


Figura 10: Representação gráfica dos resultados obtidos na genotipagem do SNP P25 no ABI PRISM 7000.

Maior facilidade de interpretação através da observação das curvas de amplificação.

A interpretação dos resultados obtidos, para este mesmo SNP no LightCycler®, não é fácil, podendo originar alguns erros. Torna-se complicada a distinção entre as duas variantes alélicas, sendo que apenas uma análise cuidada da variação das fluorescências permitirá efectuar a genotipagem correcta.

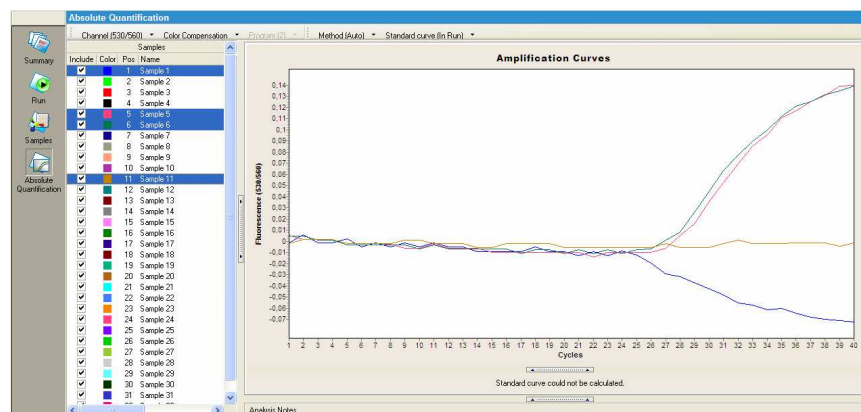


Figura 11: Variação das curvas de fluorescência (530/560) para o SNP P25. As amostras 5 e 6 apresentam a variante alélica A, a amostra 11 o alelo C, mas a genotipagem da amostra 1 é duvidosa.

A dificuldade encontrada para a genotipagem do SNP P25, devido às suas características especiais, foi mais facilmente ultrapassada com a utilização do SmartCycler. Nesta plataforma os resultados foram obtidos através da variação da fluorescência, seleccionando a opção que permite observar a variação dos dois fluoróforos em simultâneo, podendo-se fazer uma tipagem segura (Figura 12).

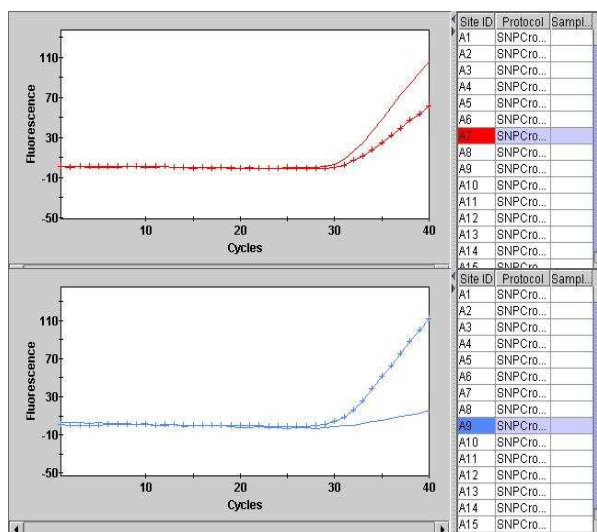


Figura 12: Variações da fluorescência para os dois fluoróforos que marcam as duas variantes alélicas para duas amostras com diferente genotipagem.

Utilizando esta plataforma, é possível fazer a análise da tabela com os dados da PCR, mais propriamente dos valores de Ct (*Cycle Threshold*). O Ct é o valor correspondente ao ciclo, a partir do qual, durante a reacção de amplificação, é detectada fluorescência acima da linha de base. Este é um valor muito importante nestas abordagem de 5' nucleasase.

Trata-se de outra possibilidade de análise dos resultados, e que mostrou ser a mais correcta e exacta para a genotipagem do SNP P25.

