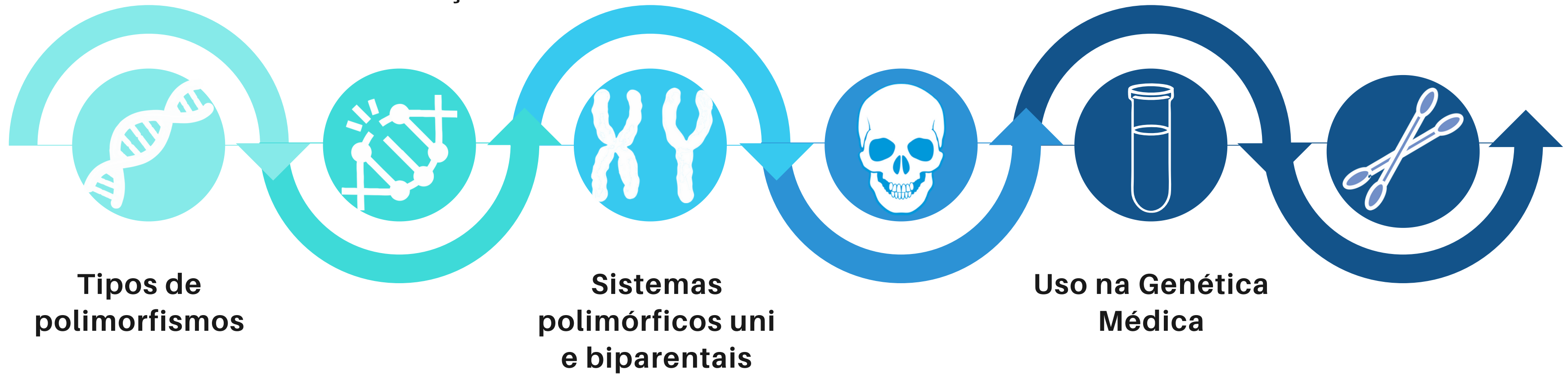


GENOMA HUMANO: SISTEMAS POLIMÓRFICOS DE INTERESSE ANTROPOLÓGICO, MÉDICO E FORENSE

Mecanismos
de origem e
manutenção

Uso na
Antropologia

Uso Forense

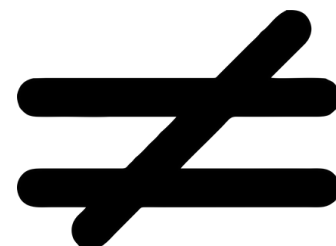




O QUÉ É UM POLIMORFISMO?

Uma sequência é polimórfica quando ela apresenta no mínimo **dois alelos** na população, cada um com **frequência $\geq 1\%$** .

Pode ser **bialélica** ou **multialélica**.



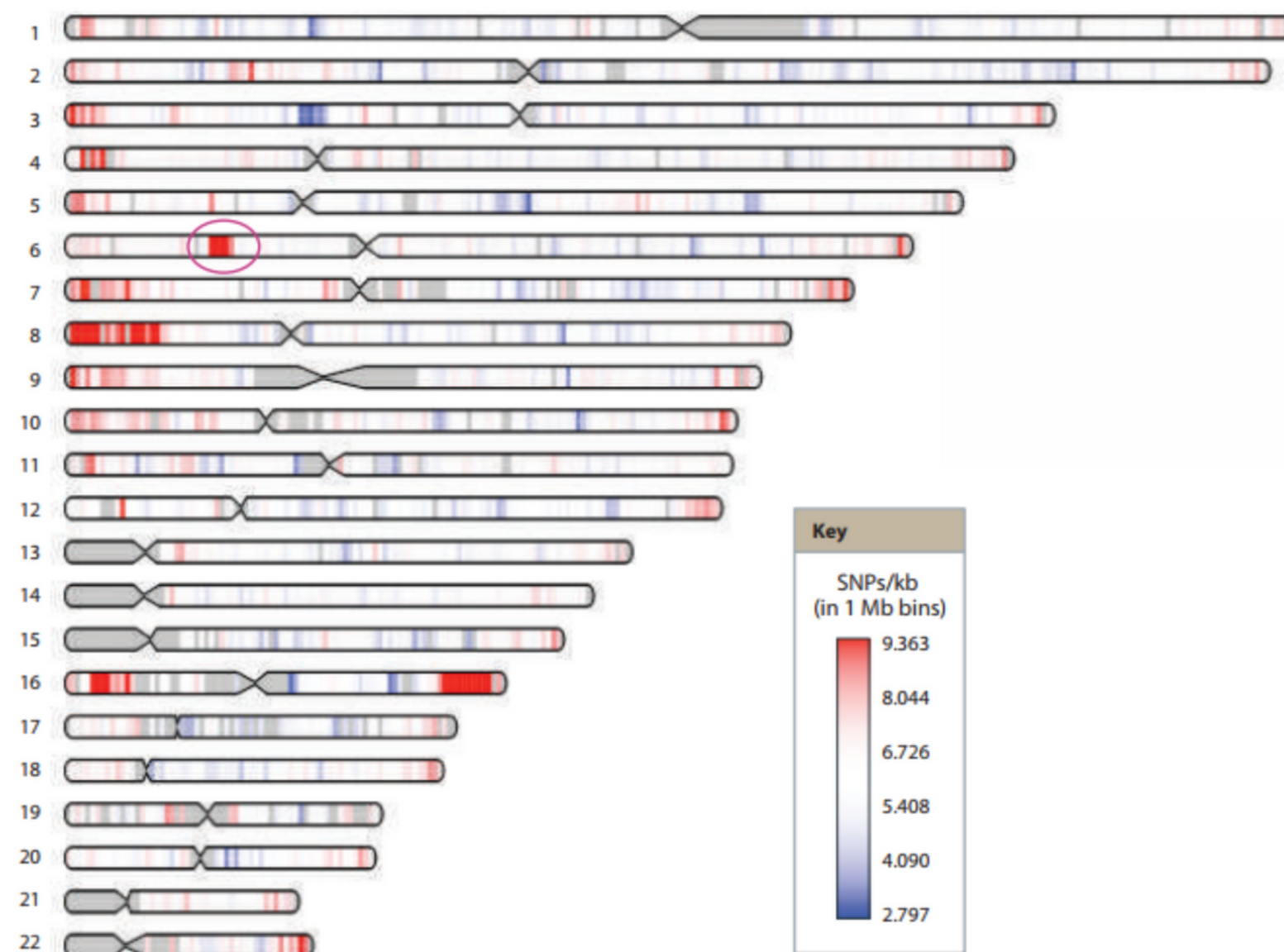
Variante: diferença menos frequente entre cópias de genoma.

Mutação: mudança ***de novo*** (que ocorre pela primeira vez).

Monomórfica = Monoalélica.



Polimorfismo é um conceito qualitativo (sim ou não) o que pode variar é o **grau de polimorfismo** (mais ou menos alelos).



SNPs por região genômica

Em destaque em vermelho a região HLA no braço curto do cromossomo 6, com alto grau de polimorfismo.

Fonte: Jobling et al., 2014



O que é necessário para que uma mutação se torne um polimorfismo?

Mutações são o resultado de danos ao DNA e falhas no sistema de reparo. A diversidade num dado locus depende tanto da sua taxa de mutação como da seleção natural e da história populacional.

- Devem ocorrer na **linhagem germinativa** (células que dão origem aos gametos).
- Devem permitir a **sobrevivência** e **não afetar a reprodução**. Ex.: Mudanças no número de cromossomos (aneuploidias) decorrem de falhas na segregação cromossômica na meiose por isso não são herdáveis.
- Cada alelo deve alcançar uma frequência $\geq 1\%$ na população.



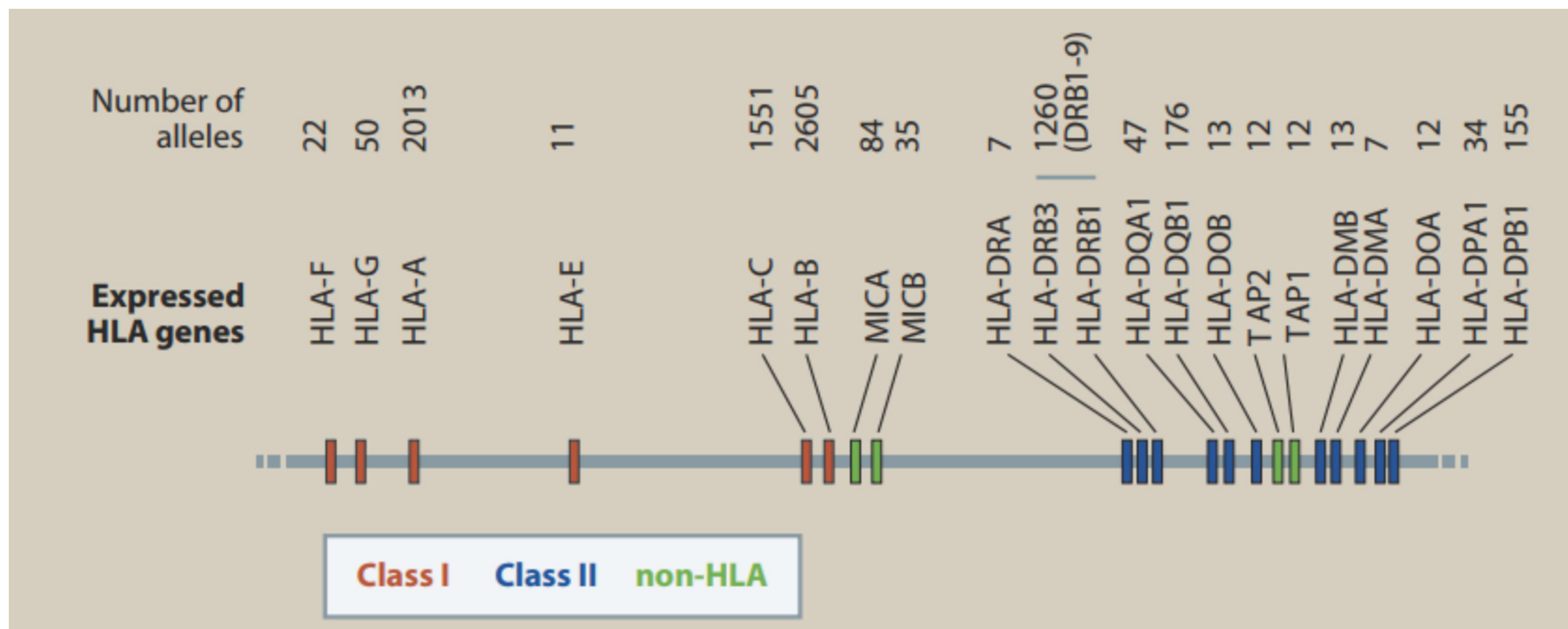
Alelos do locus HLA

Um exemplo de seleção balanceadora

Os genes HLA (**Antígeno Leucocitário Humano**, *Human Leucocyte Antigen*) apresentam antígenos aos linfócitos T, estando envolvidos com a **resposta imune adaptativa**. **Diferentes alelos** de cada gene individual **variam na habilidade de apresentar antígenos** de diferentes patógenos.

O locus HLA (3,6 Mb no cromossomo 6) está dividido em 3 regiões, chamadas "classes".

Nas regiões de classe I e II, antigos **eventos de duplicação** geraram distintos genes expressos e pseudogenes.

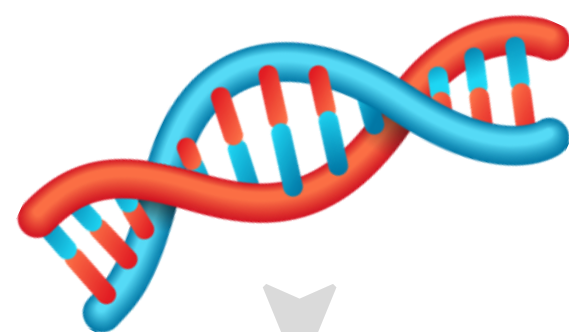


Os diferentes alelos são mantidos na população por **seleção balanceadora**, permitindo que a mesma esteja preparada para a sobrevivência a **patógenos diversos**. Hoje são os principais determinantes da **histocompatibilidade**.

Fonte: Jobling et al., 2014

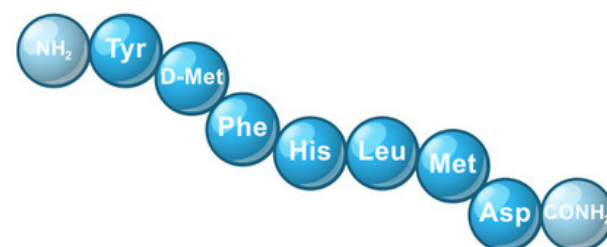


Tipos de Polimorfismos



Marcadores de DNA

Detectam diretamente as variações na sequência de nucleotídeos



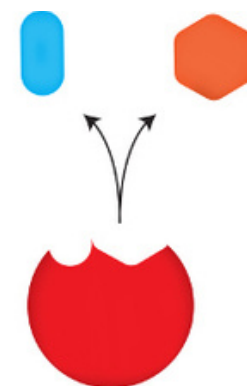
Marcadores Clássicos

Detectam a variação genética indiretamente, através da análise do produto proteico



Marcadores Imunohistoquímicos

Baseiam-se no reconhecimento de um antígeno (produto da variação no DNA) por um anticorpo



Marcadores Bioquímicos

Baseiam-se na detecção da variação na cadeia de aminoácidos da proteína. Pode ser por eletroforese ou por reação enzimática.



TIPOS DE POLIMORFISMO GENÉTICO

SNPs

Single Nucleotide Polymorphism
Polimorfismo de Nucleotídeo Único

Substituições ou indels
(inserção/deleção)
de uma **única base**.

