



# Análise de haplótipo e ancestralidade

Prof<sup>a</sup> Dr<sup>a</sup> Sântia Iole Belangero

Curso de Introdução à Bioinformática Aplicada à  
Genética



# Haplótipo

Definição:

- conjunto de alelos de um mesmo locus, os quais são, preferencialmente, segregados juntos.





# Haplótipos

- . Alguns mais raros, outros mais frequentes
- . Pode ser raro em uma população e comum em outra
- .  $3 \times 10^7$  combinação de haplótipos de HLA que poderiam ocorrer nunca foram vistos em caucasianos

Por que???





# Por...

- . Desequilíbrio de ligação
  - . baixas taxas de recombinação meiótica
- . Influências ambientais (pressão seletiva positiva para algumas combinações)
- . Fatores históricos (quanto tempo a população foi estabelecida, imigrações)



# Ligação

- . Tendência que os alelos, em um *locus* do mesmo cromossomo, têm de serem transmitidos como uma unidade íntacta na meiose
- . Quando ocorre uma taxa de 1% de crossing over na distância entre dois alelos = 1 cM



# Equilíbrio de Ligação

Alelos em equilíbrio de ligação

- . Quando a frequência de cada alelo dentro do haplótipo é = frequência daquele alelo na população
- . Taxa de recombinação alta = equilíbrio de ligação



