



# VARIAÇÕES NO NÚMERO DE CÓPIAS (CNVs)

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

Curso de Introdução à Bioinformática Aplicada à Genética

# HAPMAP

- Comparar genomas de diferentes populações para catalogar variantes humanas
- Identificação de SNPs



# RESULTADOS DO HAPMAP

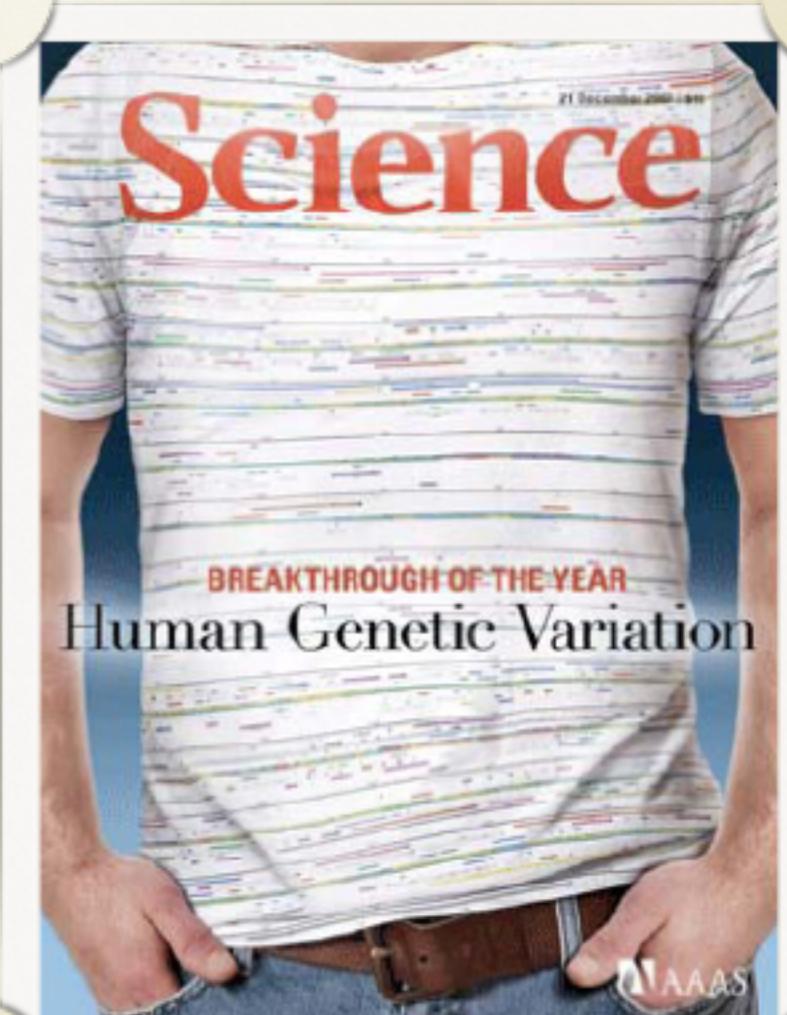
- SNPs não eram a principal fonte de variação genética
- Revelou uma inesperada “variação” no genoma
- Copy Number Variation (CNV)
  - compreendem a maior fração de variantes nos humanos
  - envolvem milhões de nucleotídeos
  - heterogeneidade entre cada genoma (diversidade humana e susceptibilidade a doenças)

# RESULTADOS DO HAPMAP

- Dados encontrados dos SNPs no HapMap
  - 15 milhões de sítios no genoma onde eles podem ocorrer
  - Cada pessoa tem 3,5 milhões de SNPs herdados e 30 novos que surgem das replicações

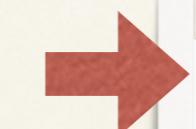
# RESULTADOS DO HAPMAP

- Dados de CNVs do HapMap
  - 1447 catalogadas em 270 indivíduos (260 milhões de bases)
  - 12% genoma está suscetível a variação no número de cópias
  - Cada pessoa tem 1500 CNVs herdadas (média de 20kb)
  - CNVs podem surgir na meiose
  - Humanos 99,5% de similaridade