VNTR

Variable Number of Tandem Repeats
Número Variável de Repetições em Tandem

Polimorfismo de **número de cópias** de uma sequência **repetitiva**

- **Satélites** (≥ 100 pb)
- **Minissatélites** (8-100 pb)
- **Microssatélites** (1-7 bp)

Elementos Transponíveis

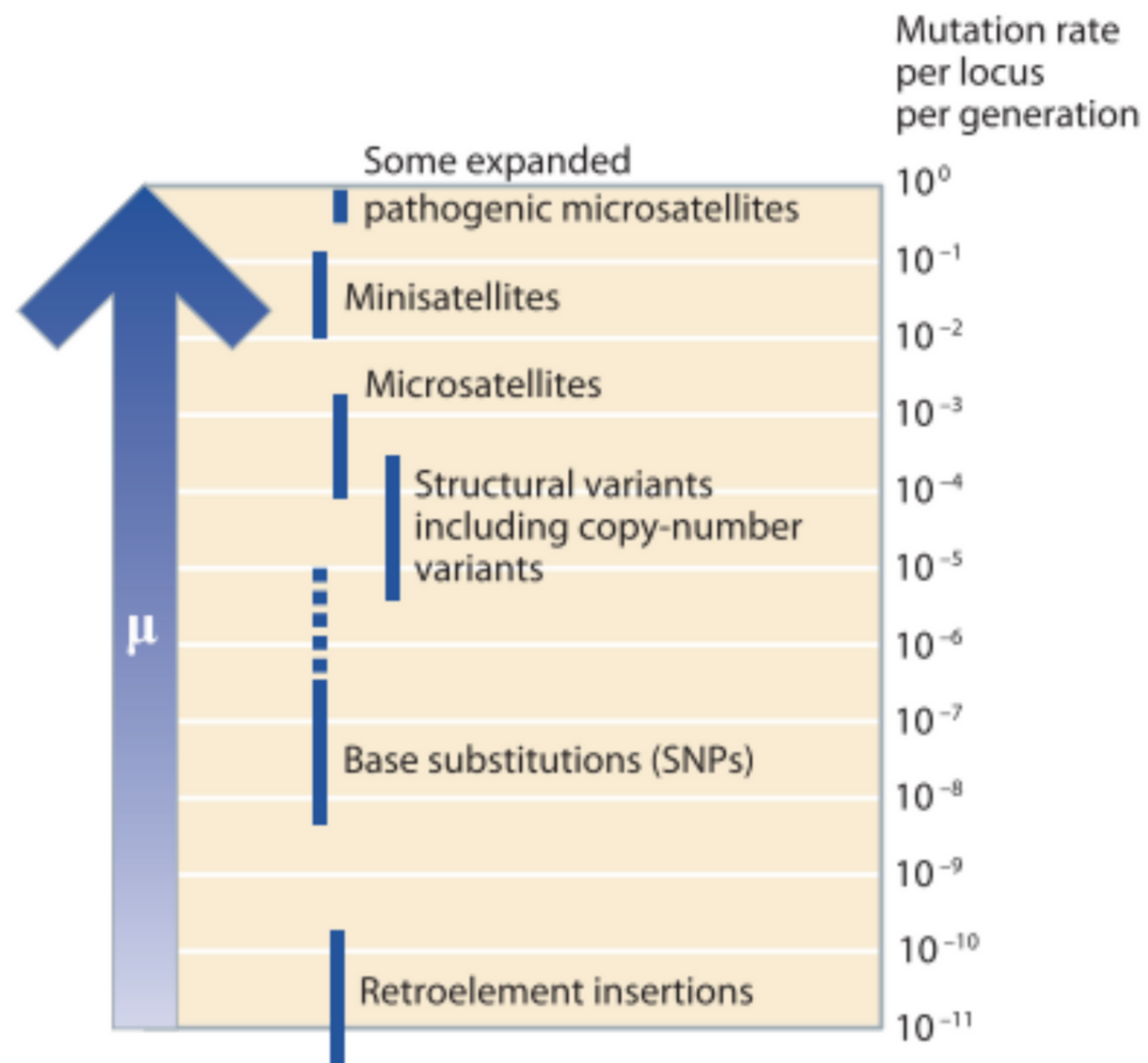
Também chamados elementos moveis ou repetições dispersas. Sequências que inserem cópias em outras partes do genoma.
45% do genoma

Variantes Estruturais

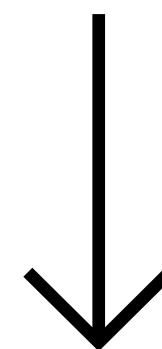
Inversões e translocações balanceadas ou imbalances genômicos (**indels**) abrangendo ≥ 1 kb



TIPOS DE POLIMORFISMO GENÉTICO



Velocidade de Mutação
varia segundo o tipo de polimorfismo

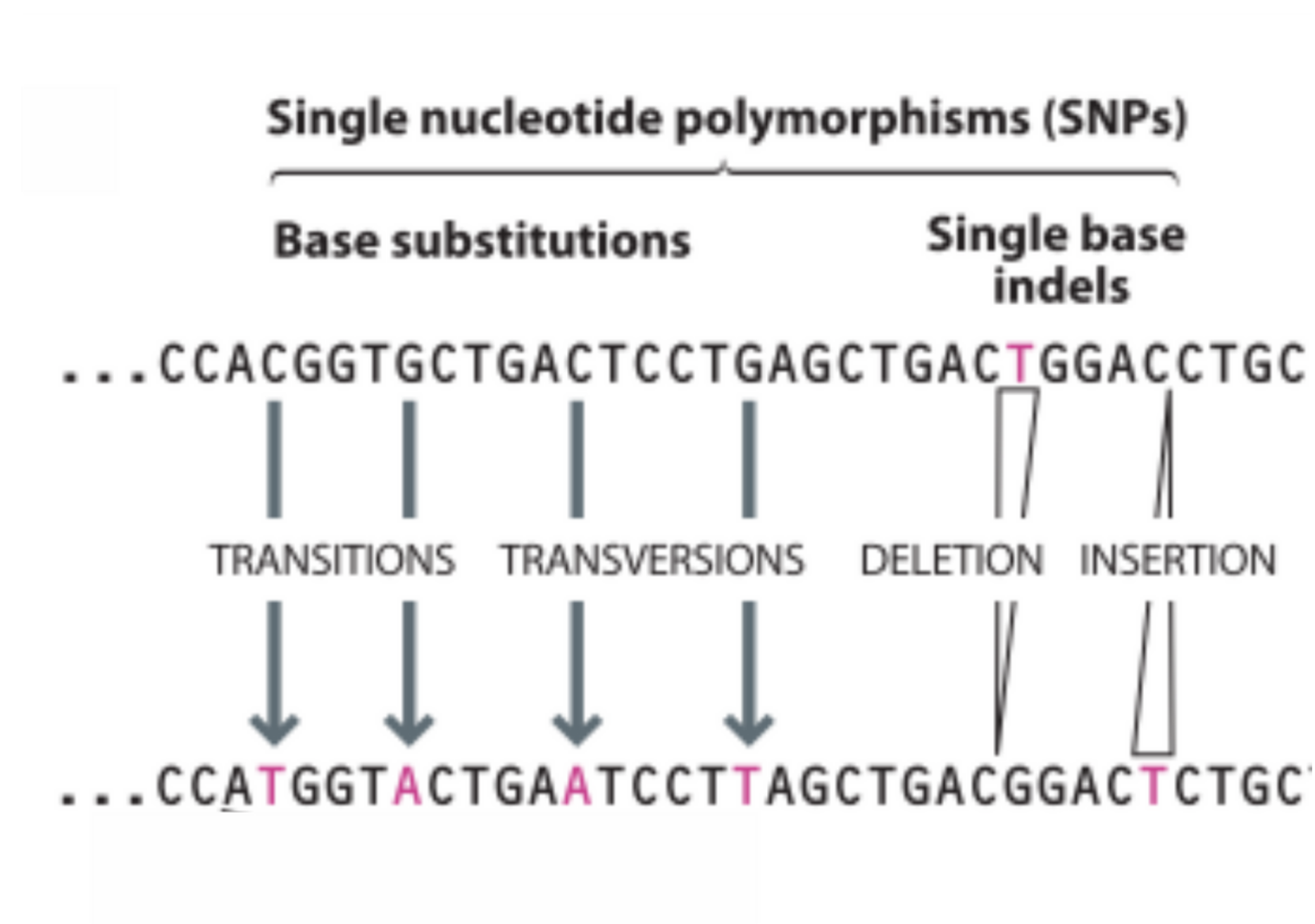


Afeta a sua aplicabilidade nas
distintas áreas de estudo



SNP

Polimorfismo de Nucleotídeo Único





SNP

Polimorfismo de Nucleotídeo Único

Características

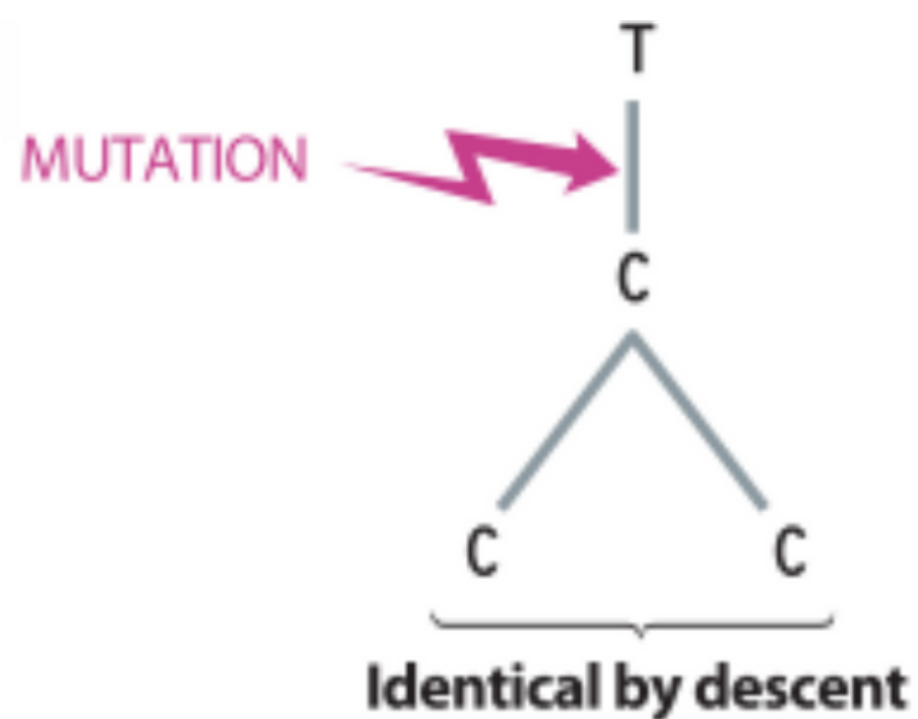
- **Nomenclatura:** rs, por Reference SNP. Ex.: rs334 (anemia falciforme).
- Ocorrem em todos os cromossomos (autossomos, Y, X, DNAMt)
- **Baixa velocidade de mutação.**
- São comumente **bialélicos** (casos de trialélicos são raros).
- Apresentam **identidade por ascendência** (IBD, identity by descent).
- O **alelo ancestral** pode ser deduzido olhando o alelo existente nos grandes macacos. O outro alelo se denomina **alelo derivado**.
- São descritos pelo **MAF = Minor Allele Frequency**, ou a frequência do alelo menos comum.
- Quando ambos os alelos ocorrem com **frequência $\geq 5\%$** se denominam **variantes comuns**.



SNP

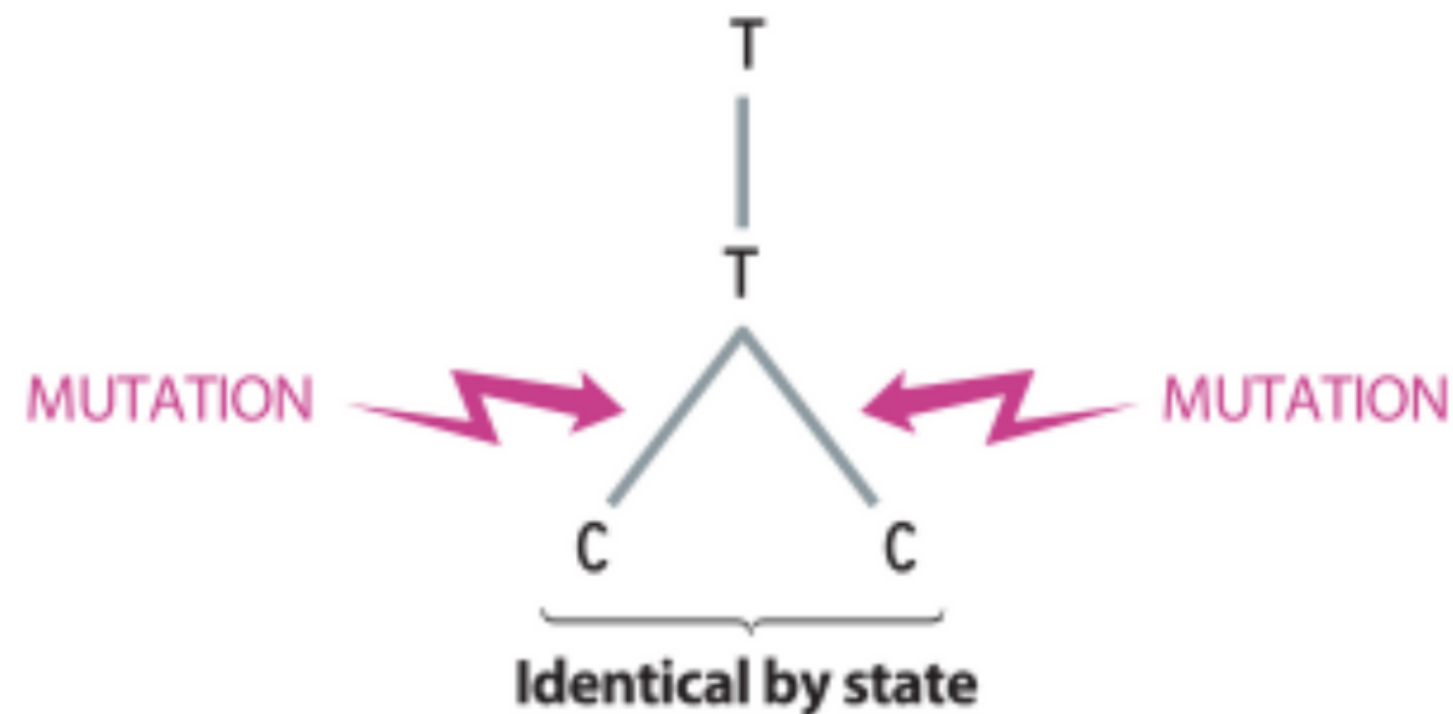
Polimorfismo de Nucleotídeo Único

SNPs
apresentam **identidade por
descendência**



≠

Identidade por Estado





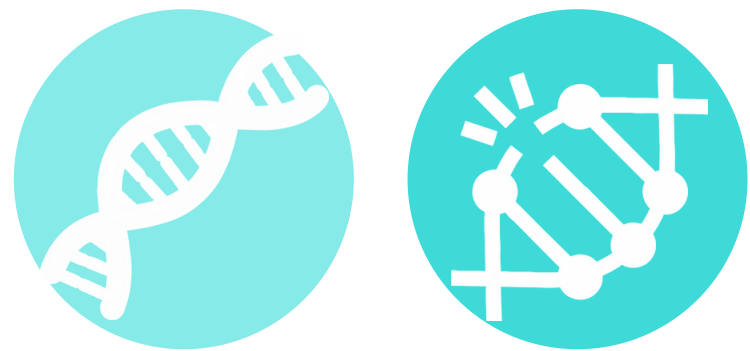
SNP

Polimorfismo de Nucleotídeo Único

Os genomas sequenciados até a data mostraram, em média, ~3-4 milhões de SNPs por genoma.