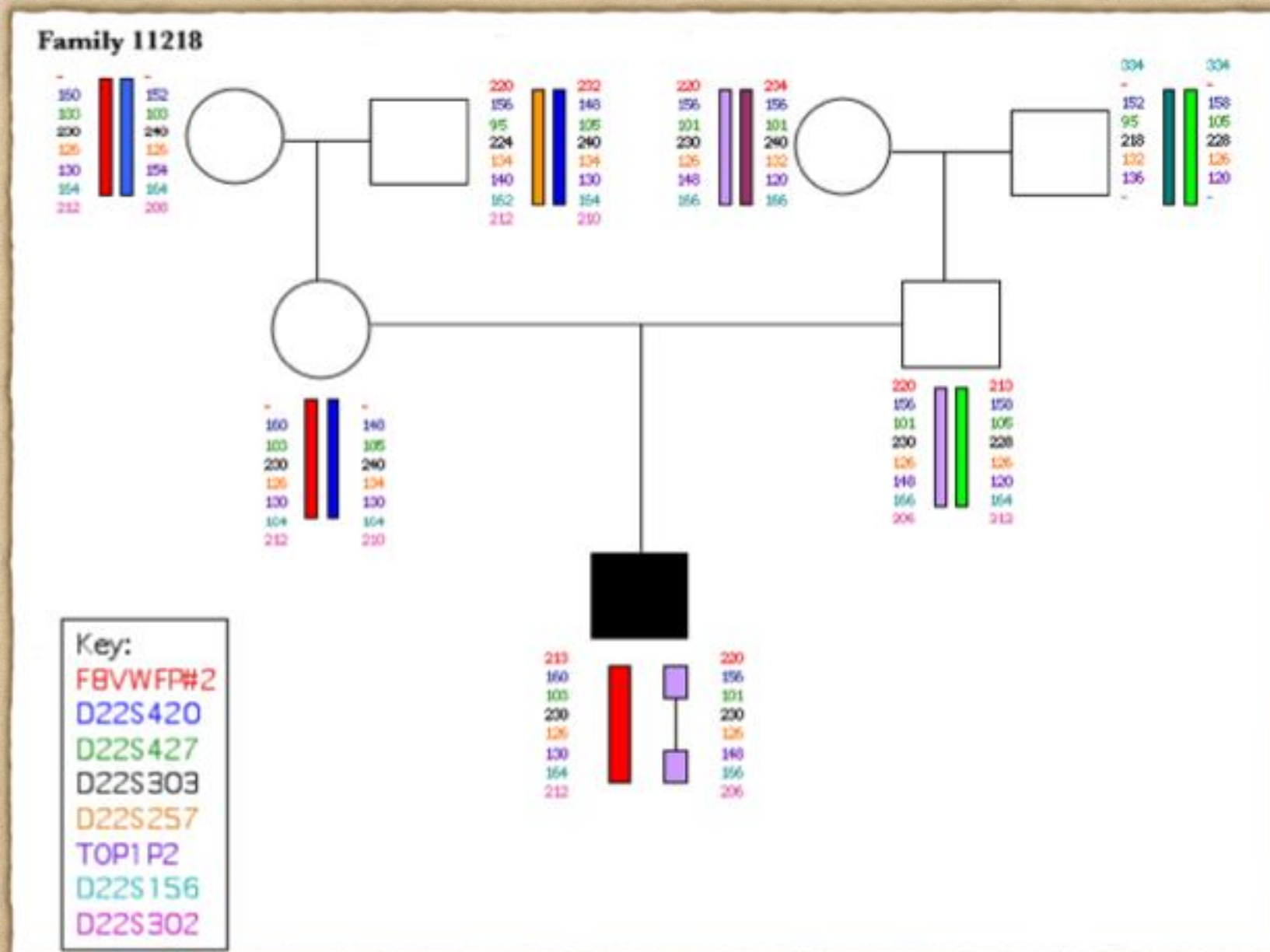
# Desequilíbrio de Ligação

Alelos em desequilíbrio de ligação

- . Quando a frequência de cada alelo dentro do haplótipo é diferente da frequência daquele alelo na população
- . Taxa de recombinação baixa = desequilíbrio de ligação