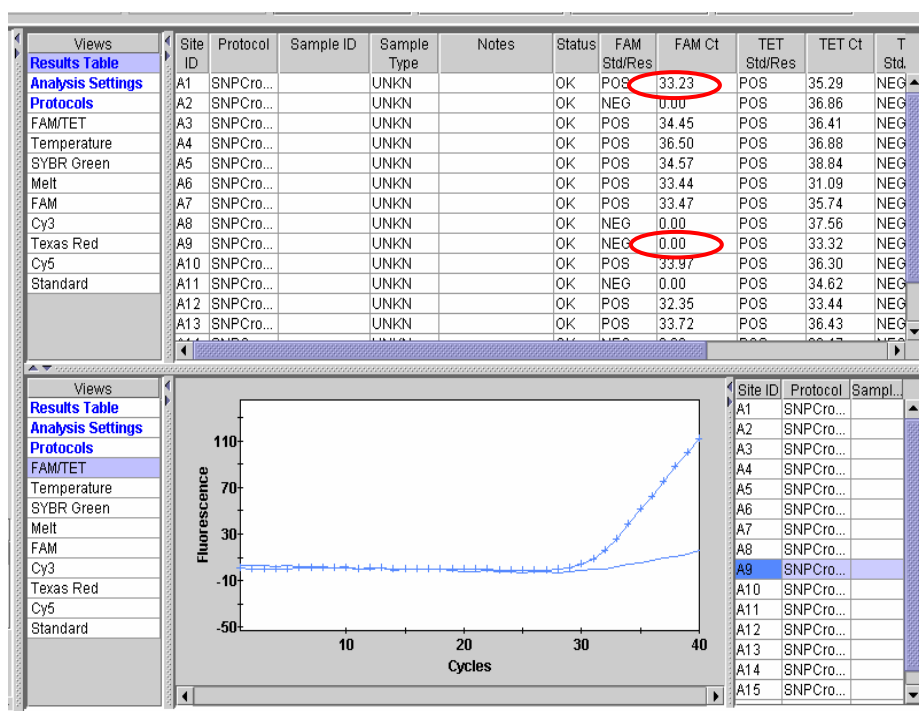


Figura 13: Apresentação gráfica dos resultados e dados da PCR apresentados pelo SmartCycler. Facilidade de análise dos Ct para cada fluoróforo, e consequentemente, genotipagem de cada amostra.

4.5 - Resultados obtidos utilizando as três plataformas

As genotipagens obtidas para as 33 amostras utilizadas, correspondentes a 33 indivíduos, e após a realização das PCRs para os 8 SNPs em estudo, encontram-se resumidas na tabela 3.

A sua observação permite verificar que, de um modo geral, o método utilizado é reprodutível nas três plataformas (Tabela 3). Alguns dos SNPs são mais difíceis de genotipar, originando, consequentemente, um maior número de tipagens duvidosas, sendo que estas dúvidas surgem com maior frequência no LightCycler® e no ABI PRISM 7000. Podemos constatar que na primeira plataforma, as incertezas de genotipagem foram de 13,0% para o SNP M45, 9,6% para o P25 e 3,2% para o M172 e para o M173. Em relação ao ABI PRISM 7000, para os SNPs M9, M45, P25 e M170 também surgiram tipagens duvidosas, mas com menor frequência, sendo estas de, respectivamente, 12,5%, 9,6%, 6,0% e 3,0% .

Analisando os resultados destas duas plataformas, verificou-se que dois dos polimorfismo caracterizados, o M45 e o P25, apresentam leituras que suscitam dúvidas com maior frequência.

Importa evidenciar que, em todas as genotipagens realizadas na plataforma Smart Cycler, não ocorreu qualquer dúvida na caracterização da totalidade dos SNPs nas 33 amostras. Pode, ainda, observar-se que a tipagem de dois dos SNPs (M172 e M173) não foi possível no ABI PRISM 7000, pois não permitiu a distinção dos possíveis genótipos. Este problema, provavelmente, ocorreu devido a características específicas dos *primers* utilizados e poderia talvez ter sido ultrapassado usando *primers* diferentes ou outras diluições em relação aos que estavam a ser usados ou, ainda, haver necessidade da optimização do próprio protocolo.

4.6 - Análise de sensibilidade da metodologia

Para testar a sensibilidade da metodologia utilizada nas diferentes plataformas, foi realizada a genotipagem de dois SNPs, o M45 e o M207, em misturas com diferentes proporções

de DNA masculino e feminino (1:1, 1:50, 1:150, 1:300) e, também, em amostras de DNA masculino (e DNA padrão) em distintas concentrações (25 – 0,5 – 0,165 – 0,08 µg/µl).

Amostra	AB - DNA		SC - DNA		LC - DNA		AB - Padrão		SC - Padrão		LC - Padrão	
	M45	M207	M45	M207	M45	M207	M45	M207	M45	M207	M45	M207
1:1	V	V	V	V	V	V	V	V	---	V	V	---
1:50	V	V	---	---	---	V	---	V	---	---	---	---
1:150	V	V	---	---	---	---	---	---	---	---	---	---
1:300	V	---	---	---	---	---	---	---	---	---	---	---
25 µg/µl	V	V	V	V	V	V	V	V	---	V	V	---
0,5 µg/µl	---	---	---	---	---	---	V	V	---	---	V	---
0,165 µg/µl	---	V	---	---	---	---	V	V	---	---	---	---
0,08 µg/µl	---	---	---	---	---	---	V	V	---	---	---	---

Tabela 4: Genotipagem dos SNPs M45 e M207 em misturas com diferentes proporções de DNA masculino e feminino, para diferentes concentrações de DNA masculino, nas distintas plataformas. Legenda: AB - ABI PRISM 7000; SC – SmartCycler; LC - LightCycler®; DNA – uma das 33 utilizadas; Padrão – amostra de DNA padrão; V – genotipagem possível; — - sem possibilidade de genotipagem.

A análise da Tabela 4 permite deduzir que foi possível a caracterização dos SNPs em misturas de DNA masculino e feminino em proporções de 1:1 até 1:50 parecendo, no entanto, que esta sensibilidade do método varia quer em relação ao SNP específico em análise (melhor detecção no caso do M207), quer nas diferentes plataformas (melhor desempenho do ABI PRISM 7000). O mesmo se observa relativamente às distintas concentrações de DNA, pois o ABI PRISM 7000 mostra maior sensibilidade na detecção do SNP para menores concentrações de DNA.

Observa-se, ainda, que a qualidade do DNA utilizado poderá também influenciar a sensibilidade do método. A detecção dos SNPs foi mais eficaz quando foi utilizada a amostra de DNA padrão.

Importa salientar que estes resultados deveriam ser confirmados em, pelo menos, mais duas experiências idênticas à realizada. Deste modo, não subsistiria a dúvida de que alguns dos resultados observados foram fruto do acaso, ou de qualquer problema pontual.