Estudos em populações europeias mostraram que é esperado encontrar um SNP a cada 1250 pb. Para africanos essa frequência aumenta a um SNP a cada 1006 pb.

A diversidade genética africana é a maior do mundo.



SNP

Mecanismos de origem

Os SNPs se originam por **mutação e falha no sistema de reparo.**

As mutações podem ocorrer por **mecanismos endógenos**, por agentes **físicos ou químicos.**

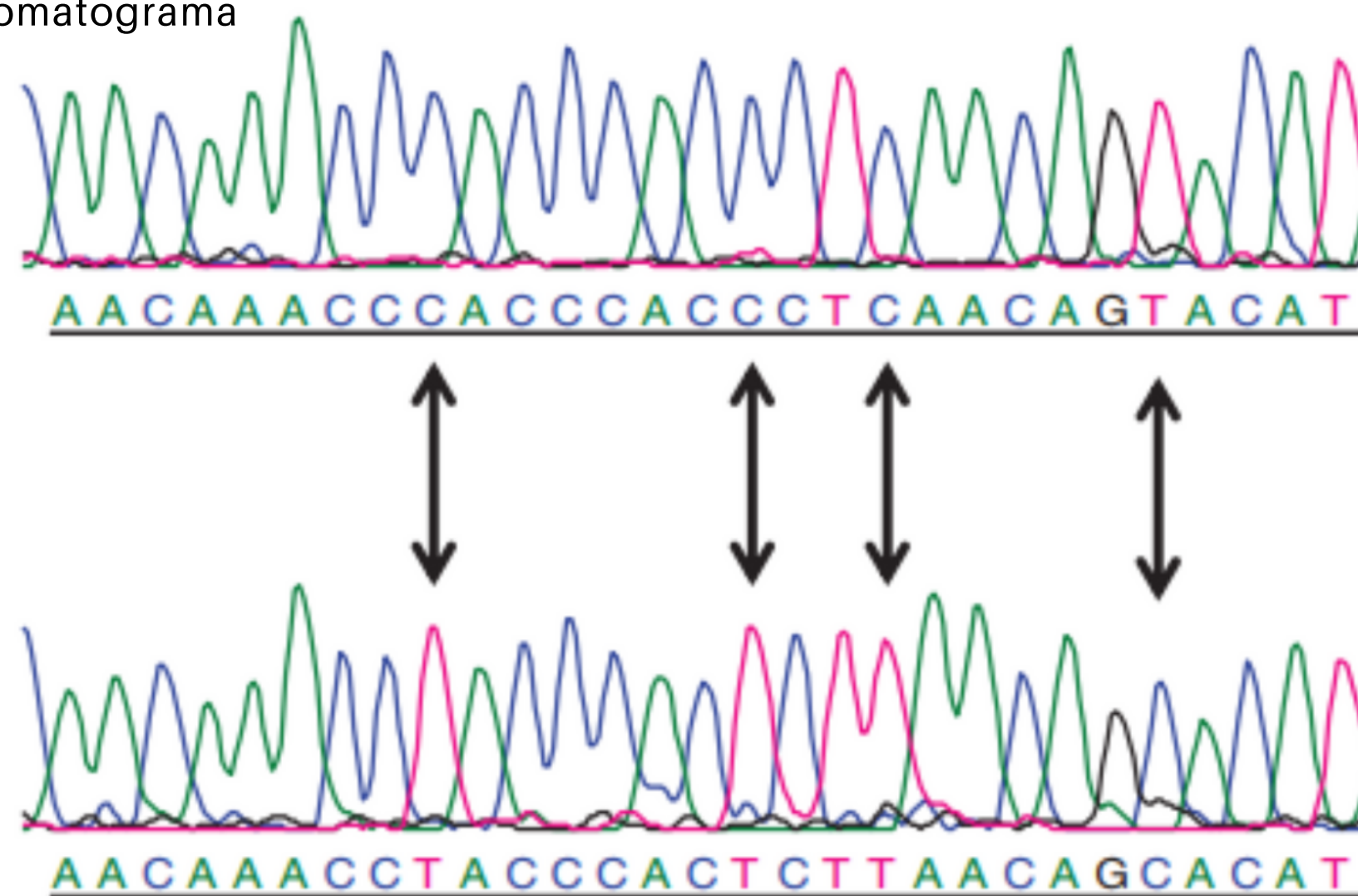
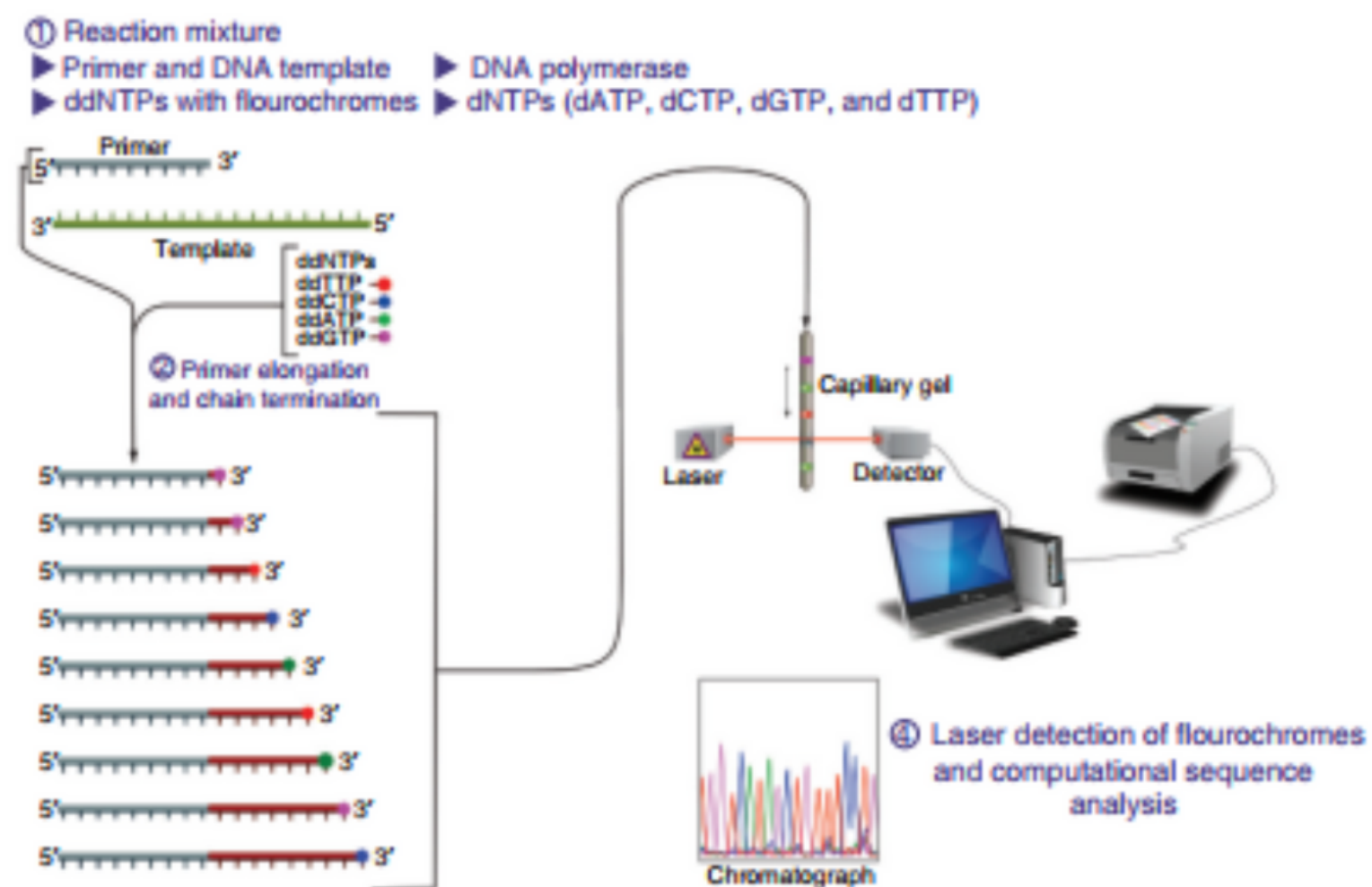


SNP: Métodos de estudo

Sequenciamento de Sanger ou método dos dideoxy

Também chamado de sequenciamento de primeira geração. A técnica permite detectar cada nucleotídeo na medida que este é adicionado durante a reação de polimerização. São utilizados dideoxy nucleotídeos (ddNTPs) marcados cada um com uma cor fluorescente diferente. Quando um ddNTP é adicionado pela polimerase a reação termina, gerando um fragmento com o comprimento equivalente à posição do nucleotídeo e uma cor específica para cada base. Ao final todos os fragmentos são separados em um tubo capilar e detectados pelo sequenciador.

Cromatograma





SNP: Métodos de estudo

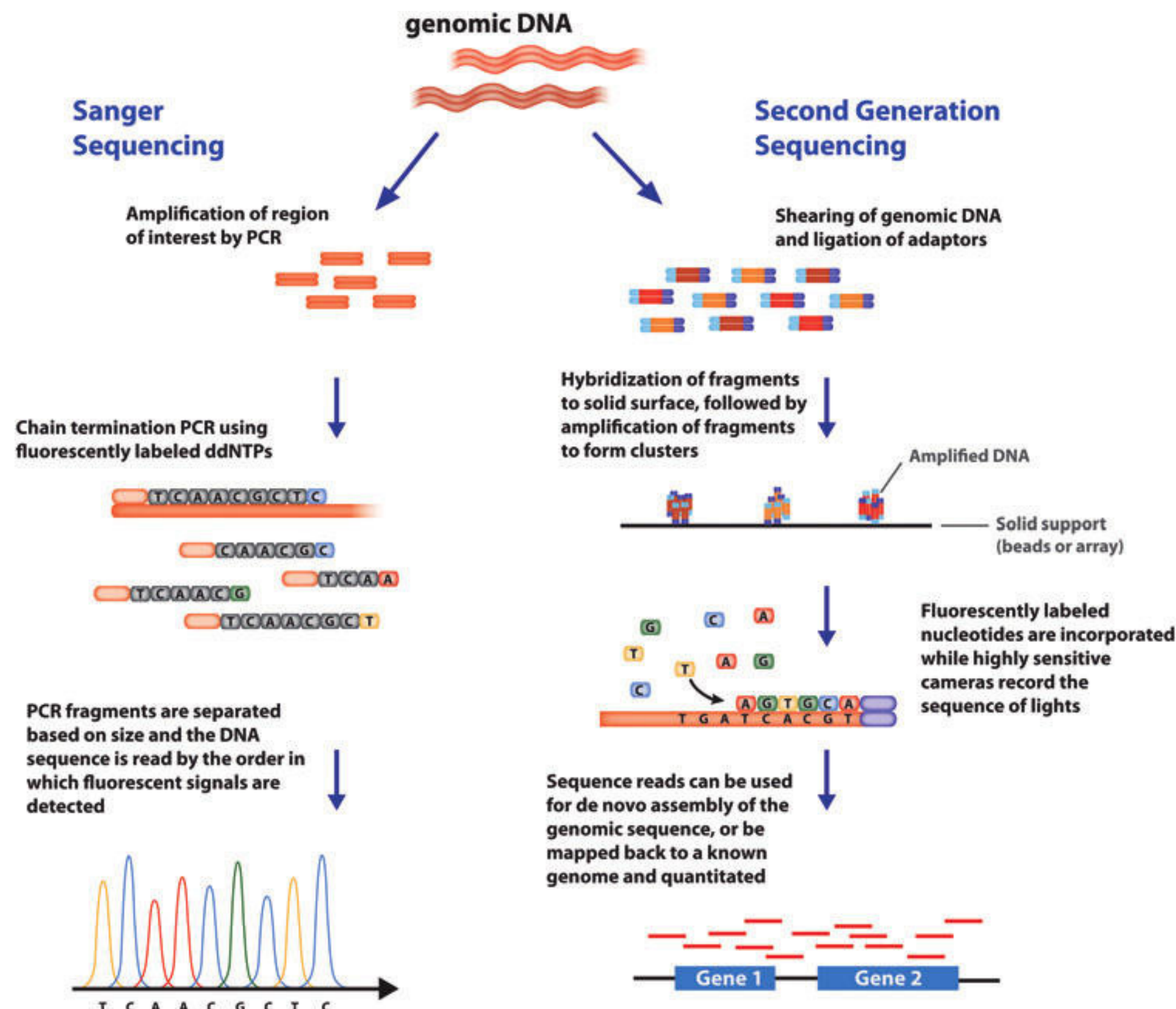
Sequenciamento de Nova Geração (NGS, *Next Generation Sequencing*)

Também chamado de sequenciamento de segunda geração.

O termo inclui vários métodos que diferem dependendo da empresa que oferece o serviço, já que este requiere de um equipamento adequado. As características gerais do NGS incluem uma etapa de fragmentação do DNA genômico (que não precisa ser amplificado previamente como no sequenciamento de Sanger) e uma segunda etapa de imobilização dos fragmentos numa superfície sobre a qual será realizada a reação de sequenciamento. Esses métodos permitem **sequenciar genomas completos ou grandes fragmentos e, inclusive, muitas amostras numa mesma reação.**

Novas abordagens permitem o sequenciamento apenas regiões exônicas (codificadoras de proteínas).

