



POLIMORFISMOS GENÉTICOS E IDENTIFICAÇÃO HUMANA: O DNA COMO PROVA FORENSE

Leonardo Arduino Marano¹, Aginaldo Luiz Simões¹, Silviene Fabiana de Oliveira², Celso Teixeira Mendes-Junior^{3*}

¹ Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 14049-900, Ribeirão Preto-SP, Brasil

² Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, 70910-900, Brasília-DF, Brasil

³ Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 14040-901, Ribeirão Preto-SP, Brasil

Palavras-chave: genética forense, STRs, SNPs

*** Autor correspondente:**

Celso Teixeira Mendes-Junior

Departamento de Química

Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto

Universidade de São Paulo

Avenida dos Bandeirantes, 3900

14040-901, Ribeirão Preto-SP, Brasil

E-mail: ctmendes@ffclrp.usp.br

As técnicas de biologia molecular permitiram a caracterização da variabilidade do genoma humano e estão sendo aplicadas na área da ciência forense para identificar indivíduos que possam ser a fonte de material biológico associado a um crime. A análise de marcadores polimórficos de DNA é atualmente reconhecida mundialmente como uma técnica forense padrão para a investigação e detecção de uma ampla variedade de crimes, desde pequenos delitos até crimes hediondos (estupros ou latrocínios). A obtenção de perfis genéticos a partir de material biológico coletado em cena de crime apresenta enorme utilidade quando comparado com o perfil de DNA de eventuais suspeitos, auxiliando a polícia a estabelecer uma conexão entre o criminoso e a cena do crime, ou até mesmo a eliminar suspeitos na fase de inquérito, durante as investigações (Bond, 2007). Os polimorfismos de DNA também vêm sendo empregados para a identificação de vítimas de guerra ou de desastres de massa (Alonso et al., 2005).

O DNA presente em cromossomos é composto por regiões codificantes e não-codificantes. Os marcadores genéticos polimórficos empregados na atividade de identificação humana normalmente encontram-se presentes na região não codificante, quer seja entre genes ou dentro de genes (em *introns*). A localização cromossômica de um gene ou marcador genético (segmento de DNA) é

denominada *locus*. O segmento de DNA que ocupa um mesmo *locus* em diferentes cromossomos homólogos pode ser idêntico ou diferente, devido à ocorrência de mutações ao longo do tempo. As possibilidades alternativas de segmento de DNA que ocupam um determinado *locus* são denominadas alelos. Considerando que um indivíduo apresenta um par de cada cromossomo autossômico (um deles de origem paterna e o outro de origem materna), ele apresentará dois alelos em cada *locus*, o que determinará o seu genótipo. O termo heterozigose está relacionado à presença de dois alelos distintos (genótipo heterozigoto), enquanto que homozigose se refere à ocorrência de dois alelos idênticos (genótipo homozigoto) (Butler, 2005). O termo polimorfismo refere-se à presença de mais de um alelo de um mesmo gene ou segmento de DNA em uma população; um *locus* é dito polimórfico quando o seu alelo mais comum tem frequência igual ou inferior a 99% (Cavalli-Sforza et al., 1994).

A evolução da genética forense acompanhou os avanços do estudo da variabilidade genética, tendo o início ocorrido há, mais ou menos, um século, quando o grupo sanguíneo ABO foi descoberto por Karl Landsteiner (Jobling e Gill, 2004). Entretanto, a aplicabilidade forense do grupo sanguíneo ABO é limitada, uma vez que existem apenas quatro fenótipos ABO possíveis (A, B, AB e O), e cerca de 40% da população mundial apresenta o tipo O. Portanto, esse sistema é útil apenas para excluir um suspeito como fonte de origem de material biológico encontrado em cena de crime, no caso de incompatibilidade entre as amostras; por outro lado, se houver compatibilidade, não se pode concluir que o suspeito é a fonte de origem do material biológico (Butler, 2005).