4.7 - Aproximação a estudo populacional

Com a genotipagem de apenas 33 amostras, correspondentes a 33 indivíduos, efectuou-se a determinação da frequência dos haplogrupos, evidenciando-se o facto do número de amostras para um estudo deste tipo ser manifestamente insuficiente.

Os 8 SNPs analisados nas 33 amostras permitiram a definição de 5 haplogrupos, seguindo a nomenclatura sugerida pelo YCC [1].

A maioria das amostras pertencem ao haplogrupo R1b (61%), tendo em consideração o estudo do SNP P25. Os outros haplogrupos representados, nesta pequena amostragem, são o E3a (3%), o I (9%), o J2 (24%) e o R1 (3%) (Tabela 5 e Figura 14).

População	M2	M170	M172	M9	M45	M207	M173	P25	Haplogrupo	N
Norte de Portugal	+	-	-	-	-	-	-	-	E3a	1
	-	+	-	-	-	-	-	-	I	3
	-	-	+	-	-	-	-	-	J2	8
	-	-	-	+	+	+	-	-	R1	1
	-	-	-	+	+	-	+	+	R1b	20

Tabela 5: Haplogrupos resultantes da genotipagem de 8 Y-SNPs numa amostra de 33 indivíduos do Norte de Portugal. Legenda: + - variante mutante; ¾ - variante ancestral.

Am	M9			M9	M2			M2	M45			M45	M207			M207	P25			P25	M170			M170	M172			M172	M173			M173
	AB	LC	SC	Geral	AB	LC	SC	Geral	AB	LC	SC	Geral	AB	LC	SC	Geral	AB	LC	SC	Geral	AB	LC	SC	Geral	AB	LC	SC	Geral	AB	LC	SC	Geral
42	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	C	C	C	C	A	A	A	A		T	T	T		C	C	C
43	G	G	G	G	A		A	A	A		A	A	G	G		G	A			A	A			A								
45	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G		C	C	C
47	C	C	C	C	A		A	A		G		G	A	A		A	A			A	C			C								
48	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
49	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
17R	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
18R	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
92	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
94	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
95	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
96	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
98	C	C	C	C	A	A	A	A	A	G	A	G	G	A	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A
7c	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
8c	G	G	G	G	A	A	A	A	A	A	A	A	G	G		G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
10c	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	T	T	T		C	C	C	
16c	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
17c	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	T	T	T		C	C	C	
23c	C	C	C	C	A	A	A	A		A	A	A	A	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
24c	C	C	C	C	A	A	A	A		A	A	A	A	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
33c	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
43c	C	C	C	C	A	A	A	A	A	A	A	G	A	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
76c	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G		C	C	C	
91c	G	G	G	G	A	A	A	A	A	A	A	A	G	G		G	C	A	A	A	A	A	A	A	G	G	G		C	C	C	
93c					A		A	A		G	A	G	G	A	A	A	C	C	C	C	C	C	C	C	G	G	G		A	A	A	
97c	C	C	C	C	A	A	A	A		A	A	G	G	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
1d	C	C	C	C	G	G	G	G		A	A	G	G	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
2d	C	C	C	C	A	A	A	A			A	G	G	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
3d	C	C	C	C	A	A	A	A		G	A	G	G	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
4d	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A		A	A	G	G	G		C	C	C	
5d	C	C	C	C	A	A	A	A	A	G	A	G	G	A	A	A	C	C	C	C	A	A	A	A	G	G	G		A	A	A	
6d	G	G	G	G	A	A	A	A	A	A	A	A	G	G	G	G	A	A	A	A	A		A	A	G	G	G		C	C	C	
9d	C	C	C	C		A	A	A		G	A	G	G	A	A	A	C	C	C	C	C		C	C	G	G	G		A	A	A	

Tabela 3: Genotipagens nas diferentes plataformas das 33 amostras utilizadas. Legenda: Am – amostras; : AB - ABI PRISM 7000; SC – SmartCycler; LC - LightCycler®; ■ - dúvidas de genotipagem; ■ - genotipagens não realizadas



Figura 14: Árvore filogenética (modificada YCC 2002) definida com os 8 polimorfismos bialélicos analisados. Os haplogrupos designados à direita, representam todos os que são possíveis de definir com estes 8 SNPs. A distribuição dos diferentes haplogrupos encontra-se representada à direita da árvore.

V - Discussão

Nos últimos anos, a utilização de marcadores bialélicos, nomeadamente de SNPs, na área forense, tem vindo a sofrer um grande incremento.

A vantagem da análise do cromossoma Y em Genética e Biologia Forense, baseia-se no facto de este se encontrar unicamente em indivíduos do sexo masculino. Assim, neste âmbito o estudo deste cromossoma é utilizado, fundamentalmente, na resolução de casos criminais, em especial em agressões sexuais e em investigação de filiação, em que os descendentes são do sexo masculino e não se dispõe de amostras biológicas do pretenso pai, mas em que se pode recorrer a familiares com ele aparentados pela linha paterna.