Fonte: Bunnik, 2013



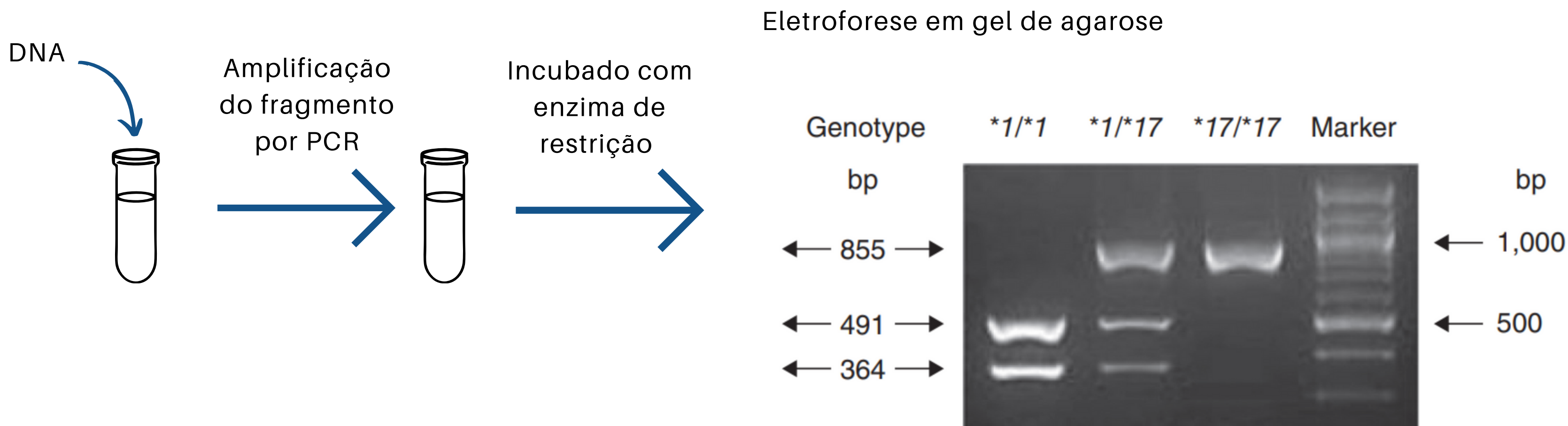


SNP: MÉTODOS DE ESTUDO

PCR-RFLP

Restriction Fragment Length Polymorphism = Polimorfismo de comprimento dos fragmentos de restrição

Após a amplificação do fragmento que contém o SNP este é incubado com enzimas de restrição que reconhecem e cortam a sequência ao redor dele. A presença de um dos alelos evita o reconhecimento e o corte pela enzima, gerando um fragmento maior e reconhecível pelo seu tamanho em um gel de eletroforese.



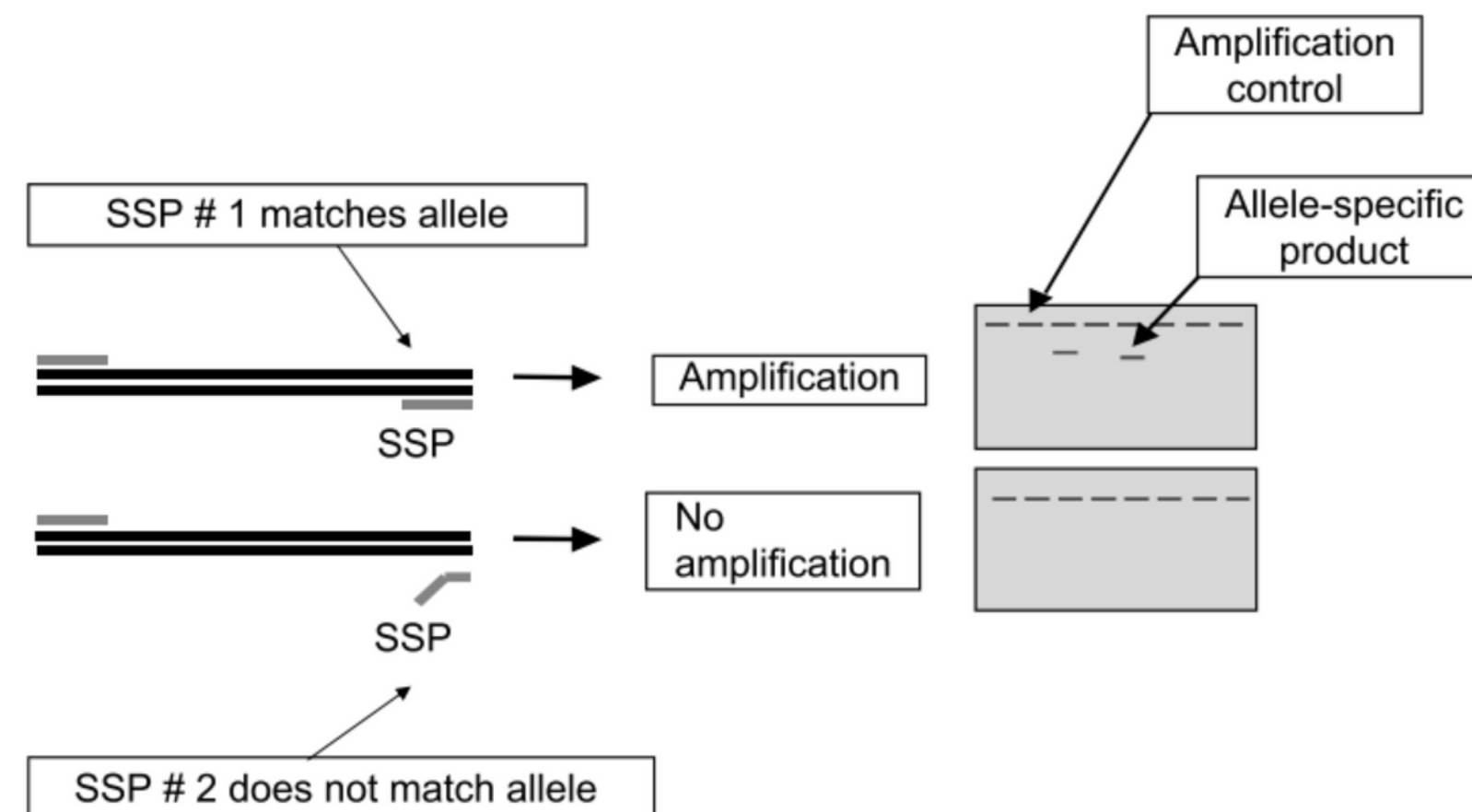


SNP: MÉTODOS DE ESTUDO

PCR-SSP

Sequence Specific Primers= Iniciadores Específicos de Sequência

Esta técnica utiliza iniciadores específicos para cada um dos alelos de um SNP. Desenham-se iniciadores cujos extremos 3' devem hibridizar cada um com apenas um dos alelos do SNP, e um terceiro *primer* comum para ambas as reações. Após a reação os tamanhos dos fragmentos se visualizam por eletroforese. Assim, indivíduos homozigotos terão amplificação a partir de apenas um dos primers enquanto indivíduos heterozigotos terão amplificação a partir de ambos os primers, observándose duas bandas no gel.

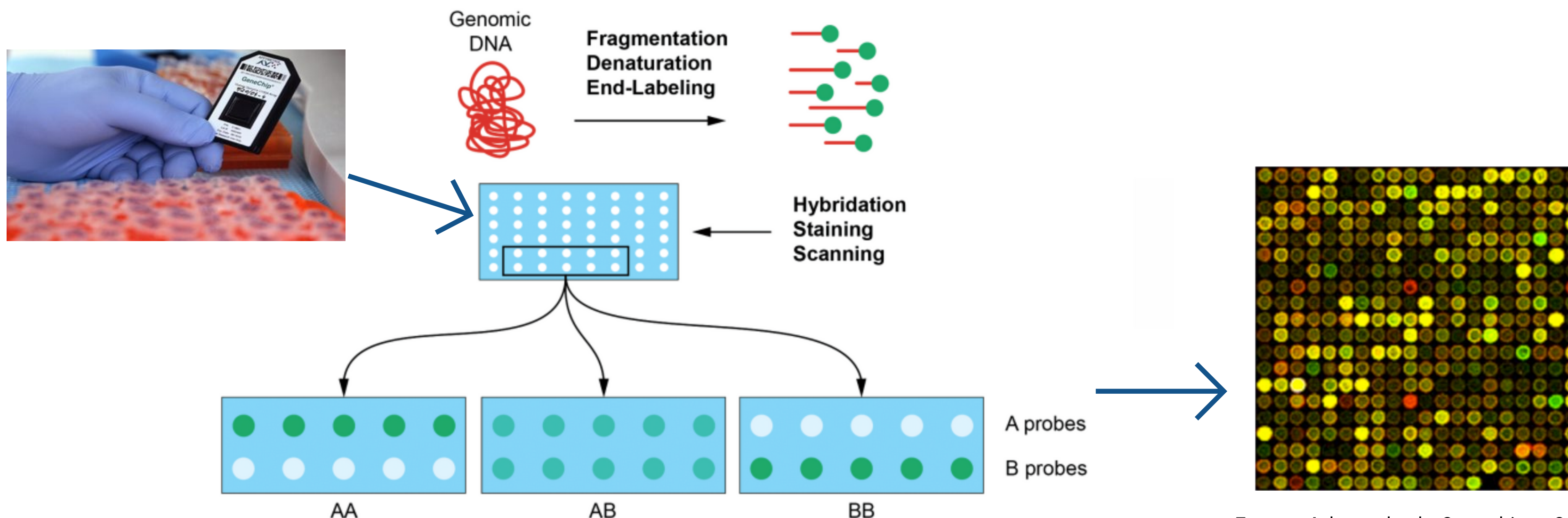




SNP: MÉTODOS DE ESTUDO

Microarrays = Microarranjos de DNA

Também chamados "*chips*". Permitem a genotipagem de milhares de SNPs para cada amostra, geralmente abrangendo "Tag" SNPs de todo o genoma (Genome-Wide). Os chips são produtos comerciais e desenhados para a genotipagem de conjuntos de SNPs pre-selecionados (não permitem a detecção de novos SNPs). Em um *chip*, cada poço possui sondas marcadas com fluorescentes diferentes para cada alelo (verde ou vermelho). O sinal amarelo indica a hibridização com ambas as sondas (o indivíduo é heterozigoto para esse locus). Cada chip permite a genotipagem de uma amostra mas eles são vendidos por quantidades (≥ 100 chips).



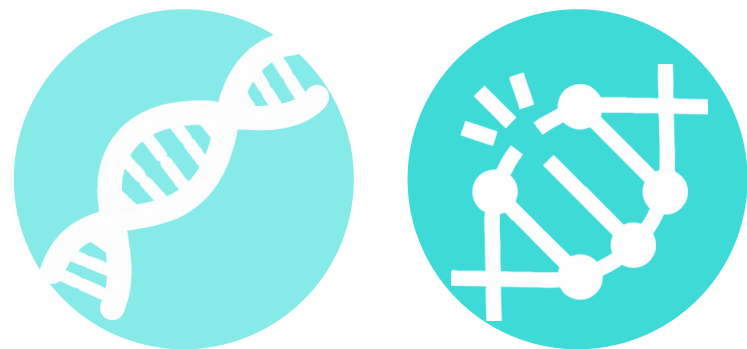


MICROSSATÉLITES

ou Short Tandem Repeats (STR)

Características

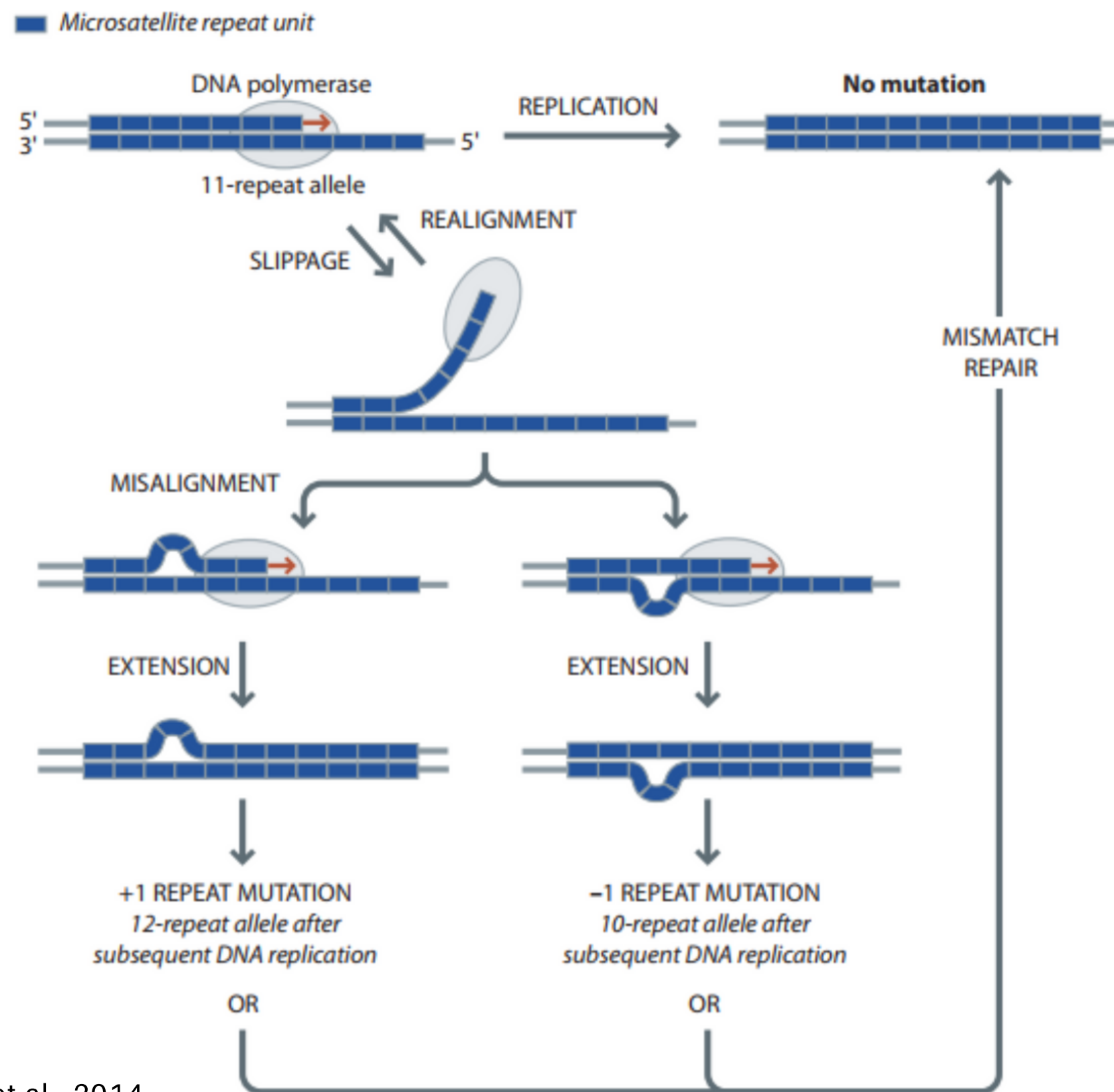
- São repetições de sequências de entre **1-7 bp** e, usualmente, **10-30 cópias**.
- Ocorrem em **autossomos, Y, X e mtDNA**.
- **Alta velocidade de mutação**.
- São comumente **multialélicos**, sendo os seus alelos os distintos números de cópias.
- Alelos com o mesmo tamanho em distintos indivíduos não são necessariamente idênticos por descendência e sim, **idênticos por estado** (evolução convergente).
- Combinações de STRs podem produzir perfis genotípicos únicos para um indivíduo.
- O **alelo ancestral** não pode ser deduzido olhando o alelo existente nos grandes macacos.