Em 1984, Alec Jeffreys revolucionou a genética forense ao implementar polimorfismos genéticos denominados VNTRs (*Variable Number of Tandem Repeats*) em uma técnica designada DNA *fingerprinting*. Os VNTRs, também conhecidos como minissatélites, são

marcadores genéticos altamente polimórficos gerados por um arranjo em sequência (*in tandem*) de múltiplas cópias de um pequeno segmento de DNA (composto por mais de 10 bases). A análise era feita por meio da técnica de *Southern Blotting*, empregando sondas que reconheciam alelos de diferentes marcadores simultaneamente (sondas multilocais). Isso gerava um padrão de múltiplas bandas característico de cada indivíduo, denominado DNA *fingerprinting*. Pouco depois, ao invés de utilizar sondas multilocais, a comunidade científica passou a concentrar os esforços na análise de minissatélites identificados por sondas que reconheciam o par de alelos de um único minissatélite (sondas unilocais). Essa abordagem contribuiu para simplificar as interpretações de perfis de DNA e os cálculos estatísticos envolvidos (Jobling e Gill, 2004). Embora essa técnica não seja frequentemente utilizada atualmente, o termo DNA *fingerprinting* continua sendo empregado para designar perfis de marcadores genéticos polimórficos, independente de sua natureza.

O desenvolvimento da técnica de PCR (*Polymerase Chain Reaction*) proporcionou uma revolução no processo de identificação humana por DNA, ao permitir, em poucas horas, a obtenção de bilhões de cópias de um determinado *locus* presente em um pequeno fragmento de DNA. A amplificação do DNA por PCR resultou em um grande aumento de sensibilidade, fazendo com que materiais biológicos degradados, encontrados em pequenas quantidades em cenas de crime, pudessem ser analisados com sucesso. Dessa forma, marcadores genéticos polimórficos analisáveis em pequenos fragmentos de DNA (o que não é o caso dos minissatélites) passaram a representar a base que fundamenta a identificação humana por DNA (Butler, 2005; Goodwin et al, 2007; Jobling e Gill, 2004).

Dentre os principais marcadores genéticos polimórficos amplificáveis por PCR, os STRs (*Short Tandem Repeats*) e os SNPs (*Single Nucleotide Polymorphisms*) são aqueles que apresentam maior aplicabilidade forense (Figura 1). Este artigo tem por objetivo apresentar características que tornam esses polimorfismos genéticos de DNA imprescindíveis para a atividade de identificação humana.

STRs (*Short Tandem Repeats*)

Os STRs, também conhecidos como microssatélites, são polimorfismos genéticos gerados por um arranjo em sequência (*in tandem*) de múltiplas cópias de um pequeno segmento de DNA (2 a 6 pares de base apenas). Os STRs mais informativos podem apresentar mais de uma dezena de alelos diferentes, sendo que cada alelo conterà um número específico de cópias da unidade de

repetição (Figura 1). Portanto, a diversidade de alelos observada em STRs se deve à variabilidade do número de unidades repetidas contidas no segmento de DNA. Desse modo, dois indivíduos apresentam grande probabilidade de apresentarem alelos distintos em um determinado *locus*. Milhares desses polimorfismos já foram identificados em humanos, sendo que algumas estimativas apontam para a existência de cerca de um milhão de STRs distribuídos pelo genoma humano (Butler, 2005; Goodwin et al, 2007).

Devido à diversidade observada em cada *locus*, tais marcadores são altamente informativos. Os STRs mais polimórficos apresentam elevado poder de discriminação (probabilidade de que dois indivíduos selecionados aleatoriamente possuam genótipos distintos) e baixa probabilidade de match (probabilidade de que dois indivíduos selecionados aleatoriamente possuam genótipos idênticos). Cerca de 20 STRs autossômicos, selecionados dentre aqueles mais polimórficos, são normalmente empregados nas atividades humanas (Butler, 2005; Goodwin et al, 2007).

Após uma série de estudos relacionados à variabilidade dos STRs, o FBI (*Federal Bureau of Investigation*), agência governamental norte-americana, selecionou um conjunto de 13 STRs que passaram a compor o sistema CODIS (*Combined DNA Index System*) em 1997. Em conjunto, esses 13 marcadores apresentam probabilidade de match de aproximadamente $1,7 \times 10^{-15}$, sendo que o perfil composto pelos genótipos mais frequentes de cada um dos 13 marcadores apresenta uma probabilidade de ocorrência estimada em cerca de 1 em 160 bilhões (Butler, 2005). Em termos práticos, esses valores asseguram que cada indivíduo da população mundial (exceto no caso de gêmeos idênticos) apresente um perfil de marcadores genéticos polimórficos exclusivo no que se refere aos 13 CODIS. Os 13 marcadores podem ser analisados simultaneamente em um procedimento laboratorial automatizado, o que facilita sua aplicação forense até mesmo em situações nas quais a análise de um grande número de indivíduos é necessária.