Actualmente, são várias as tecnologias disponíveis para a genotipagem de SNPs, e mesmo as existentes encontram-se em constante evolução. A escolha da melhor tecnologia a utilizar para diferentes aplicações específicas irá depender dos objectivos finais; isto é, do número de marcadores a caracterizar em simultâneo e do número de amostras a analisar. No que diz respeito à área da Genética Forense, é certo que seria interessante o uso das diferentes tecnologias, cada uma delas apropriada a determinada situação da rotina laboratorial. Ou na resolução de casos criminais, em que a quantidade e a qualidade do DNA são um ponto crítico, ou na conclusão de casos de filiação biológica e, também, para a construção de bases de dados genéticas com interesse criminal. Assim, são vários os aspectos a ter em conta na sua área forense, tais como a sensibilidade do método, a reprodutibilidade, a precisão, a capacidade para a realização de PCR *multiplex* e as suas possibilidades em relação ao número de amostras e de SNPs a genotipar. É ainda importante não esquecer a flexibilidade permitida por cada tecnologia, o tempo necessário a despende e os custos envolvidos, considerando quer a plataforma a utilizar, quer o preço da genotipagem. Um importante factor limitante no campo da Genética Forense, e que afectará qualquer método e tecnologia seleccionada, é a quantidade e a qualidade do DNA disponível.

A estratégia para a genotipagem de SNPs seguida neste trabalho, a PCR em Tempo Real com utilização de sondas TaqMan, mostrou ser possível nas três plataformas em ensaio. Foi viável a elaboração de um protocolo geral, quer relativamente à constituição da mistura de amplificação, quer às condições de PCR utilizadas, aplicável nos três aparelhos de PCR em Tempo Real. A realização da PCR no ABI PRISM 7000 e no SmartCycler foi viável com a utilização das mesmas condições a todos os níveis, com resultados positivos.

A execução da PCR no LightCycler[®] pressupõe o ajuste das condições normalmente usadas. Nesta plataforma a mistura de amplificação é colocada em capilares de vidro. Esta particularidade permite variações mais rápidas de temperatura, ideais para uma diminuição no tempo global de cada PCR. No entanto, o vidro dos capilares, por vezes, adsorve a Taq DNA polimerase existente na mistura de amplificação, não permitindo assim que esta seja usada na extensão. Esta contrariedade foi ultrapassada através da adição de albumina à referida mistura, sendo que os capilares, deste modo, irão absorver preferencialmente esta, permitindo a utilização da Taq polimerase presente na reacção. Esta modificação encontrava-se já descrita noutros trabalhos, em que se utilizava a mesma mistura de amplificação (TaqMan Universal PCR Master Mix 1x - *Applied Biosystems*) e também recorrendo ao LightCycler[®] para realizar a PCR [11,22]. Esta alteração de protocolo mostrou ser eficiente no desenrolar da amplificação e, consequente, na caracterização dos SNPs.

A metodologia mostrou ser suficientemente robusta para permitir a sua aplicação em diferentes aparelhos de PCR em Tempo Real.

A análise da tabela referente às genotipagens efectuadas, nas diferentes plataformas, permite avaliar a metodologia, tendo em conta determinadas características, tais como a precisão e a especificidade. Efectuando a análise de cada polimorfismo, considerados individualmente, pode observar-se que as genotipagens efectuadas são, de um modo geral, concordantes nas três plataformas, revelando a precisão e especificidade da técnica de PCR em

Tempo Real. Surgem, no entanto, situações pontuais, como aquela que se observa para o M45, em que a genotipagem não é concordante. Estas situações poderão ocorrer devido a características específicas da sequência da região a amplificar ou dos próprios *primers* utilizados, que façam com que a amplificação não ocorra tão eficazmente, levando a erros de genotipagem. Podem ser, por exemplo, *primers* menos estáveis ou regiões de amplificação com sequências nucleótídicas complicadas para o desenho de *primers* e de sondas com as melhores características. Estas dificuldades poderão ser ultrapassadas com a confirmação de genotipagem, através da realização de nova PCR, com a correcta conservação dos *primers* utilizados e com a efectuação de uma genotipagem cuidadosa.

Observando ainda a mesma tabela (tabela 3), verifica-se que as dificuldades de genotipagem foram pouco frequentes. Estas ocorreram, maioritariamente, no ABI PRISM 7000 e no LightCycler®. Em relação às genotipagens efectuadas no SmartCycler, não foram encontradas estas ambiguidades. A explicação para este facto poderá passar pela própria forma de apresentação dos resultados finais de cada aparelho, o que permite uma melhor ou mais deficiente interpretação dos resultados.

Relativamente à sensibilidade do método, esta foi testada para diferentes concentrações de DNA masculino, para DNAs com diferente qualidade e, também, para misturas de DNA feminino e masculino (tabela 4).

Em relação à concentração, foi possível a detecção de SNPs para concentrações de DNA até 8ng/μl. Todas as genotipagens efectuadas neste trabalho, para os diferentes polimorfismo e nas três plataformas, foram realizadas usando DNA com concentrações de aproximadamente 10ng, possibilitando a detecção dos vários SNPs sem dificuldades, encontrando-se descrita por outros autores, a caracterização de SNPs utilizando concentrações de DNA de 1ng, [39] recorrendo à minisequenciação. Usando esta mesma metodologia, outros autores, referem que com a PCR *multiplex*, são caracterizados SNPs com concentrações de

DNA entre 50pg-1ng [24]. Para a metodologia da PCR em Tempo Real, em LightCycler® e com a utilização de sondas FRET, Lareu *et al.* [17] descrevem a detecção de SNPs com concentrações de DNA de 100pg, evidenciando a elevada sensibilidade da técnica.