MICROSSATÉLITES

Mecanismo de origem

Aceita-se que os STR se originam pelo **escorregamento das fitas** repetitivas durante a replicação.



Fonte: Jobling et al., 2014



MICROSSATÉLITES

Métodos de estudo

A técnica fundamental para o estudo dos STR é a PCR combinada com outras técnicas:

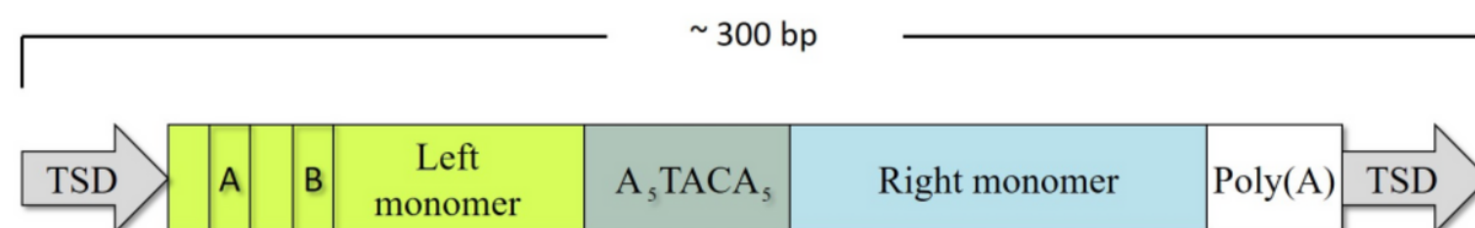
- **PCR + eletroforese:** utilizam-se iniciadores que flanqueiam a região polimórfica. Após a amplificação, o produto de PCR é visualizado através de um gel de agarose e os distintos alelos são diferenciados pelo tamanho das bandas.
- **PCR + sequenciamento:** um dos iniciadores da PCR é marcado com fluorescência no extremo 5'. Após a amplificação, os produtos marcados fluorescentemente são sequenciados em sequenciador capilar (o mesmo utilizado no sequenciamento de Sanger) permitindo a determinação do comprimento de cada fragmento. A marcação de distintos microssatélites com diferentes cores permite a sua genotipagem simultânea (técnica denominada de **multiplex**).



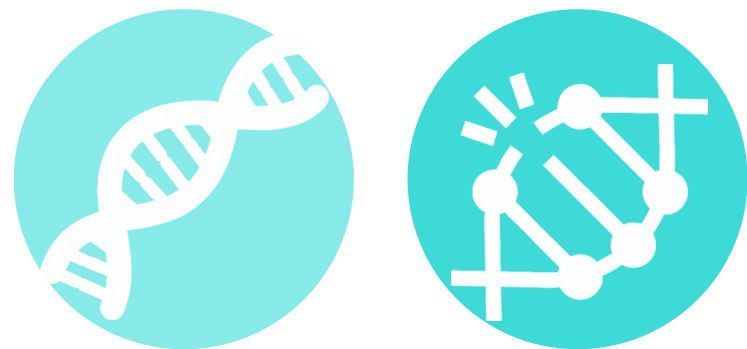
INSERÇÕES *Alu*

Características

- São **retrotransposons** curtos (~300 pb) derivados de uma sequência original que se transpõe periodicamente a outras regiões do genoma.
- Transposições *Alu* têm ocorrido desde a origem dos primatas e por isso todos os humanos possuem certas cópias em regiões específicas do genoma (*Alu* não polimórficas ou **fixas**), enquanto outras ocorreram mais recentemente originando **polimorfismo de inserção *Alu* (AIP)**.
- **Velocidade de transposição muito baixa.**
- Apresentam **identidade por descendência.**
- O **estado ancestral é sempre a ausência** da sequência *Alu*.
- Ocorrem nos autossomos, X e Y.
- Existem > 1 milhão *Alus* no genoma humano, mas apenas ~2000 são ativas.



Fonte: Kim et al., 2016

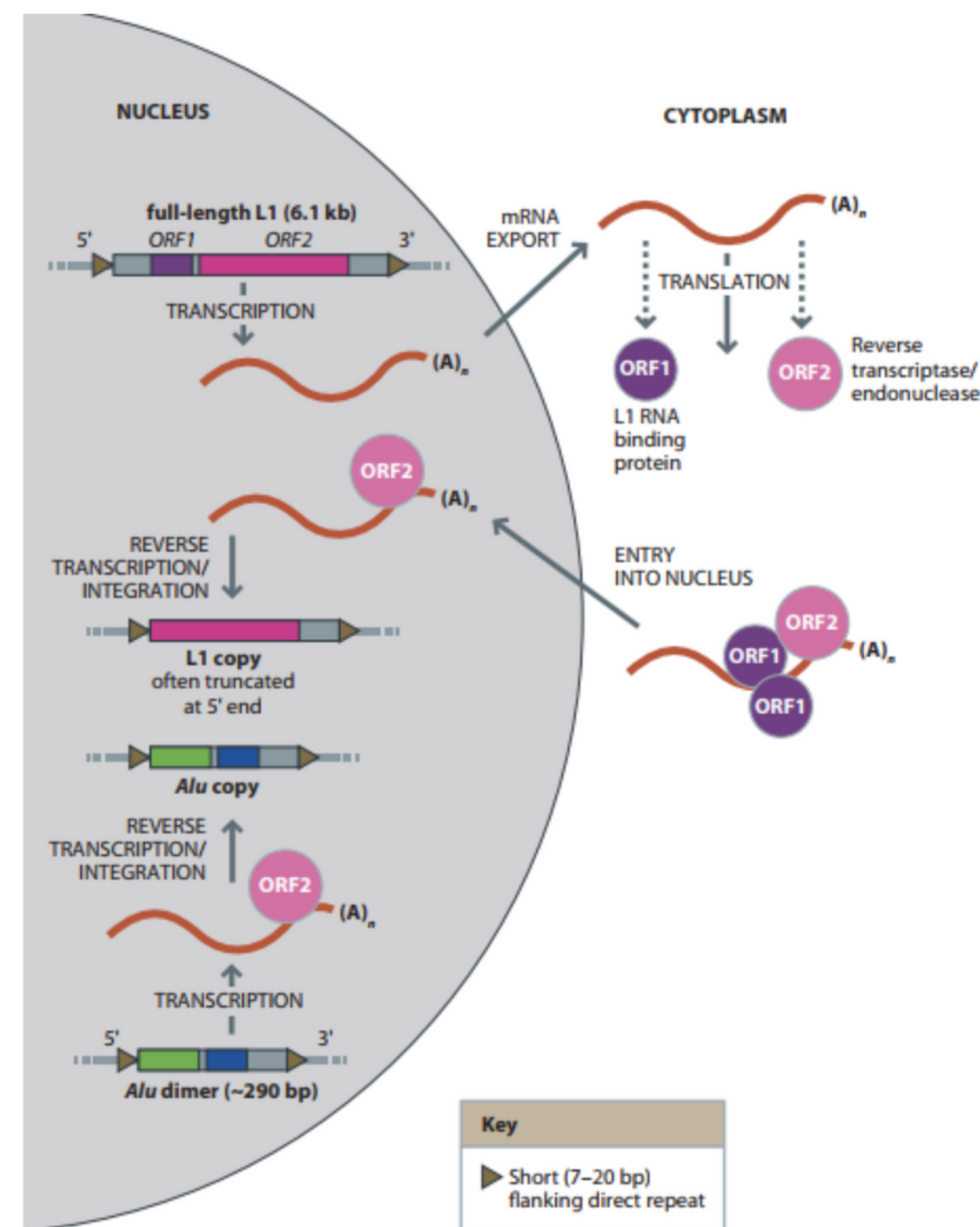


INSERÇÕES *Alu*

Mecanismo de origem

"Copiado e colado"

- São retrotransposon não autônomos: RNA polimerase III realiza a sua transcrição. Utiliza a maquinaria enzimática de outro elemento transponível (L1) para se reinserir em outro local. A endonuclease de L1 reconhece e corta uma sequência consenso no sítio de inserção e a retrotranscriptase de L1 retrotranscreve o RNA *Alu* para DNA, que pode então ser inserido na nova posição.

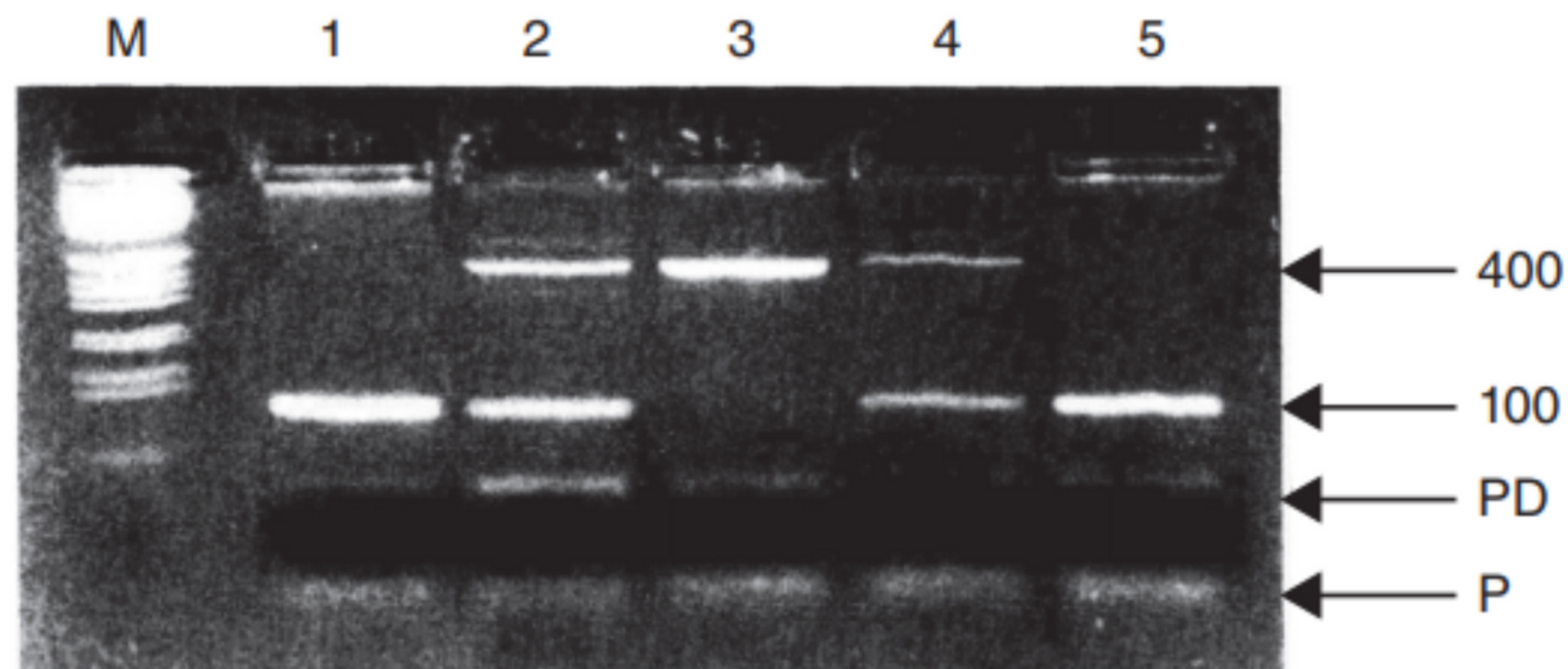




INSERÇÕES *Alu*

Método de estudo

Devido à presença de regiões consenso que flanqueiam o elemento *Alu* e a sua sequência relativamente curta, eles podem ser facilmente detectados por **PCR + eletroforese** em gel de agarose. Em presença da inserção, a PCR amplifica uma banda 300 pb maior do que na ausência da mesma.



banda de 400 pb: presença de *Alu*
banda de 100 pb: ausência de *Alu*
PD: dímero de *primer*
P: *primer* sobrando
M: marcador de peso molecular

Fonte: Jobling et al., 2014



Sistemas polimórficos uni e biparentais

A **forma de herança** determina a aplicabilidade de um marcador genético para diversos estudos.

Marcadores biparentais: autossomos, cromossomo X.

Marcadores uniparentais: cromossomo Y, DNAm.



Sistemas polimórficos biparentais

- Autossomos: 99,9% do DNA da célula.
- São os marcadores mais estudados para entender as **bases genéticas de fenótipos**.
- O seu estudo mostra a **história completa de miscigenação** de um indivíduo/população.
- Sofrem **recombinação**.