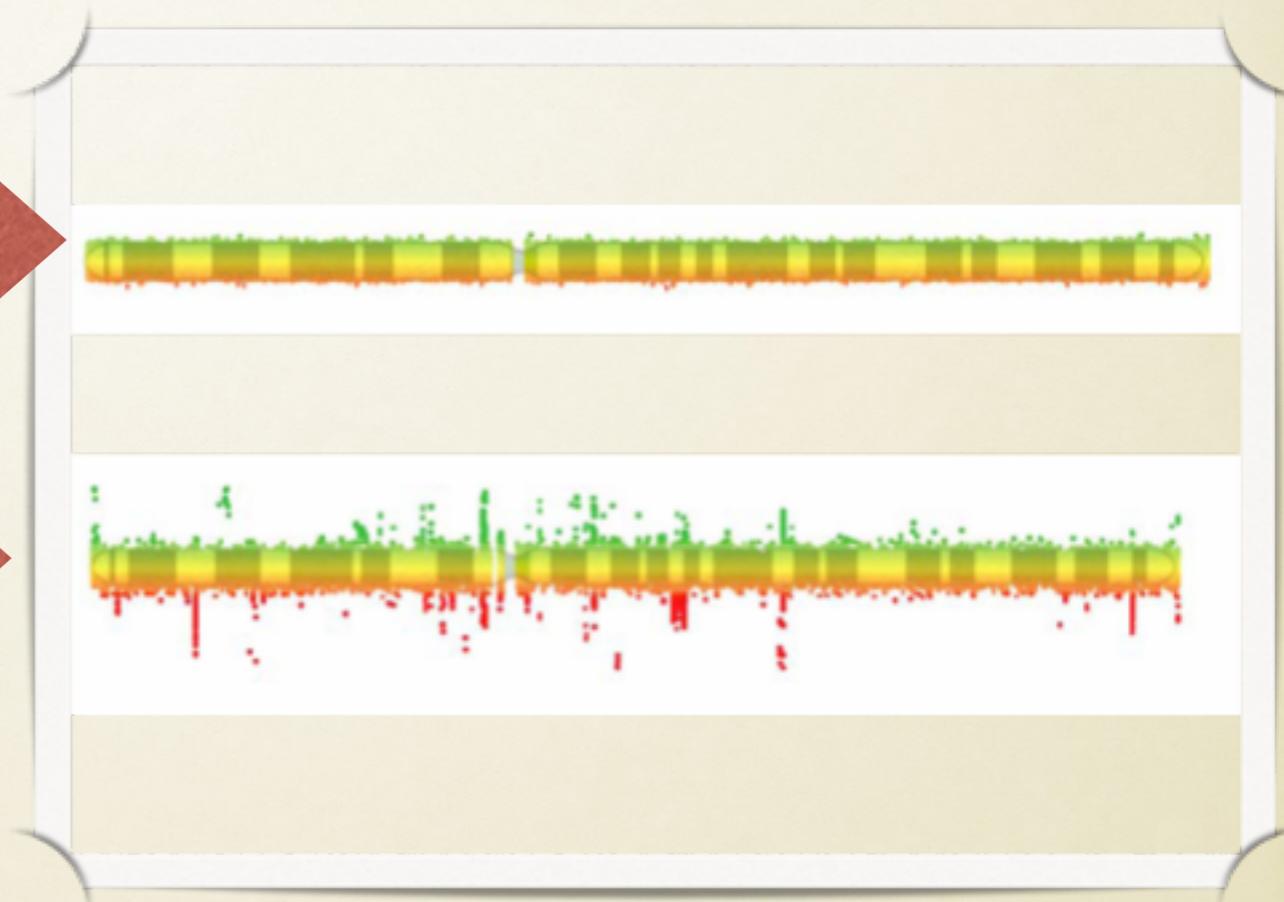


# VARIANTES GENÔMICAS

Se genoma diferisse só por SNP



Mas...