Ainda relativamente à concentração de DNA necessária para a detecção de um SNP, pode observar-se, neste trabalho, que o ABI PRISM 7000 foi a plataforma com melhor desempenho, proporcionando resultados com menores concentrações de DNA. Não podendo esquecer o carácter apenas informativo desta experiência, seria necessário um maior número de repetições da mesma, para eliminar quaisquer dúvidas. A melhor capacidade de resposta do ABI PRISM 7000, perante esta aparente dificuldade, poderá ser explicada pelo facto do processo ser mais lento na realização da PCR. Esta característica concede um maior período de actuação da Taq DNA polimerase, permitindo a detecção de menores quantidades de DNA.

Uma possível vantagem da utilização de SNPs, quando comparada com a utilização dos tradicionais STRs, relativamente à sua aplicação na Genética Forense, consiste precisamente na sua maior sensibilidade, possibilitando a caracterização de DNA masculino em misturas de DNA masculino-feminino.

Para a avaliação desta vantagem, usando a metodologia e os SNPs referidos neste estudo, foram realizadas 4 misturas, com diferentes proporções de DNA masculino-feminino (1:1; 1:50; 1:150; 1:300). Foi possível a detecção de SNPs para proporções de 1:300 (masculino-feminino) ocorrendo, contudo, uma maior conformidade nos resultados das três plataformas para proporção de 1:50. Também nesta situação de análise de misturas, a plataforma que apresentou melhores resultados foi o ABI PRISM 7000. A maior sensibilidade devido à menor rapidez do processo, pode continuar a ser a explicação possível.

No trabalho já referido anteriormente por Lareu *et al.* [17], em que utilizam sondas FRET no LightCycler®, foi efectuada esta mesma avaliação de sensibilidade. Com a metodologia

utilizada descrevem a detecção de misturas nas mesmas proporções das conseguidas neste trabalho.

Os resultados obtidos seriam, provavelmente, mais variáveis se as misturas utilizadas fossem de misturas de fluidos masculino-feminino, e não de misturas de DNAs padrão.

Relativamente à importância da qualidade de DNA na análise destes polimorfismos, a experiência realizada neste trabalho não permite obter informação conclusiva. A amostra de DNA utilizada foi extraída com Chelex, e os resultados obtidos para esta amostra não se diferenciam dos observados para a amostra padrão. Sabe-se, no entanto, que a maior facilidade na análise destes polimorfismos em amostras de DNA antigas e degradadas, é uma das vantagens em relação à tradicional análise de STRs [9].

Para além dos aspectos técnicos a evidenciar neste trabalho, importa também referir algumas características especiais relacionadas com alguns dos polimorfismos analisados, neste caso, o P25.

Diferentes SNPs do cromossoma Y podem definir um mesmo haplogrupo; por isso, é importante uma selecção dos mais apropriados. As bases de dados de SNPs são cruciais nesta selecção, pois será importante ter em conta um dado recentemente descrito que sugere que 40% do cromossoma Y consiste em duplicações inter e intra cromossomais, com 90-98% de sequências idênticas [5]. Diferenças de um nucleótido entre nucleótidos da mesma posição de regiões duplicadas, são erradas e frequentemente descritas como SNPs. Estima-se que cerca de 10% dos SNPs referenciados nas bases de dados, consistem de facto em variações em sequências parálogas, [27]. Trata-se de sequências com origem em fenómenos de duplicação que podem ter até 10 kilobases de comprimento, com uma homologia de cerca de 95% (PSVs – *Paralogous Sequence Variants*). Uma propriedade importante das sequências parálogas é o facto de estas serem muito propensas à conversão génica; ou seja, a informação contida na sequência duplicada não é feita de modo recíproco de uma cópia para a outra, levando à

ocorrência de diferenças em nucleótidos na mesma posição. Estas diferenças não serão SNPs, mas surgem devido a este fenómeno [2]. Por isso, os SNPs a utilizar em Genética Forense, deverão ser aqueles que ocorrem em posições únicas do cromossoma.

O P25 é um SNP do cromossoma Y descrito como uma variação C->A na região q11.223 [12], muito utilizado em estudos populacionais, e que define o haplogrupo R1b, um dos mais frequentes em populações da Europa Ocidental. Existem três cópias da sequência do P25 na região AZFc do cromossoma Y, a mais distal das quais é a g3, que contém o alelo A derivado, enquanto as mais proximais, a g1 e a g2, possuem o alelo C ancestral.

Esta ocorrência, leva, em termos práticos a que, quando está presente a variante derivada A, surja sempre sinal da variante ancestral C. Este sinal é, no entanto, de menor intensidade relativamente ao sinal obtido para o alelo A. Esta circunstância é devida à homologia de apenas 98% entre os *primers* utilizados e uma das sequências parálogas confirmadas por Blast. Esta ausência de homologia total, resulta numa menor eficácia da acção da Taq DNA Polimerase e, consequentemente, a uma amplificação menos eficiente.

Na genotipagem deste SNP, nas amostras utilizadas neste trabalho, surgiram algumas dificuldades, essencialmente para as plataformas ABI PRISM 7000 e LightCycler®. Para a primeira, as dificuldades foram solucionadas através da análise das respectivas curvas de amplificação. Já para a segunda, a diferenciação alélica não foi possível sendo, por isso, necessário efectuar uma genotipagem mais cuidadosa. A caracterização deste mesmo SNP no SmartCycler fez-se sem problemas, quer através da análise das curvas de variação de fluorescência quer, de um modo mais seguro e exacto, a partir dos valores de Ct atribuídos na tabela com os dados da corrida da PCR, que é uma alternativa do *software* de análise ligado a esta plataforma.

Importa ainda referir que se encontram também descrito, para este polimorfismo, a ocorrência de mutações recorrentes ou “*back-mutations*”, como resultado de, pelo menos, dois fenómenos de conversão génica nas sequências palindrómicas [2].