Sistemas polimórficos uniparentais

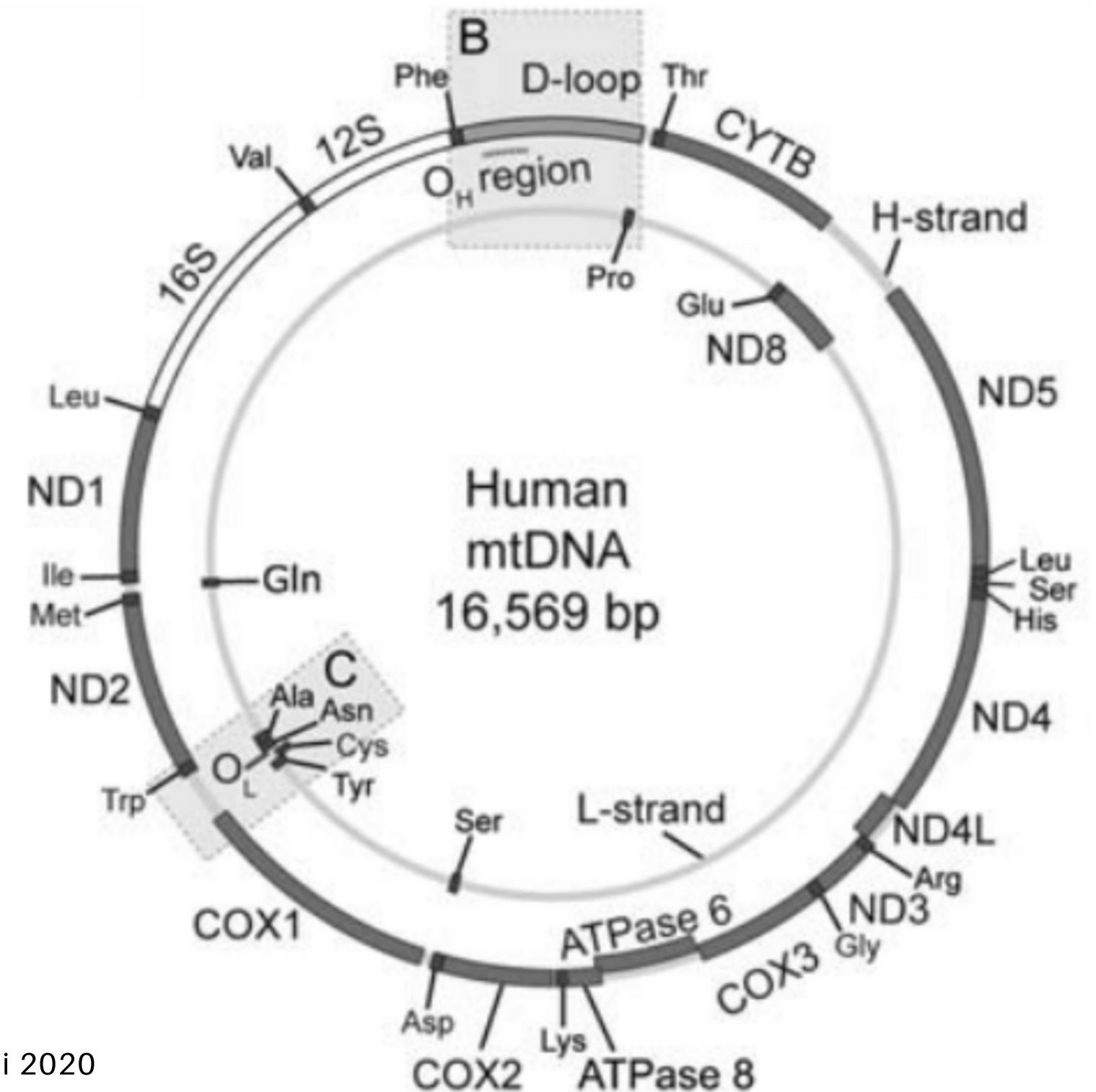
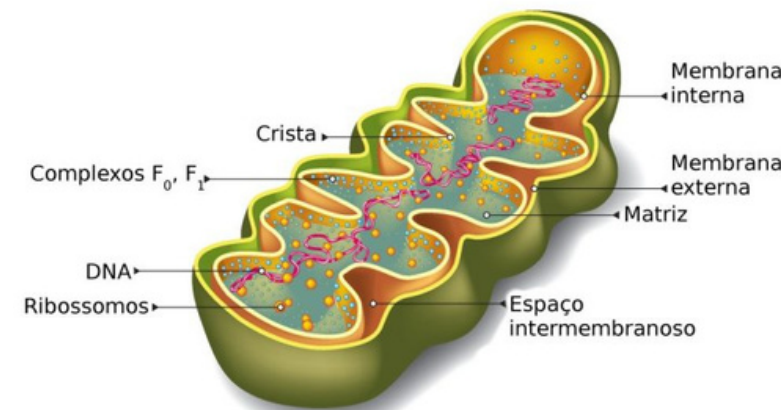
- **DNAmt, cromossomo Y.**
- **Não sofrem recombinação**, a única fonte de diversidade é a mutação.
- Uteis para conhecer **linhagens**.
- O seu estudo mostra a **história parcial de miscigenação** de um indivíduo/população.
- Permitem detectar **vieses sexuais nas contribuições ancestrais** de uma população e outros fenômenos demográficos Ex.: patri ou matrilocidade.



Sistemas polimórficos uniparentais

DNAmt

- **16.569 pb.**
- **Genoma circular, compacto e sem íntrons** (DNA não codificador é apenas 7%).
- **Não possui histonas.**
- Codifica apenas **37 genes**:
 - 13 envolvidos na fosforilação oxidativa.
 - 2 RNAr
 - 22 RNAt.
- Código genético diferente do nuclear.



Fonte: Gasparre and Porcelli 2020



Sistemas polimórficos uniparentais

DNAmt

- **Alto número de cópias** (cada célula somática possui centos de milhares de cópias).
- **Herança materna.**
- Marcadores estudados: **SNPs e indels, sequenciamento completo** e ou apenas da **região de controle**.
- **Velocidade de mutação é 10 X maior** do que a do genoma nuclear. É possível estimar essa velocidade diretamente desde pedigris.
- Velocidade de mutação varia entre as distintas partes do DNAmt, sendo **10 X maior** nas regiões hipervariáveis I e II (HVS I e HVS II) **na região de controle**.
- Envolvimento da mitocôndria na fosforilação oxidativa que produz **radicais livres mutagênicos**.
- **Maior taxa de replicação** que o DNA nuclear.
- Em cada célula há uma coexistência de cópias *wild type* (tipo selvagem) e mutadas = **heteroplasmia** (o oposto é **homoplasmia**). Porém, o indivíduo apresenta um genoma mitocondrial predominante (o mais comum).

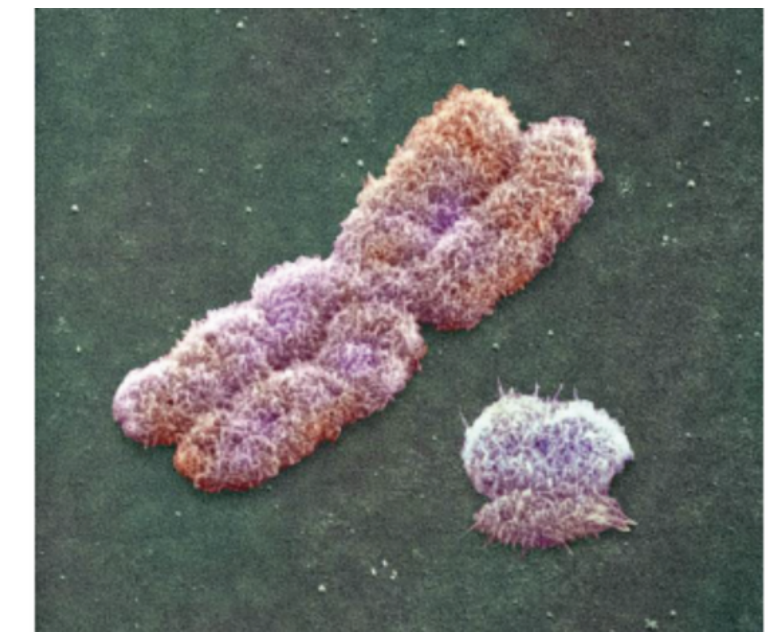
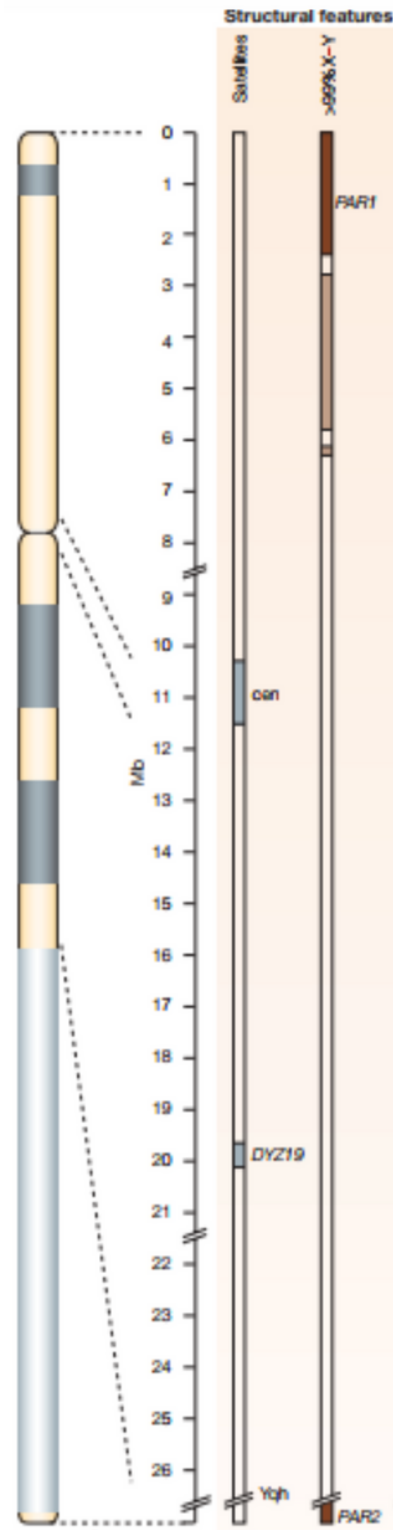


Sistemas polimórficos uniparentais

Cromossomo Y

- Genoma haploide de **60 Mb** (2% do genoma humano haploide).
- ~**95 %** DNA **não recombinante** (NRY = *non-recombining Y*), haplótipos são passados intactos de pai para filho.
- Única fonte de variação é a **mutação**.
- ~**5%** é constituído pelas **regiões pseudoautossômicas (PAR1 e PAR2, teloméricas)** as quais recombina com regiões homólogas no cromossomo X.
- > **50%** de DNA **não codificante**.
- Tendência a **degenerar**.
- ~**200 genes**, ~**55 ativos** principalmente envolvidos com **desenvolvimento sexual e fertilidade masculina**. Alguns possuem homólogos no cromossomo X.
- Inclui o gene **SRY**, determinante do sexo.
- Marcadores estudados: **STRs, Alu, indels, SNPs, sequenciamento completo**.

Fonte: Jobling and Tyler, 2003



Fonte: Stoneking, 2017



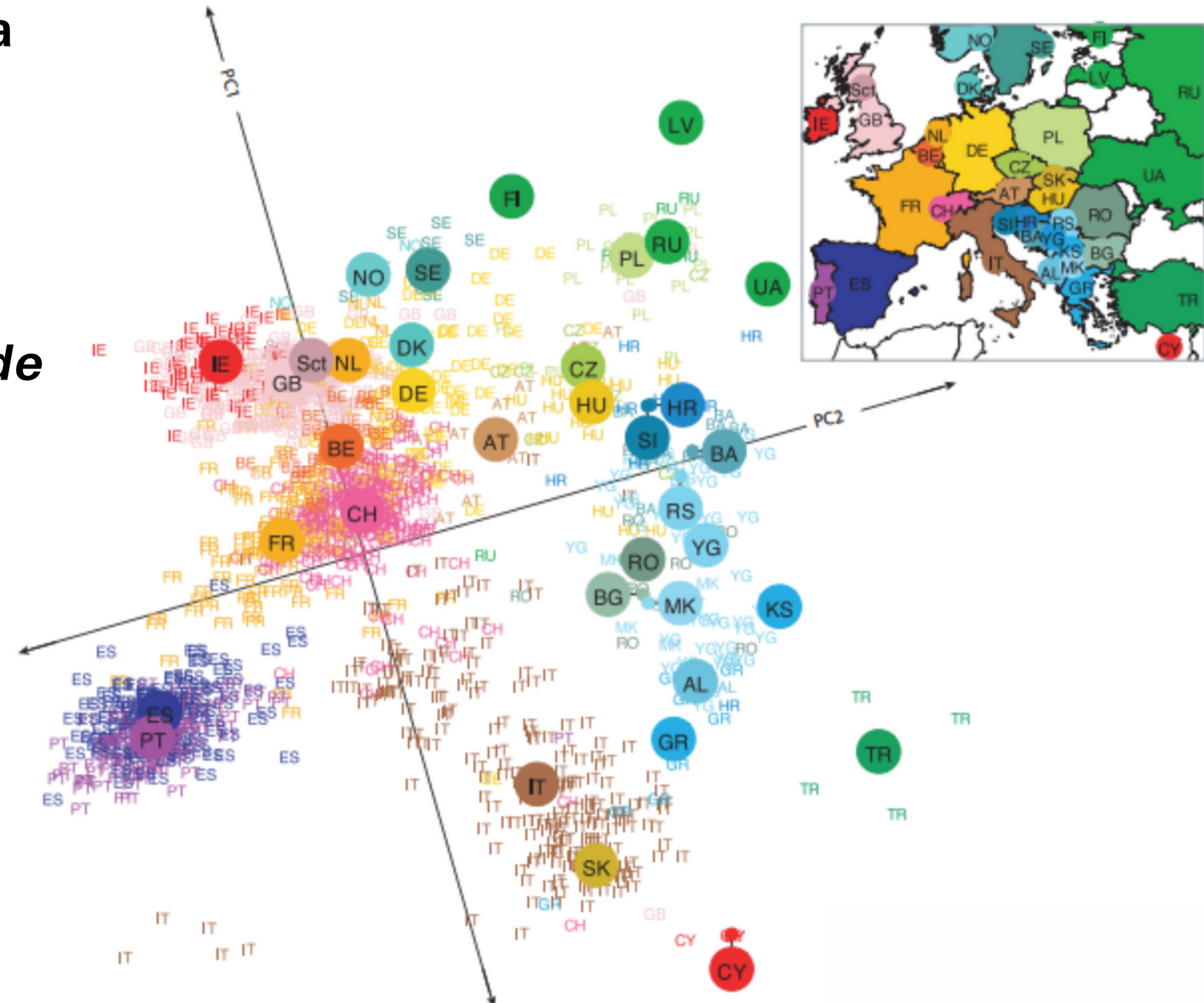
Aplicações na Antropologia



Uso na Antropologia

Marcadores Biparentais

- Estudos de **ancestralidade e estrutura populacional**.
- Marcadores utilizados:
 - Conjunto limitado: **AIMs (ancestry informative markers)**
 - Marcadores genômicos (**Genome-Wide SNPs**)