# TIPOS DE CNV

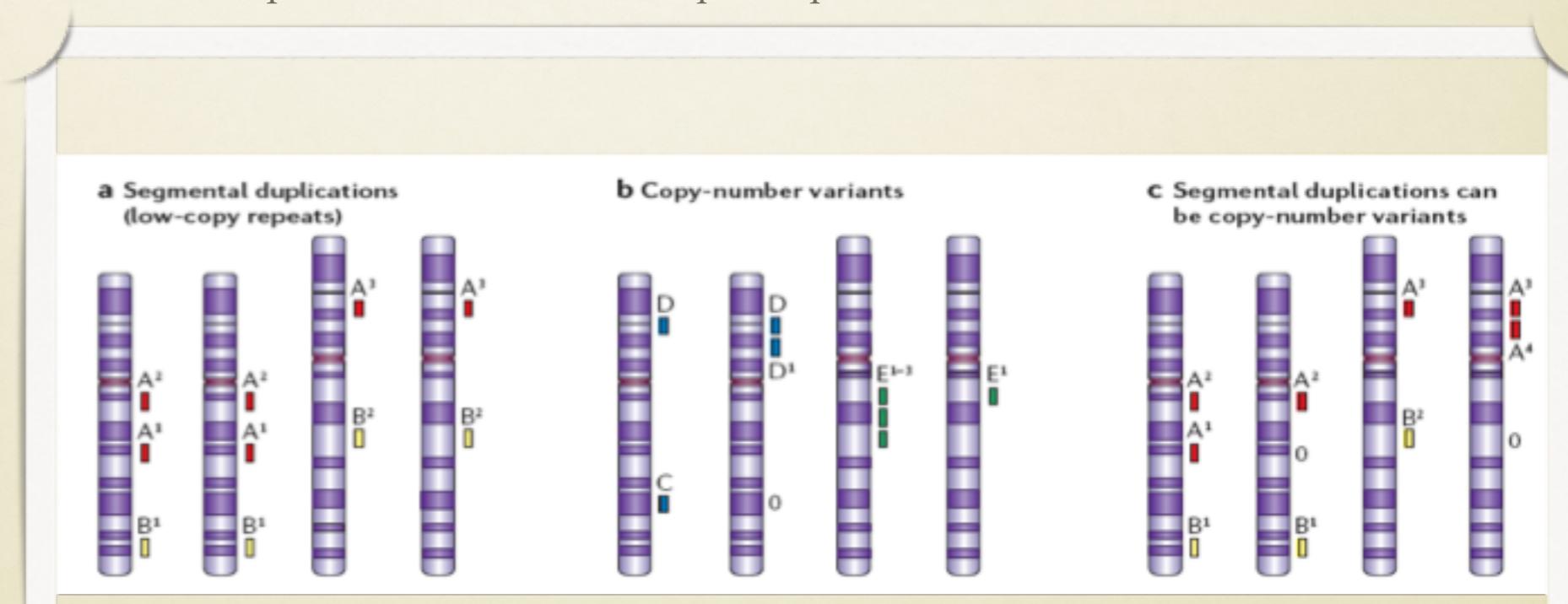
- **Copy-number variant (CNV)**
  - >1 kb – variam quanto ao número de cópias comparado com genoma referência

- **Copy-number polymorphism**

- CNV em mais de 1% da população.

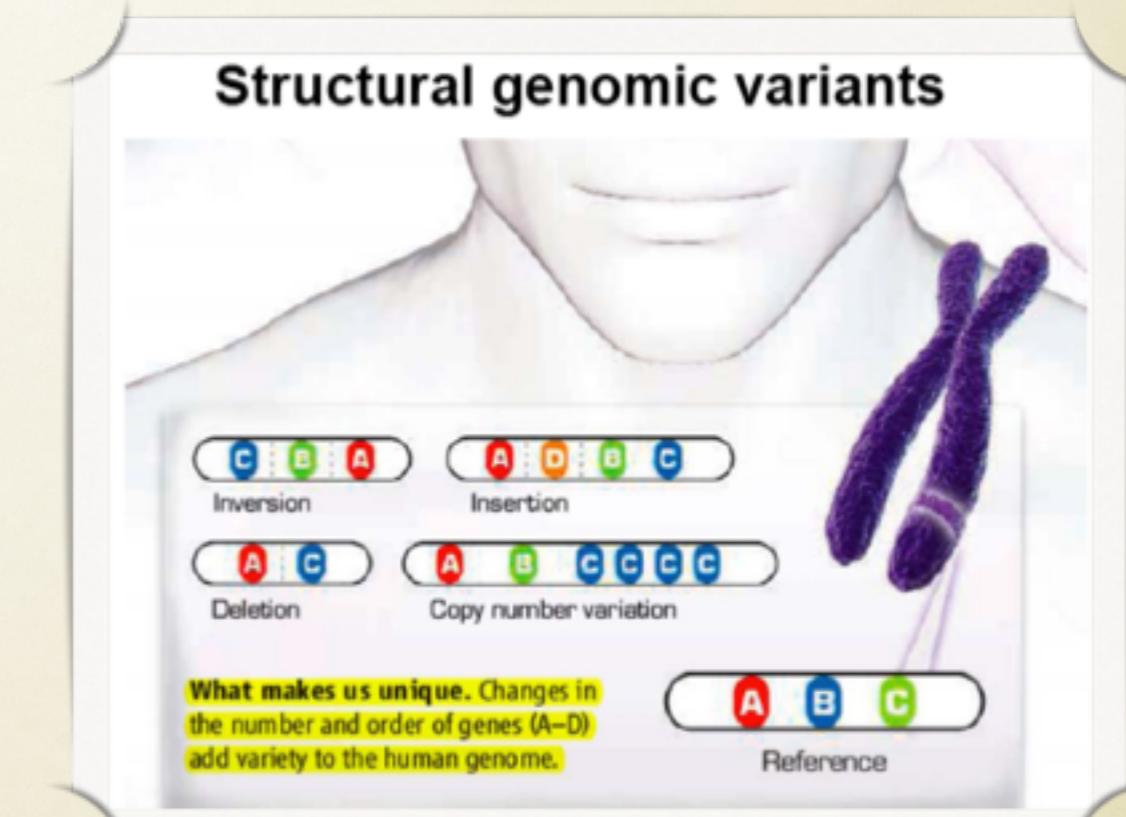
- **Duplicação segmental ou low-copy repeat (LCR)**