Assim, em virtude do que foi referido em relação a este polimorfismo, tem sido recomendado, tanto a nível de estudos populacionais, como na área forense, análise suplementares ou a substituição deste polimorfismo pelo SNP M269 [2].

Com o objectivo de validar um protocolo para a caracterização de SNPs pela técnica de PCR em Tempo Real, foram caracterizadas neste estudo 33 amostras de indivíduos do Norte de Portugal relativamente a 8 SNPs do cromossoma Y.

Esta análise permitiu a obtenção de resultados que podem constituir uma aproximação a um futuro estudo populacional. Na amostra estudada observou-se um total de 5 haplogrupos definidos de acordo com a árvore do YCC [1]. Das 33 amostras estudadas, 20 pertencem a um dos haplogrupos mais frequentes das populações europeias, o R1b. Das amostras restantes, 8 pertencem ao haplogrupo J2, 3 ao haplogrupo I e as duas restantes distribuem-se pelos haplogrupos R1 e E3a. Apesar destes serem haplogrupos muito frequentes em populações europeias, não são exclusivos destas. O haplogrupo R1 também tem sido encontrado no Leste Europeu, na Ásia Central e no Norte da Índia e, ainda, como resultado do fluxo de genes durante a época colonial, na Austrália e na América. O haplogrupo J2 caracteriza também populações da Ásia Central e do Sul da Índia [33]. Pode ser observada na tabela 6, uma comparação dos resultados deste estudo com os obtidos para outras populações provenientes de diferentes áreas geográficas.

A maior frequência encontrada neste estudo para o haplogrupo R1b, está em concordância com o já descrito, para uma população portuguesa com a mesma proveniência, por Beleza *et al.* [6].

Haplogrupo YCC	Portugal (N=33)	Europa (N=68)	Leste Asiático (N=114)	Norte de África (N=52)	Americanos Brancos (N=114)
R1	1	25			13
R1a			1	5	6
R1b	20		3	3	54
E3a	1			1	
E3b		15	1	13	3
J2	8	14		10	6
I	3	3			24
F*		7	6	11	
G				5	3
K*		4			1
P*				2	
A-E					
P*					

Tabela 6: Resultados da genotipagem de Y-SNP em populações com diferente proveniência geográfica

[3,35,39].

Este mesmo resultado foi obtido em populações europeias relativamente próximas da portuguesa, como a Italiana e a Espanhola [6-8,23,24].

Uma diferença encontrada entre os estudos anteriormente referidos e este, e que importa salientar, é a ausência do haplogrupo E3b no nosso estudo, uma vez que este é descrito como sendo o segundo mais frequente em populações europeias. Esta ausência será apenas devida ao acaso, resultante de um enviesamento provocado pelo reduzido número de amostras utilizado. O prolongamento deste estudo, que consistiu na análise de 88 indivíduos, efectuado *a posteriori*, é indicativo de que na realidade o haplogrupo E3b é o segundo mais frequente [25]. Por isso, poder-se-á inferir que o aumento da amostragem, assim como do número de SNPs a analisar, permitirá uma melhor caracterização desta população.

VI - Conclusão

O trabalho realizado permite concluir que:

O uso da PCR em Tempo Real é um bom método para a caracterização de SNPs, com utilidade na resolução de casos forenses. As suas principais vantagens são o tempo reduzido para a realização da análise, ser uma técnica reprodutível, precisa e específica, obviando as dificuldades inerentes ao estabelecimento das condições ideais de amplificação e análise dos produtos amplificados de vários polimorfismo em simultâneo.

A rapidez e a capacidade de se processar, em paralelo, um elevado número de amostras, bem como de executar, em simultâneo ensaios distintos, permite ultrapassar a aparente desvantagem da impossibilidade de se analisar concomitantemente diversos SNPs. A principal desvantagem prende-se apenas com a dificuldade da realização da PCR *multiplex*.

Quanto ao desempenho das três plataformas usadas neste trabalho, o SmartCycler foi a que demonstrou maior facilidade na caracterização dos SNPs analisados, bem como exactidão, não havendo necessidade de repetição de ensaios. Esta plataforma, apesar da desvantagem do pequeno número de amostras possível de genotipar num só ensaio, foi a que apresentou resultados mais fiáveis.

Em relação ao ABI PRISM 7000, apresentando esta plataforma boas características para ser utilizada em situações onde se torna necessária a genotipagem de um elevado número de amostras, demonstrou ainda ser bastante eficiente na detecção de SNPs em situações com DNA de baixa concentração e qualidade. As suas principais desvantagens serão talvez a mais baixa flexibilidade em termos de protocolos a utilizar, e o tempo despendido na sua realização.

Relativamente ao LightCycler[®], os resultados obtidos neste trabalho são demonstrativos das limitações desta plataforma, tendo em vista o objectivo pretendido.

O estudo destes polimorfismo bialélicos do cromossoma Y, pode ser importante especialmente na análise de amostras relacionadas com agressões sexuais. Esta importância reside no facto da fracção masculina, cuja identificação genética é imprescindível neste tipo de

crimes, na generalidade das situações, estar em pequena quantidade, constituindo a menor fracção quando presente em misturas com fracção feminina, circunstância que ocorre na maioria dos casos.