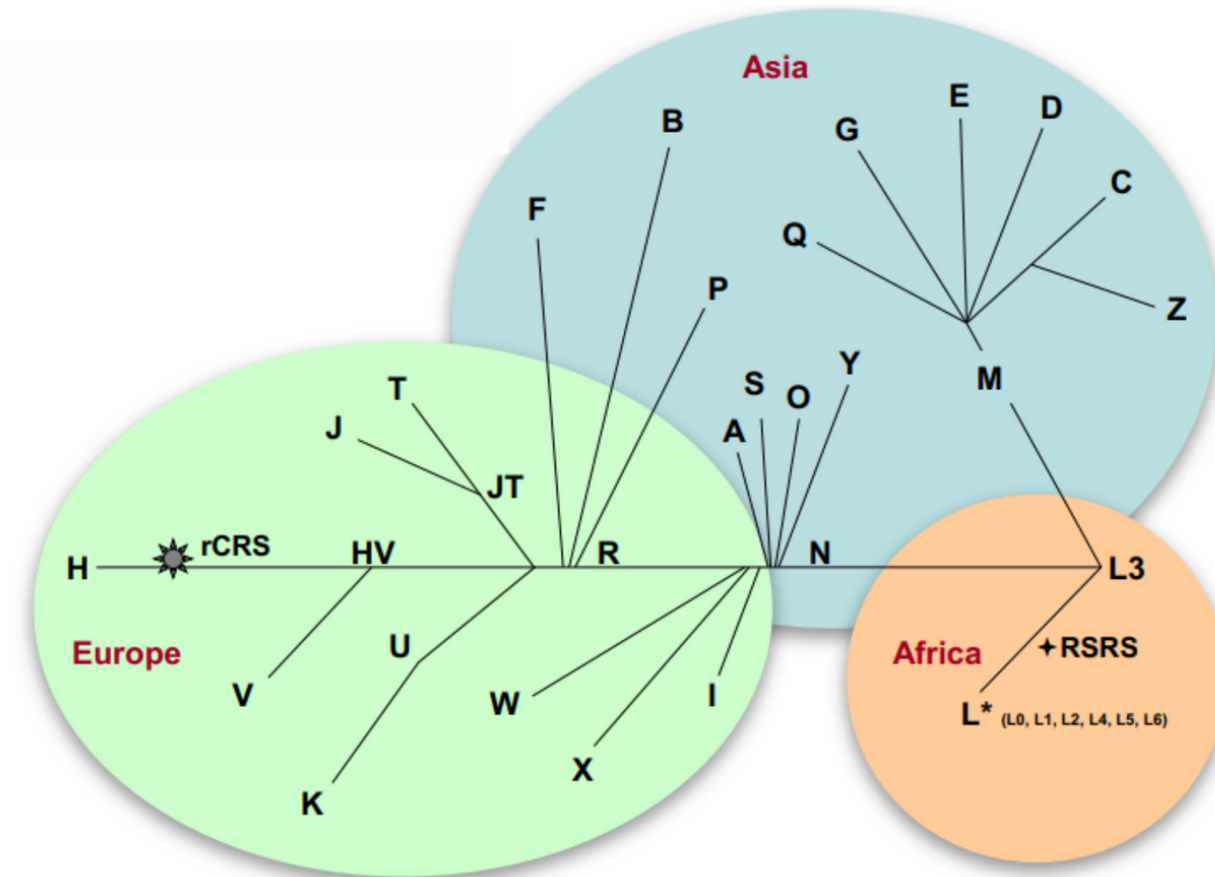
Fonte: Jobling et al., 2014



Uso na Antropologia

DNAm

- Sequenciado em 1981 (antes que o genoma nuclear).
- Estudo de **haplogrupos** = **haplótipo** + **grupo**.
- Membros de um **haplogrupo** compartilham mutações específicas que derivam, por descendência, de um **ancestral comum**.
- Cada haplogrupo forma um **clado** da árvore filogenética.
- Mutações fora das que definem cada haplogrupo determinam maiores **níveis de resolução** do haplótipo.
- Os clados são **geográfico-específicos**.
- Útil para estudos de DNA antigo por ser mais facilmente recuperável de ossos e amostras antigas.



Fonte: mitomap.org

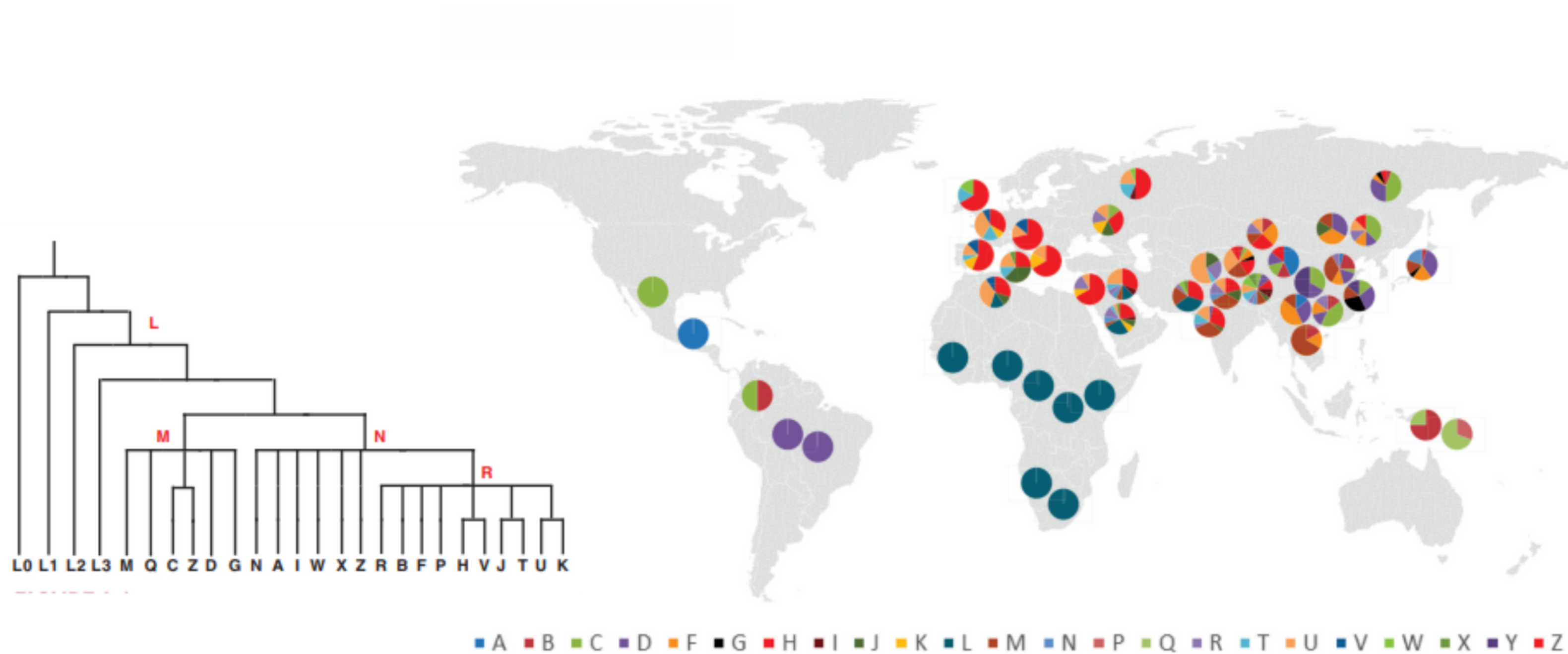


Fonte: eeaf.org



Uso na Antropologia

DNAm



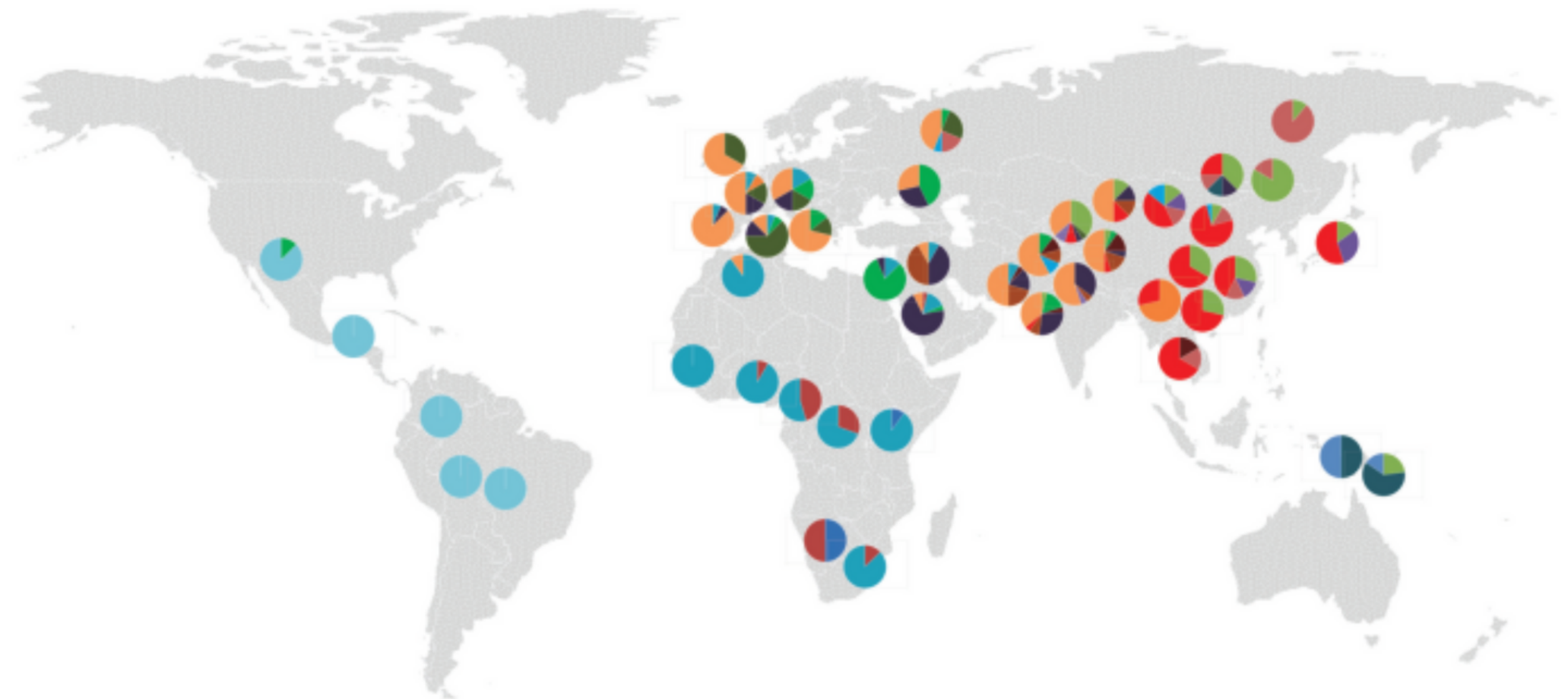
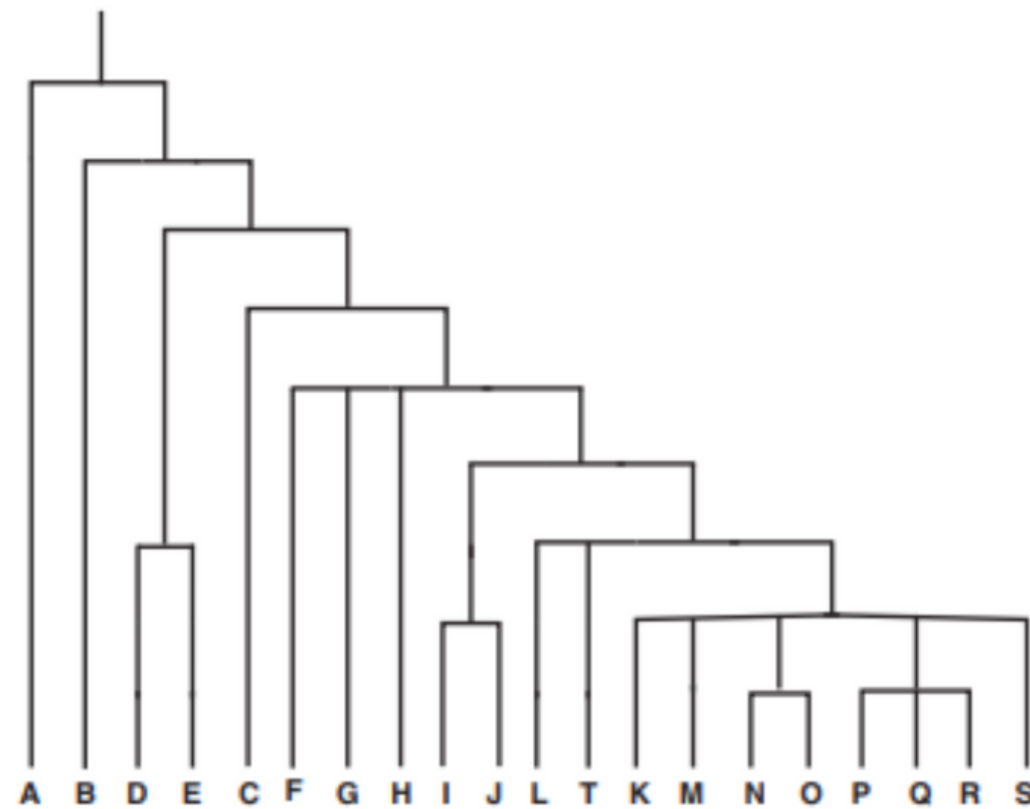
Fonte: Stoneking, 2017



Uso na Antropologia

Cromossomo Y

- O seu uso é similar ao do DNAm, porém mais limitado pela maior dificuldade da sua recuperação (menor número de cópia e maior suscetibilidade a degradação).
- Assim como com o DNAm, **SNPs**, **indels** e **Alu** definem **haplogrupos geográfico-específicos**.
- **STR** permitem adicionar **resolução aos haplogrupos** e analisar **processos demográficos mais recentes**.