- > 1 kb que apresenta 2 ou mais cópias por genoma haploide
  - Identidade de >90% das sequências de DNA.
  - Elas são frequentemente variáveis quanto ao número de cópias e podem ser também CNVs.



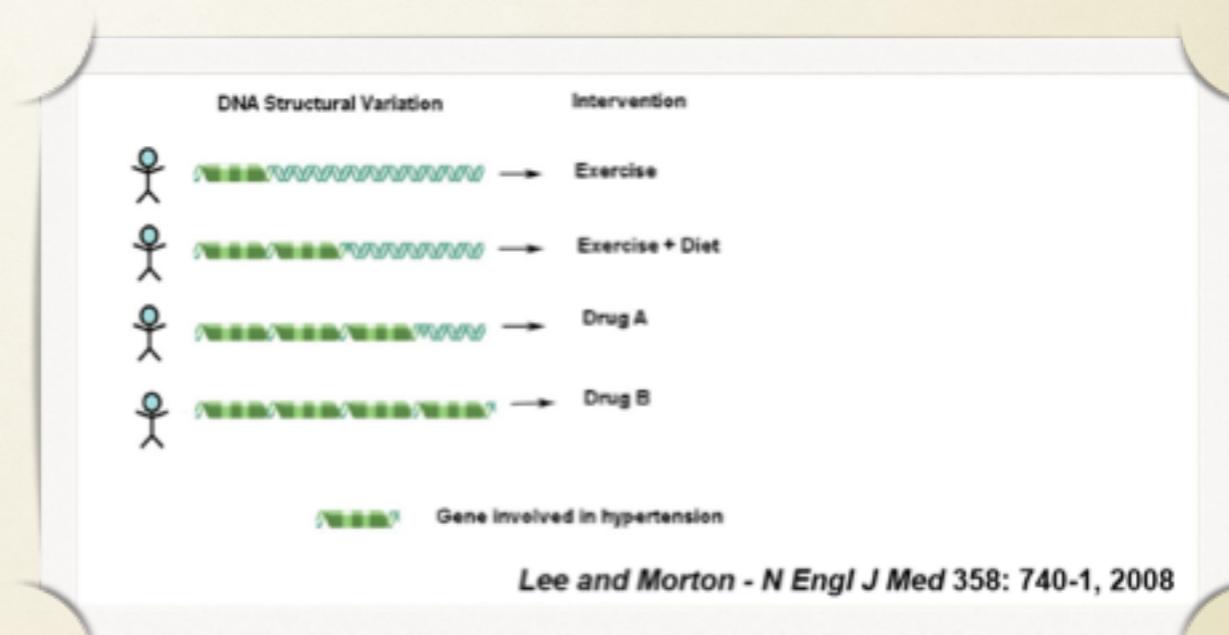
# IMPORTÂNCIA FENOTÍPICA DAS CNVs

- CNVs - podem ter efeitos fenotípicos
  - (ex: microdeleções e duplicações que causam doenças)
- ‘Normais’ X ‘Causadoras de doenças’
- De novo X Herdada
- Estudos populacionais (frequência da variante)



# IMPORTÂNCIA MÉDICA DAS CNVs

- Testes clínicos
- Genética de população e evolução
- Farmacogenômica e intervenções clínicas
  - Resposta imune e inflamação
  - Susceptibilidade a doenças complexas: diabetes, cardiovasculares, câncer e esquizofrenia
  - Resistência a infecções