O sistema CODIS consiste em um banco de dados de perfis de DNA de criminosos e amostras encontradas em cenas de crimes, interligando todos os 50 estados norte-americanos. Atualmente, mais de 8 milhões de perfis de DNA de criminosos e agressores e de 311 mil perfis de DNA de amostras forenses encontradas em cenas de crime estão armazenados no sistema. Mais de 114 mil correspondências entre perfis forenses e de criminosos (hits) foram efetuados até o momento, contribuindo para o sucesso de mais de 112 mil investigações (<http://www.fbi.gov/hq/lab/html/codis1.htm>, acessado em 07 de maio de 2010). É importante ressaltar que o FBI e a Polícia

Federal brasileira assinaram convênio em maio de 2009 para cessão do sistema CODIS ao governo brasileiro, o que permitirá à Polícia Federal criar um banco de dados nacional com amostras de DNA de criminosos, suspeitos e vítimas.

SNPs (*Single Nucleotide Polymorphisms*)

Os SNPs, polimorfismos de nucleotídeo único, podem ser definidos como polimorfismos criados por uma mutação de ponto, resultantes da substituição de um nucleotídeo por outro, proporcionando o surgimento de diferentes alelos (Figura 1). Esses polimorfismos estão presentes por todo o genoma nuclear e no DNA mitocondrial, tendo sido identificados mais de 10 milhões de SNPs até o presente.

Apesar desses marcadores poderem apresentar até quatro alelos (A, C, T ou G), eles são em sua maioria bialélicos, tornando-os menos informativos para a identificação humana do que os marcadores STRs conhecidos, uma vez que os mesmos podem apresentar mais de dez alelos por *locus* (Brookes, 1999; Goodwin et al, 2007). Desse modo, entre 50 e 100 SNPs são necessários para se alcançar o poder de discriminação dos 13 CODIS (Butler, 2005; Goodwin et al, 2007). No entanto, os SNPs podem ser extremamente vantajosos em algumas situações.

Dentre as vantagens, podemos citar a facilidade de esses polimorfismos serem analisados em larga escala por meio de tecnologias modernas e sua facilidade de adaptação a novas tecnologias de automação laboratorial. Deste modo, milhares de SNPs podem ser analisados simultaneamente em um único indivíduo. Além disso, SNPs podem ser de grande utilidade na análise de amostras degradadas, como as encontradas na cena de um crime ou em um cenário de desastre de massa (Alonso et al., 2005). Nesses casos, o material biológico pode se encontrar tão destruído, que os fragmentos de DNA estejam presentes em tamanhos menores do que aqueles necessários para uma genotipagem correta de STRs. Considerando que os STRs requerem normalmente a amplificação de fragmentos com tamanho variando de 100 a 450 pares de base, sua análise se torna prejudicada em amostras de DNA degradado em que fragmentos desse tamanho são dificilmente encontrados. Como o alvo da tipagem de SNPs é somente de um nucleotídeo, fragmentos extremamente pequenos (a partir de 33 pares de base) presentes em amostras degradadas podem ser analisados de forma precisa (Alonso et al., 2005; Goodwin et al, 2007).

Os SNPs podem também fornecer informações sobre a ancestralidade dos indivíduos (Frudakis, 2008). Os SNPs cujas frequências de seus alelos diferem subs-

tancialmente entre populações de regiões geográficas distintas (europeus, africanos, asiáticos e ameríndios, por exemplo) são denominados AIMs (Marcadores Informativos de Ancestralidade). Tais marcadores podem contribuir significativamente na resolução de crimes. Em 2002, uma série de estupros seguidos por assassinato das vítimas ocorreu na Louisiana, EUA. A análise de STRs (CODIS) das amostras de DNA encontradas nas vítimas não foi suficiente para a identificação do assassino, pois seu perfil não se encontrava no banco de dados do FBI. Relatos de testemunhas apontavam para um homem caucasiano visto nas redondezas de uma das cenas de crime, levando a investigação a se focar em suspeitos brancos do sexo masculino. No entanto, um ano após o começo das investigações, a análise de um conjunto de AIMs mostrou que o suposto criminoso era um indivíduo afro-americano, apresentando cerca de 85% de ancestralidade africana e 15% de ancestralidade ameríndia. Tais informações alteraram completamente o rumo das investigações e levaram à prisão do assassino um mês após essas análises (Frudakis, 2008).