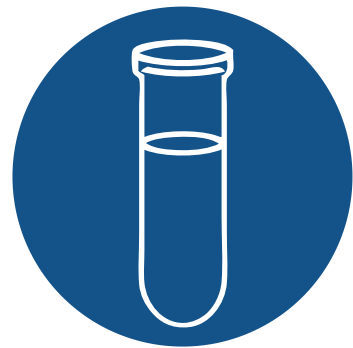


Fonte: Stoneking, 2017

■ A ■ B ■ C ■ D ■ E ■ F ■ G ■ H ■ I ■ J ■ K ■ L ■ M ■ N ■ O ■ P ■ Q ■ R



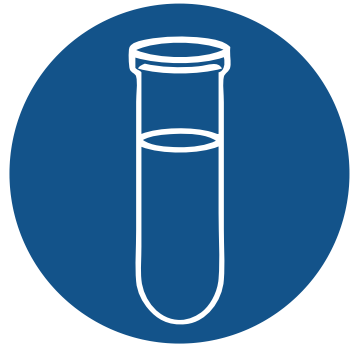
Aplicações na Genética Médica



Uso na Genética Médica: Marcadores Biparentais

- Sem conhecimento prévio: Genome Wide Association Studies ou GWAS (SNPs).
- Estudos de genes candidato (todos os tipos de marcadores).
- Sequenciamento de genoma e sequenciamento de exoma (todos os tipos de marcadores).





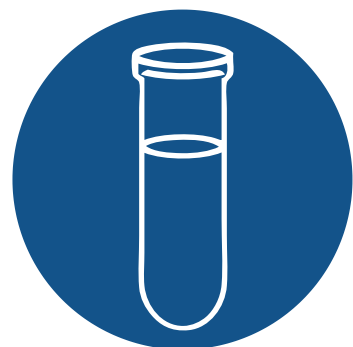
Uso na Genética Médica: Marcadores Uniparentais

DNAmt

Graças à heteroplasma, **mutações patogênicas** podem estar presentes em níveis baixos sem afetar a saúde (nível de mutação $<1\%$).

O aumento da frequência de uma mutação patogênica pode produzir **disfunção mitocondrial** afetando principalmente **tecidos com alto requerimento de ATP** como músculo esquelético e cérebro.

A **severidade de doenças** de origem mitocondrial depende do nível de heteroplasma. Existem estudos de **associação de haplogrupos com doenças**, possivelmente pela sua relação com a **ancestralidade**.



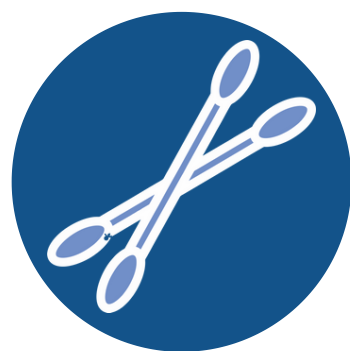
Uso na Genética Médica: Marcadores Uniparentais

Cromossomo Y

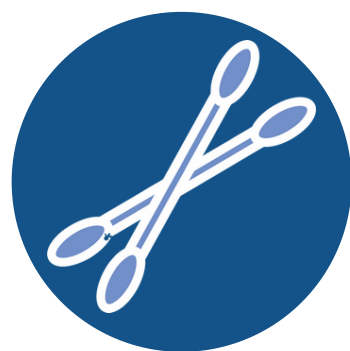
Os estudos biomédicos do cromossomo Y tem associado marcadores principalmente a doenças envolvendo **fertilidade, hipertensão, câncer de próstata e testicular.**

Table 1 | **Association studies using the Y chromosome**

Trait	Population	Y marker(s)		Association	Rep.	References
		Binary	Multi			
XX maleness	Europeans	2	MSY1	XX maleness with haplogroup Y*(xP)	–	34
Infertility/ sperm count	Japanese	3	–	Infertility and low sperm count with haplogroup D	N	31,32
Infertility	Italians	8	–	None	–	33
Infertility	Europeans	11	–	None	–	120
Infertility	Europeans	9	–	None	–	121
Infertility	Japanese	14	αH	None	–	32
Sperm count	Danes	3	–	Low sperm count with haplogroup K(xP)	–	35
Sperm count	Italians	11	–	None	–	122
Alcohol dependence	Finns	1	7	Alcohol dependence with three different lineages	–	123
Height	European-Australians	1	–	Haplogroup P 1.9 cm taller	–	124
Blood pressure	European-Australians	1	–	High blood pressure with haplogroup P	N	125,126
Blood pressure	Polish, Scots	1	–	High blood pressure with haplogroup Y*(xP)	–	125
Blood pressure	Japanese	1	–	None	–	127
Prostate cancer	Japanese	–	1	Prostate cancer with one microsatellite allele	–	128
Testicular cancer	English	7	6	None	–	129
Longevity	Sardinians	14	–	None	–	130
Autism	Norwegians, Swedes, French	10	–	None	–	131



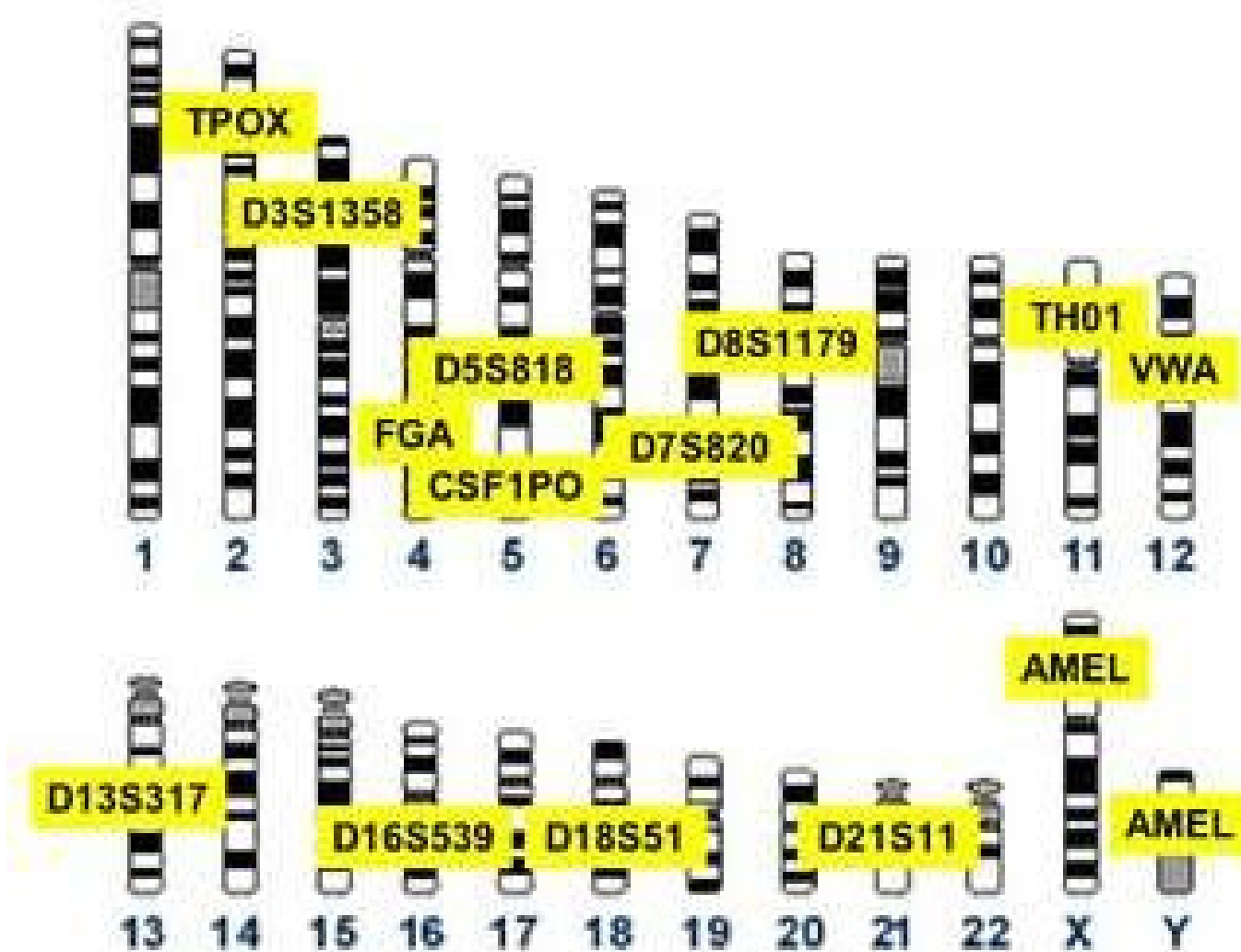
Aplicações na área forense



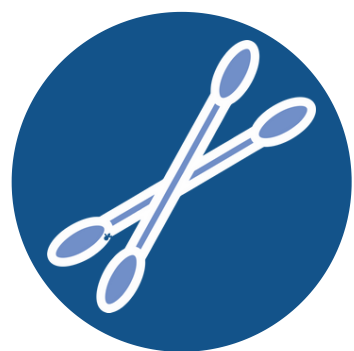
Uso Forense

Marcadores biparentais

- Utilizados para identificação de suspeitos de ambos os sexos. É analisado um conjunto de **STRs** autossômicos distribuídos de forma a representar o genoma do indivíduo.
- Atualmente existem **kits** otimizados para uso forense.



Fonte: Arruda, 2017

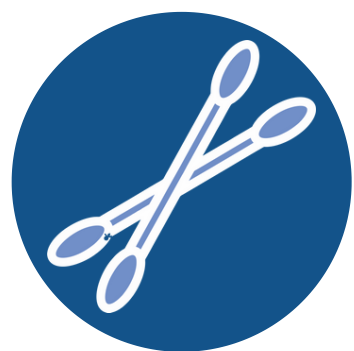


Uso Forense DNAmt

- Especialmente útil na identificação a partir de espécimes forenses com DNA degradado (ossos, cabelo, dente).
- Mais estável por ser circular e estar mais protegido da degradação.
- **Não serve para identificação** de pessoas.
- Compartilhado por todos os parentes por linha materna (o DNAmt pode ser usado na **confirmação da identidade** da pessoa desaparecida).
- Utilizado na identificação de desaparecidos e investigação de desastres em massa.

- Risco de **contaminação**.

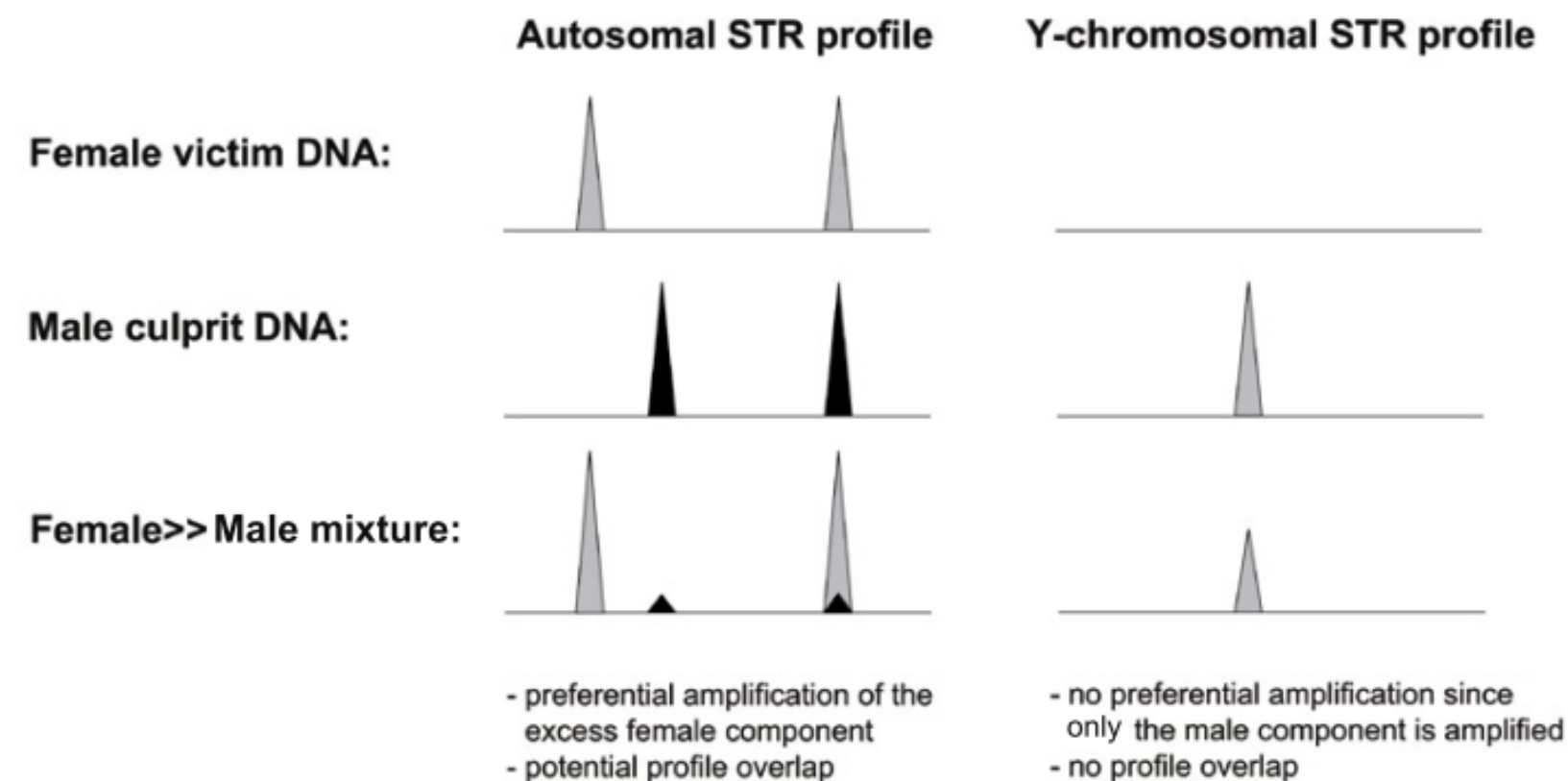




Uso Forense

Cromossomo Y

- Contribui na identificação de suspeitos masculinos através de um perfil de **STR** (Ex.: estupro).
- **Não serve para identificação** de pessoas, mas permite a exclusão de suspeitos.
- Compartilhado pelos parentes por linha paterna.
- Utilizado na sexagem de amostras pela amplificação de **genes específicos do Y** (Ex.: amelogenina) ou **indels**.
- Estimação de **paternidade**.



Fonte: Kayser 2007