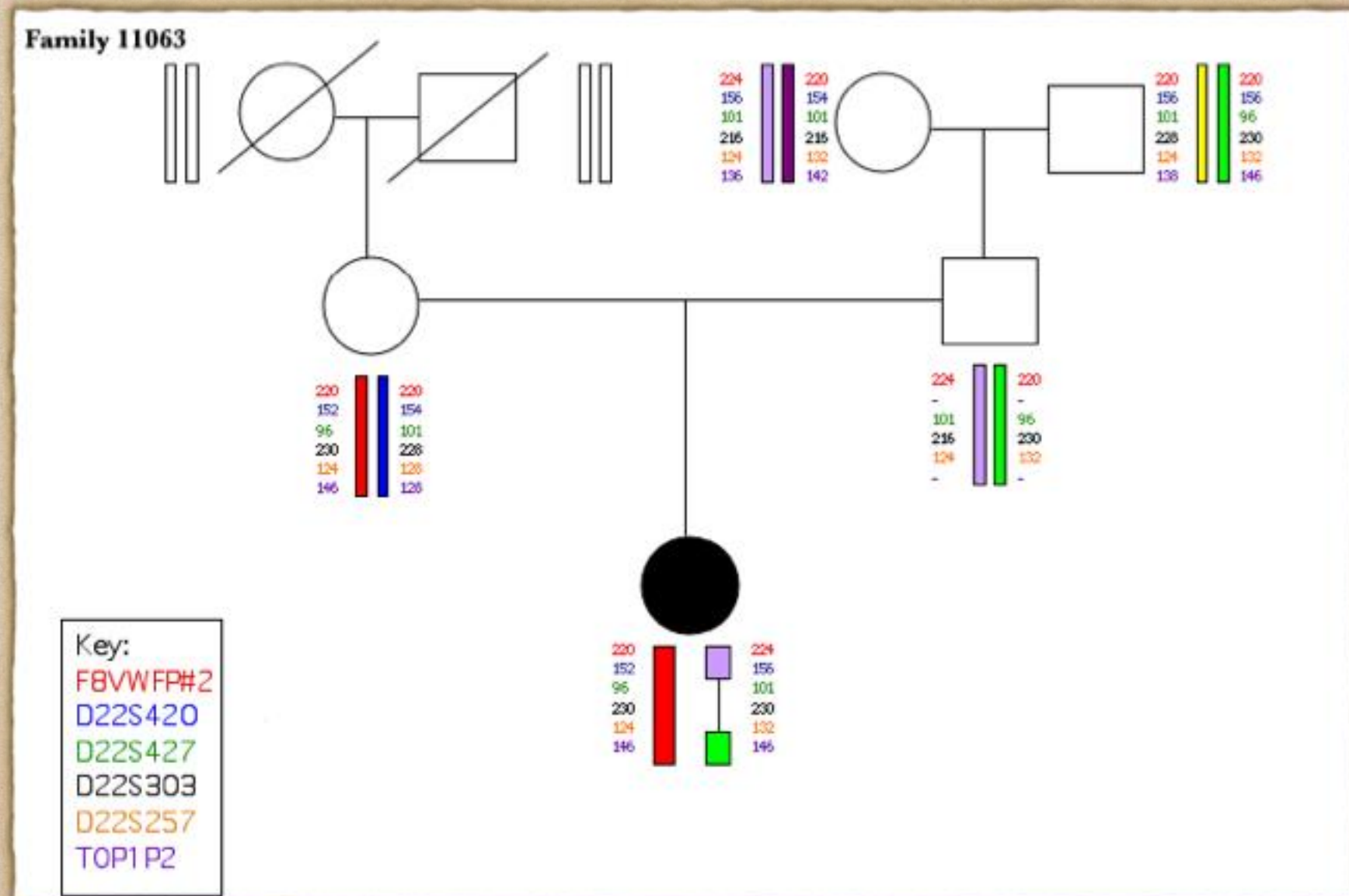


# Sem recombinação meiótica





# Com recombinação meiótica





# Estratificação populacional

## Definição

- . Presença de uma diferença sistemática na frequência alélica entre subpopulações de uma população devido à diferença de ancestralidade
- . Estrutura de uma população
- . Importante, principalmente no contexto de estudos de associação



# Estratificação populacional

## Métodos de avaliação

- AIMS (marcadores informativos de ancestralidade)
- PCA/MDS





# Marcadores informativos de ancestralidade (AIMs)

- Um marcador informativo de ancestralidade (AIM) é um polimorfismo (ou um conjunto deles) de um locus que exibe substancialmente diferentes frequências entre populações de diferentes regiões geográficas.
- Com AIMs podemos estimar a origem geográfica dos ancestrais e determinar a proporção de ancestralidade de cada região geográfica.



# Marcadores informativos de ancestralidade (AIMs)

- . Foram seleccionados dentre os 15 milhões de SNPs, aqueles que fossem úteis como AIMs
- . Ex.: alelo nulo Duffy (FY\*o)
  - . tem uma frequência de quase 100% dos africanos sub-saharianos
  - . ocorre muito raramente em populações fora desta região
  - . uma pessoa que tenha este alelo é, portanto, mais propenso a ter antepassados da África Sub-Sahariana



# Marcadores informativos de ancestralidade (AIMs)

- . Existe uma tendência à formação de cluster (pelas frequências alélicas) onde as populações vivem juntas e interagem ao longo do tempo
- . Isto é devido à endogamia dentro de grupos, ou fronteiras nacionais, culturais ou linguísticas



# Marcadores informativos de ancestralidade (AIMs)

- . Embora as diferenças genéticas entre os grupos humanos sejam pequenas, essas diferenças em determinados genes (ex. Duffy, ABCC11, SLC24A5), podem ser usadas para localizar, de maneira confiável, indivíduos dentro de clusters com base geográfica.
- . Esses clusters genéticos estão associados com grupos que historicamente têm ocupado determinadas regiões (Rosenberg et al 2002; Bamshad et al 2003)



# Métodos de análise de AIMs

- . STRUCTURE é um software para gerar clusters gênicos
- . Baseia-se na formação de cluster genético:
  - . dentro de um cluster as pessoas compartilham frequências alélicas médias entre si muito mais do que com as pessoas de outros clusters
- . O programa coloca indivíduos dentro de um número arbitrário de clusters baseado na sua similaridade genética



# Estratificação populacional por AIMs na população brasileira

Dados são do grupo do Prof Sidney Santos  
Lab. de Genética Humana e Médica  
UFPA

Santos et al., Assessing Individual Interethnic Admixture and Population Substructure Using a 48-  
Insertion-Deletion (INDEL) Ancestry-Informative Marker (AIM) Panel  
Human Mutation



# Estratificação populacional na população brasileira

- . objetivo: medir as proporções de 3 diferentes ancestralidades (africana sub-sahariana, europeia e nativo-americana) em populações misturadas
- . seleção de INDELs (polimorfismos de inserção-deleção) informativos para ancestralidade



# Estratificação populacional na população brasileira

- . Pré-seleção de 60 INDELs (AIMs) foram baseados em 3 critérios:
  - . (1) grande diferença de frequência alélica ( $d > 40\%$ ) entre africanos, europeus e nativo-americanos
  - . (2) estar em diferentes cromossomos ou em diferentes regiões físicas do mesmo cromossomo
  - . (3) tamanho variável entre 3-40 bp para permitir genotipagem simultânea de múltiplos marcadores (eletroforese de capilar)



# Estratificação populacional na população brasileira

- . 593 indivíduos de origem conhecida daqueles 3 grupos ancestrais:
  - . 189 sub-saharianos (Angola, Moçambique, Zaire, Camarões, Costa do Marfim)
  - . 161 europeus (principalmente portugueses)
  - . 243 nativos-americanos (tribos da Amazônia brasileira)
- . 280 de populações misturadas
  - . 81 do Sul do Brasil
  - . 96 Belém
  - . 103 de comunidades afro-descendentes que vivem na Amazônia