Outro aspecto de grande relevância forense relacionado aos SNPs da análise desses polimorfismos decorre de sua presença nas regiões codificadoras e promotora dos genes, ou seja, em regiões responsáveis pela codificação das proteínas, que são por sua vez grandes responsáveis pelo fenótipo do indivíduo. Dessa forma, conjuntos específicos de SNPs presentes nesses genes já se mostraram relevantes para a determinação de características físicas, como a cor da pele, cabelo e olhos, espessura dos fios de cabelo, formas da face e estatura (Frudakis, 2008; Watson, 2000). Sendo assim, é possível que, no futuro, a análise de SNPs presentes nesses genes venha a contribuir para a determinação do retrato “falado” biomolecular de um criminoso que tenha deixado vestígios de sangue (ou de outros materiais biológicos) em uma cena de crime.

Considerações finais

Apesar dos STRs serem marcadores mais informativos e tradicionalmente empregados na atividade de identificação humana, a grande disponibilidade genômica de SNPs e a evolução tecnológica na área de detecção, em conjunto com as características pertinentes a tais marcadores, estão estimulando sua introdução em muitos laboratórios forenses.

Referências bibliográficas

Alonso A., Martin P., Albarran C., Garcia P., Simon L.F., Iturralde M.J., Fernandez-Rodriguez A., Atienza I., Capilla J., Garcia-Hirschfeld J., Martinez P., Vallejo G., Garcia O., Garcia E., Real P., Alvarez D., Leon A., Sancho M. (2005). Challenges of DNA profiling in mass disaster investigations. *Croat Med J* 46, 540-548.

- Bond J.W. (2007). Value of DNA evidence in detecting crime. *J Forensic Sci* 52, 128-136.
- Brookes A.J. (1999). The essence of SNPs. *Gene* 234, 177-186.
- Butler J.M. (2005). *Forensic DNA typing: biology, technology, and genetics of STR markers*. Second edition, Elsevier – Academic Press, USA.
- Cavalli-Sforza L.L., Menozzi P., Piazza A. (1994). *The history and geography of human genes*, Princeton University Press, Princeton, USA.
- Frudakis T. (2008). *Molecular photofitting: predicting ancestry and phenotype using DNA*. Elsevier – Academic Press, USA.
- Goodwin W., Linacre A., Hadi S. (2007). *An introduction to forensic genetics*. John Wiley & Sons Ltd, England.
- Jobling M.A., Gill P. Encoded evidence: DNA in forensic analysis. *Nat. Rev. Genet.* 5, 739-752.
- Watson A. (2000). New tools. A new breed of high-tech detectives. *Science* 289, 850-854.

A) Cena de crime:

5' - . . . CTTTAATGCACTACTAATGAATGAATGAATGAATGAATGCTCAAT . . . -3'
 3' - . . . GAAATTACGTGATGATTACTTACTTACTTACTTACTTACTTACTGAGTTA . . . -5'

5' - . . . CTTTAATACACTACTAATGAATGAATGAATGAATGCTCAAT -3'
 3' - . . . GAAATTATGTGATGATTACTTACTTACTTACTTACTGAGTTA -5'

B) Suspeito:

5' - . . . CTTTAATGCACTACTAATGAATGAATGAATGAATGAATGCTCAAT -3'
 3' - . . . GAAATTACGTGATGATTACTTACTTACTTACTTACTTACTGAGTTA -5'

5' - . . . CTTTAATGCACTACTAATGAATGAATGAATGAATGAATGCTCAAT -3'
 3' - . . . GAAATTACGTGATGATTACTTACTTACTTACTTACTTACTGAGTTA -5'

Figura 1. Representação esquemática da sequência nucleotídica de um pequeno fragmento cromossômico humano obtido a partir de amostras de DNA extraídas de (A) material biológico encontrado em uma cena de crime e (B) sangue de um suspeito. Destaca-se a presença de um SNP (vermelho) e um STR (azul). Observa-se que a amostra da cena de crime é heterozigota em ambos marcadores (A/G e 5/7), enquanto que a amostra do suspeito é homozigota em ambos marcadores (G/G e 6/6). Embora a análise de vários marcadores seja necessária para que se alcance uma conclusão segura e definitiva, estes dois marcadores indicam que o suspeito não é a fonte de origem do material biológico encontrado na cena de crime.