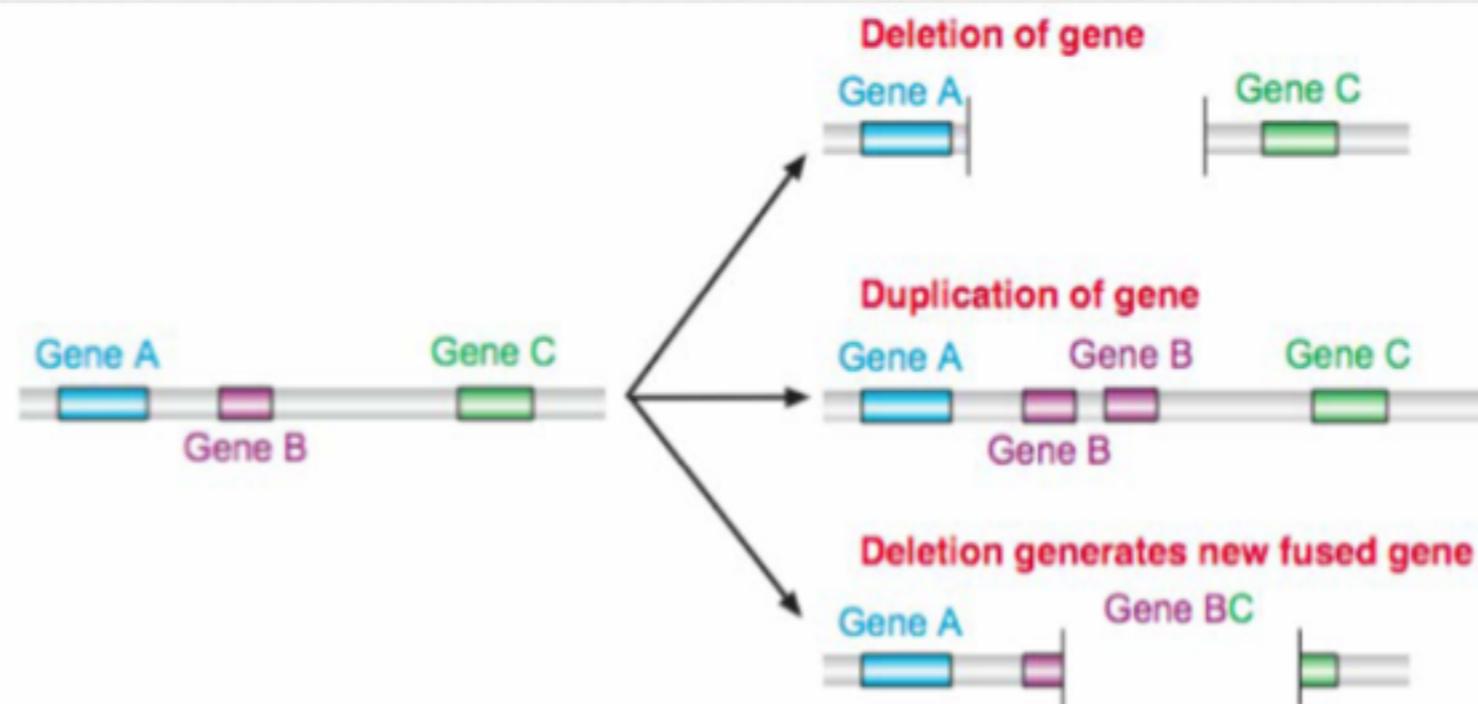


# IMPACTO DAS CNVs

- Muitas estão localizadas fora de genes
  - explica parte da heterogeneidade da espécie humana
- Quase metade das CNVs está em genes
  - a origem de muitas doenças (inclusive mendelianas)