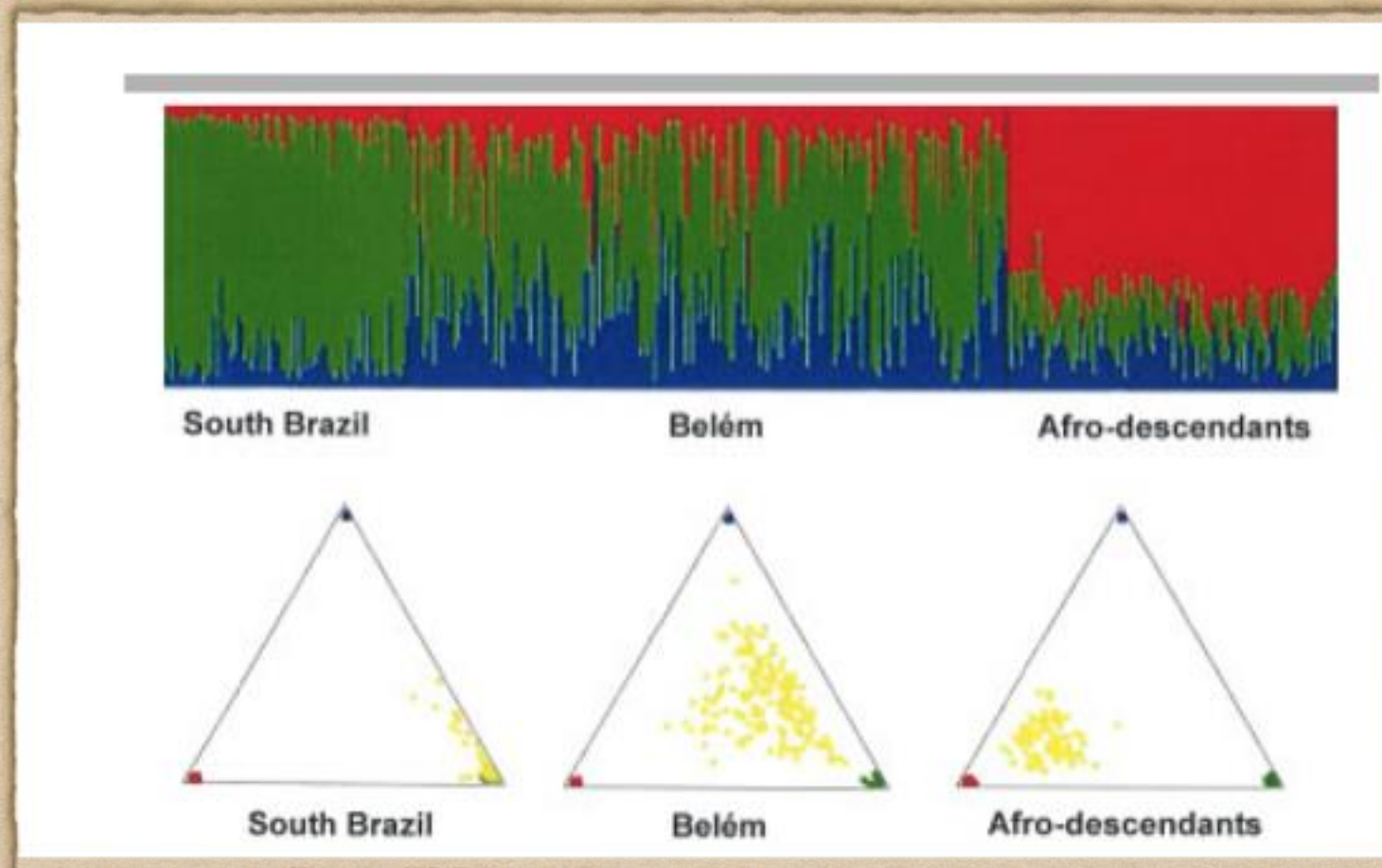
# Estratificação populacional na população brasileira

## Resultados

- . Indivíduos do sul do Brasil: quase ancestralidade europeia exclusiva
- . Comunidades afro-descendentes na região da Amazônia: > contribuição Africana com algum grau de mistura de europeus e nativo-americanos
- . Belém: contribuição significativa das 3 etnias, embora com proporção maior de europeia



# Estratificação populacional na população brasileira



- Plot de barras: Cada linha vertical representa um indivíduo e sua proporção de europeu (verde), africano (vermelho) e ameríndio (azul).
- Plot de triângulo: cada indivíduos está representado por um ponto colorido e a proporção correspondente da mistura estão indicados pela distância dos vértices do triângulo.
- Verde, vermelho e azul corresponde a indivíduos de populações acima e indivíduos de populações misturadas estão em amarelo.