A caracterização de marcadores bialélicos do cromossoma Y deverá estender-se a um maior número de polimorfismos e, em simultâneo, aos dos cromossomas autossómicos, funcionando como importante informação adicional. Recentemente, têm sido também publicados estudos relativamente à caracterização de SNPs do DNA mitocondrial, cujo principal objectivo é distinguir indivíduos que pertençam ao mesmo haplogrupo [10,16,36].

VII – Referências Bibliográficas

- [1] A nomenclature system for the tree of human Y-chromosomal binary haplogroups, *Genome Res* 12 (2002) 339-348.
- [2] S.M. Adams, T.E. King, E. Bosch and M.A. Jobling, The case of the unreliable SNP: recurrent back-mutation of Y-chromosomal marker P25 through gene conversion, *Forensic Sci Int* 159 (2006) 14-20.
- [3] F. Alessandrini, C. Turchi, V. Onofri, L. Buscemi, M. Pesaresi and A. Tagliabracci, Multiplex PCR development of Y-chromosomal biallelic polymorphisms for forensic application, *J Forensic Sci* 50 (2005) 519-525.
- [4] A. Amorim and L. Pereira, Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs, *Forensic Sci Int* 150 (2005) 17-21.
- [5] J.A. Bailey, Z. Gu, R.A. Clark, K. Reinert, R.V. Samonte, S. Schwartz, M.D. Adams, E.W. Myers, P.W. Li and E.E. Eichler, Recent segmental duplications in the human genome, *Science* 297 (2002) 1003-1007.
- [6] S. Beleza, L. Gusmao, A. Lopes, C. Alves, I. Gomes, M. Giouzeli, F. Calafell, A. Carracedo and A. Amorim, Micro-phylogeographic and demographic history of Portuguese male lineages, *Ann Hum Genet* 70 (2006) 181-194.
- [7] M. Brion, B.M. Dupuy, M. Heinrich, C. Hohoff, B. Hoste, B. Ludes, B. Mevag, N. Morling, H. Niederstatter, W. Parson, J. Sanchez, K. Bender, N. Siebert, C. Thacker, C. Vide and A. Carracedo, A collaborative study of the EDNAP group regarding Y-chromosome binary polymorphism analysis, *Forensic Sci Int* 153 (2005) 103-108.
- [8] M. Brion, B. Sobrino, A. Blanco-Verea, M.V. Lareu and A. Carracedo, Hierarchical analysis of 30 Y-chromosome SNPs in European populations, *Int J Legal Med* 119 (2005) 10-15.
- [9] J.M. Butler, *Forensic DNA Typing: Biology, Technology and Genetics of STR Markers*, 2nd Edition edn., Elsevier Academic Press, Burlington, 2005, 660 pp.
- [10] M.D. Coble, R.S. Just, J.E. O'Callaghan, I.H. Letmanyi, C.T. Peterson, J.A. Irwin and T.J. Parsons, Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians, *Int J Legal Med* 118 (2004) 137-146.
- [11] M.J. Espy, J.R. Uhl, P.S. Mitchell, J.N. Thorvilson, K.A. Svien, A.D. Wold and T.F. Smith, Diagnosis of herpes simplex virus infections in the clinical laboratory by LightCycler PCR, *J Clin Microbiol* 38 (2000) 795-799.
- [12] M.F. Hammer, T.M. Karafet, A.J. Redd, H. Jarjanazi, S. Santachiara-Benerecetti, H. Soodyall and S.L. Zegura, Hierarchical patterns of global human Y-chromosome diversity, *Mol Biol Evol* 18 (2001) 1189-1203.
- [13] S. Inagaki, Y. Yamamoto, Y. Doi, T. Takata, T. Ishikawa, K. Imabayashi, K. Yoshitome, S. Miyaishi and H. Ishizu, A new 39-plex analysis method for SNPs including 15 blood group loci, *Forensic Sci Int* 144 (2004) 45-57.
- [14] A.J. Jeffreys, V. Wilson and S.L. Thein, Individual-specific 'fingerprints' of human DNA, *Nature* 316 (1985) 76-79.
- [15] M.A. Jobling, A. Pandya and C. Tyler-Smith, The Y chromosome in forensic analysis and paternity testing, *Int J Legal Med* 110 (1997) 118-124.
- [16] R.S. Just, J.A. Irwin, J.E. O'Callaghan, J.L. Saunier, M.D. Coble, P.M. Vallone, J.M. Butler, S.M. Barritt and T.J. Parsons, Toward increased utility of mtDNA in forensic identifications, *Forensic Sci Int* 146 Suppl (2004) S147-149.