# CNVs E EXPRESSÃO GÊNICA

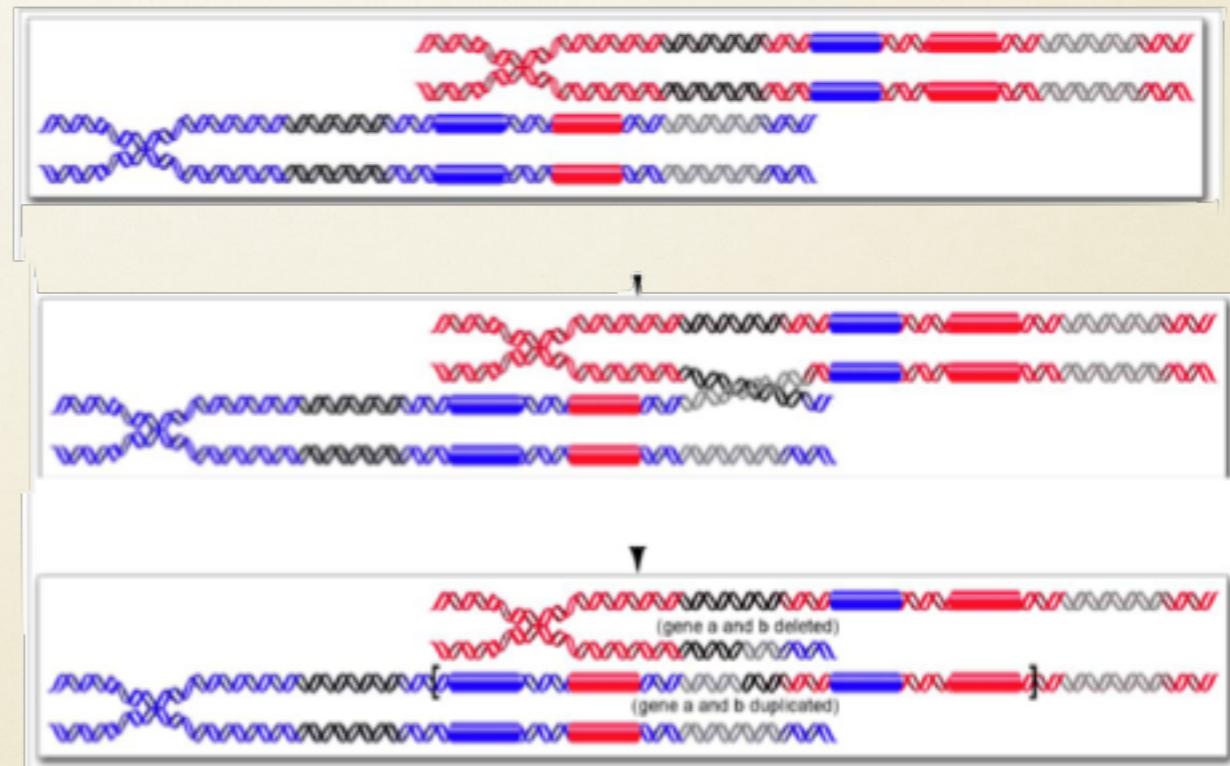


*Courtesy of Dr. Matthew Hurles, Sanger Center, UK*

# MECANISMOS DE FORMAÇÃO DAS CNVs

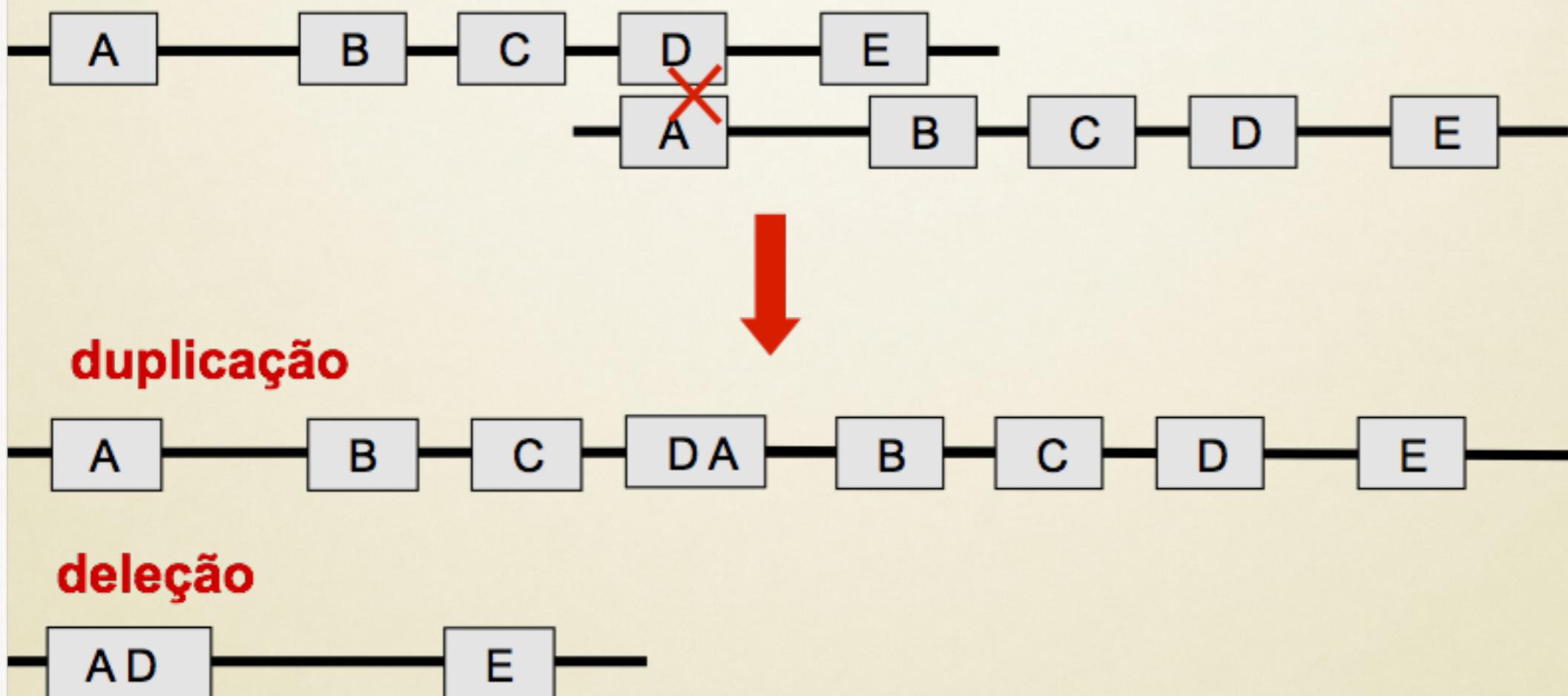
- Recombinação homóloga não alélica

(Non allelic homologous recombination - NAHR)



Lee et al. in Genomics and Personalized Medicine (In Press)