# Estratificação populacional na população brasileira

**Table 2. Global Interethnic Admixture Estimates in Three Mixed Populations\***

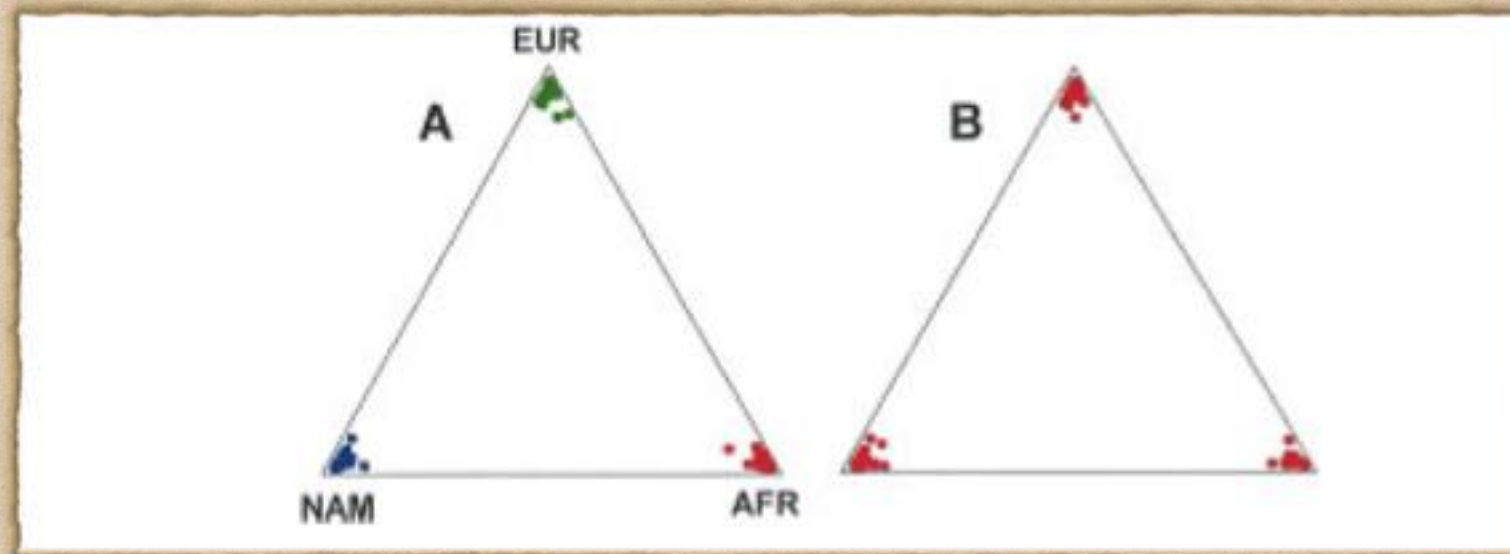
Software	South Brazil			Afro-descendants			Belém		
	AFR	EUR	NAM	AFR	EUR	NAM	AFR	EUR	NAM
ADMIXMAP	3.3	89.2	7.5	72.0	14.2	13.8	14.4	57.3	28.3
STRUCTURE	1.0	95.0	4.0	80.7	9.3	10.0	11.7	61.4	26.9
ADMIX95	0.0	91.3	8.7	69.0	14.9	16.1	14.8	54.6	30.6
ADMIX2.0	0.0	100.0	0.0	79.9	0.0	20.1	6.5	65.9	27.6

\*Percentages (%) of AFR, EUR, and NAM.

AFR, African; EUR, European; NAM, Native American.



# Estratificação populacional na população brasileira



- Seleção de 16 marcadores de ancestralidade africana (que apresentam um valor de alto 'd' entre africanos e europeus; 16 marcadores europeus e 16 marcadores nativo americano.
- Representação do painel dos 48 AIMS - distinção dos indivíduos (STRUCTURE)
- Resultados foram obtidos com e sem informação a priori da origem dos indivíduos



# Estratificação populacional

## Métodos de avaliação

- AIMS (marcadores informativos de ancestralidade)
- MDS/PCA