- [17] M. Lareu, J. Puente, B. Sobrino, B. Quintans, M. Brion and A. Carracedo, The use of the LightCycler for the detection of Y chromosome SNPs, *Forensic Sci Int* 118 (2001) 163-168.
- [18] P. Malaspina, F. Cruciani, B.M. Ciminelli, L. Terrenato, P. Santolamazza, A. Alonso, J. Banyko, R. Brdicka, O. Garcia, C. Gaudiano, G. Guanti, K.K. Kidd, J. Lavinha, M. Avila, P. Mandich, P. Moral, R. Qamar, S.Q. Mehdi, A. Ragusa, G. Stefanescu, M. Caraghin, C. Tyler-Smith, R. Scozzari and A. Novelletto, Network analyses of Y-chromosomal types in Europe, northern Africa, and western Asia reveal specific patterns of geographic distribution, *Am J Hum Genet* 63 (1998) 847-860.
- [19] C.D. Mamotte, Genotyping of single nucleotide substitutions, *Clin Biochem Rev* 27 (2006) 63-75.
- [20] T. Maniatis, E.F. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Third Edition edn., Cold Spring Harbor Laboratory, Melbourne, 1982.
- [21] J. Mengel-Jorgensen, J.J. Sanchez, C. Borsting, F. Kirpekar and N. Morling, MALDI-TOF mass spectrometric detection of multiplex single base extended primers. A study of 17 y-chromosome single-nucleotide polymorphisms, *Anal Chem* 76 (2004) 6039-6045.
- [22] A. Nitsche, N. Steuer, C.A. Schmidt, O. Landt and W. Siebert, Different real-time PCR formats compared for the quantitative detection of human cytomegalovirus DNA, *Clin Chem* 45 (1999) 1932-1937.
- [23] V. Onofri, F. Alessandrini, C. Turchi, B. Fraternali, L. Buscemi, M. Pesaresi and A. Tagliabracci, Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis, *Int J Legal Med* (2007).
- [24] V. Onofri, F. Alessandrini, C. Turchi, M. Pesaresi, L. Buscemi and A. Tagliabracci, Development of multiplex PCRs for evolutionary and forensic applications of 37 human Y chromosome SNPs, *Forensic Sci Int* 157 (2006) 23-35.
- [25] M.L. Pontes, C. Carvalho and M.F. Pinheiro, Caracterização de SNPs do Cromossoma Y numa amostra populacional do Norte de Portugal. XI Jornadas del Grupo Español y Portugués de la Sociedad Internacional de Genética Forense, Madrid, 2006.
- [26] J.J. Sanchez, C. Borsting, C. Hallenberg, A. Buchard, A. Hernandez and N. Morling, Multiplex PCR and minisequencing of SNPs--a model with 35 Y chromosome SNPs, *Forensic Sci Int* 137 (2003) 74-84.
- [27] J.J. Sanchez, M. Brion, W. Parson, A.J. Blanco-Verea, C. Borsting, M. Lareu, H. Niederstatter, H. Oberacher, N. Morling and A. Carracedo, Duplications of the Y-chromosome specific loci P25 and 92R7 and forensic implications, *Forensic Sci Int* 140 (2004) 241-250.
- [28] M.T. Seielstad, J.M. Hebert, A.A. Lin, P.A. Underhill, M. Ibrahim, D. Vollrath and L.L. Cavalli-Sforza, Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition, *Hum Mol Genet* 3 (1994) 2159-2161.
- [29] P. Shen, F. Wang, P.A. Underhill, C. Franco, W.H. Yang, A. Roxas, R. Sung, A.A. Lin, R.W. Hyman, D. Vollrath, R.W. Davis, L.L. Cavalli-Sforza and P.J. Oefner, Population genetic implications from sequence variation in four Y chromosome genes, *Proc Natl Acad Sci U S A* 97 (2000) 7354-7359.
- [30] H. Skaletsky, T. Kuroda-Kawaguchi, P.J. Minx, H.S. Cordum, L. Hillier, L.G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, A. Chinwalla, A. Delehaunty, K. Delehaunty, H. Du, G. Fewell, L. Fulton, R. Fulton, T. Graves, S.F. Hou, P. Latrielle, S. Leonard, E. Mardis, R. Maupin, J. McPherson, T. Miner, W. Nash, C. Nguyen, P. Ozersky, K. Pepin, S. Rock, T. Rohlfling, K. Scott, B. Schultz, C. Strong, A. Tin-Wollam, S.P. Yang, R.H. Waterston, R.K. Wilson, S. Rozen and D.C. Page, The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes, *Nature* 423 (2003) 825-837.

- [31] B. Sobrino, M. Brion and A. Carracedo, SNPs in forensic genetics: a review on SNP typing methodologies, *Forensic Sci Int* 154 (2005) 181-194.
- [32] P.A. Underhill, L. Jin, A.A. Lin, S.Q. Mehdi, T. Jenkins, D. Vollrath, R.W. Davis, L.L. Cavalli-Sforza and P.J. Oefner, Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography, *Genome Res* 7 (1997) 996-1005.
- [33] P.A. Underhill, G. Passarino, A.A. Lin, P. Shen, M. Mirazon Lahr, R.A. Foley, P.J. Oefner and L.L. Cavalli-Sforza, The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations, *Ann Hum Genet* 65 (2001) 43-62.
- [34] P.A. Underhill, P. Shen, A.A. Lin, L. Jin, G. Passarino, W.H. Yang, E. Kauffman, B. Bonne-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J.R. Kidd, S.Q. Mehdi, M.T. Seielstad, R.S. Wells, A. Piazza, R.W. Davis, M.W. Feldman, L.L. Cavalli-Sforza and P.J. Oefner, Y chromosome sequence variation and the history of human populations, *Nat Genet* 26 (2000) 358-361.
- [35] P.M. Vallone and J.M. Butler, Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension, *J Forensic Sci* 49 (2004) 723-732.
- [36] P.M. Vallone, R.S. Just, M.D. Coble, J.M. Butler and T.J. Parsons, A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome, *Int J Legal Med* 118 (2004) 147-157.
- [37] P.S. Walsh, D.A. Metzger and R. Higuchi, Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, *Biotechniques* 10 (1991) 506-513.
- [38] J.D. Watson and F.H. Crick, The structure of DNA, *Cold Spring Harb Symp Quant Biol* 18 (1953) 123-131.
- [39] J.H. Wetton, K.W. Tsang and H. Khan, Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs, *Forensic Sci Int* 152 (2005) 45-53.