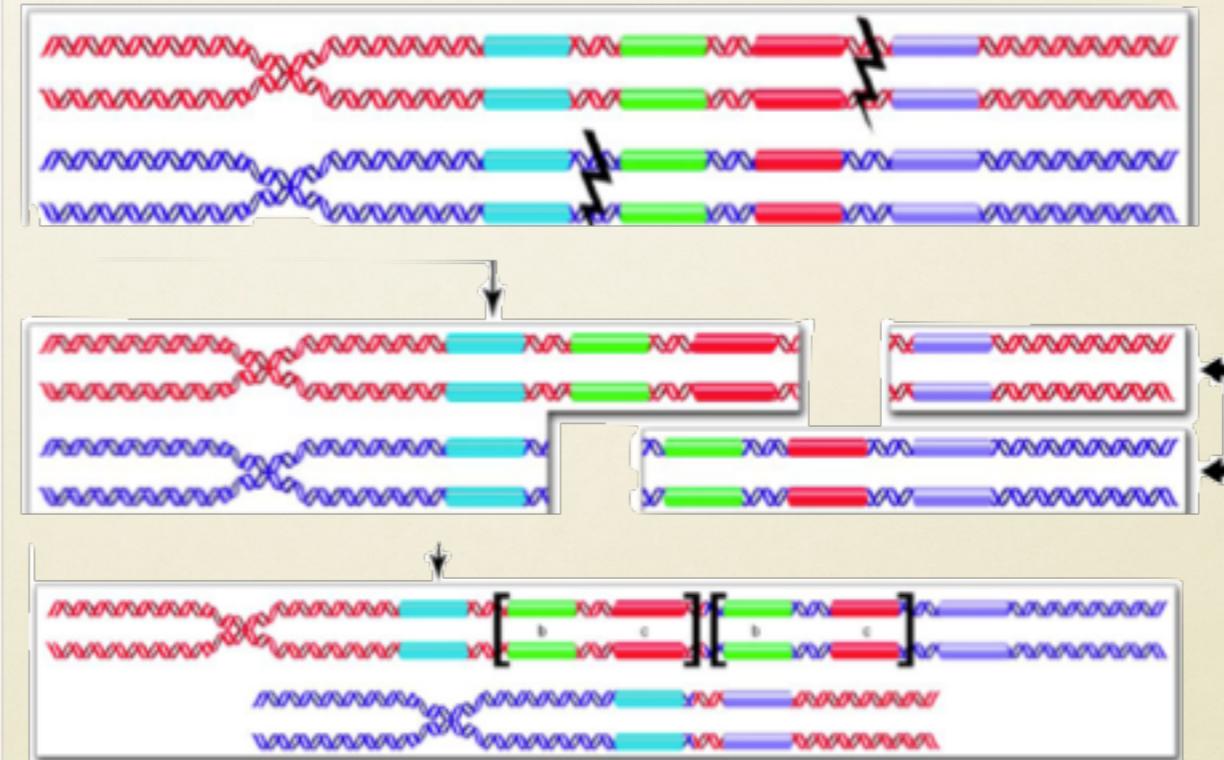
# RECOMBINAÇÃO HOMÓLOGA NÃO ALÉLICA



# MECANISMOS DE FORMAÇÃO DAS CNVs

- Junção de terminação não homóloga

(Non homologous end joining - NHEJ)



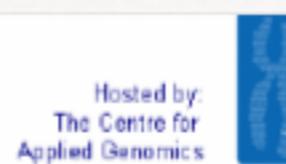
Lee et al. in Genomics and Personalized Medicine (In Press)

# BANCO DE DADOS DE CNVs

## Database of Genomic Variants

A curated catalogue of structural variation in the human genome

[About This Project](#) | [Genome Browser](#) | [Download](#) | [Links](#) | [Data Submissions](#) | [Email us](#)



Please select genome assembly:

### View Data by Chromosome

[1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) [21](#) [22](#) [X](#) [Y](#) [All](#)

### Keyword Search

Exact Match?  Yes  No

Examples: clone name, accession number, cytoband, gene

### BLAT Search

Enter sequence in FASTA format here:

### View Data by Genome



### Summary Statistics

Total entries: [27799 \(hg18\)](#)

CNVs: 17641

Inversions: 498

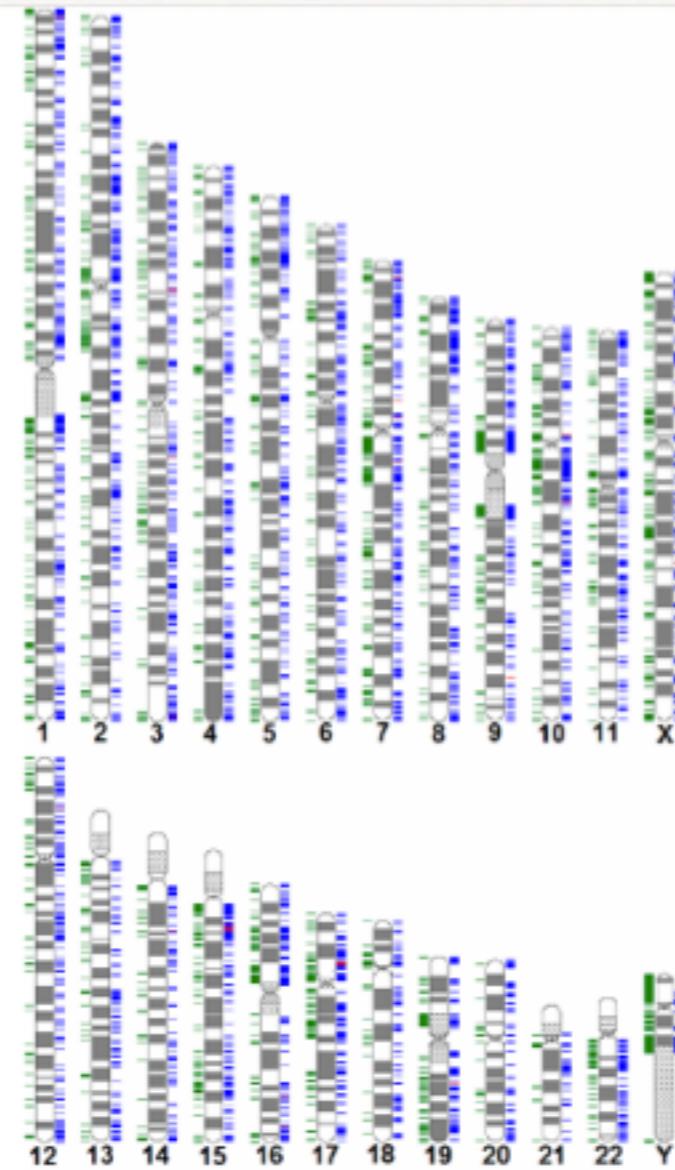
Indels (100bp-1Kb): 9660

Total CNV loci: 5672

Articles cited: 49

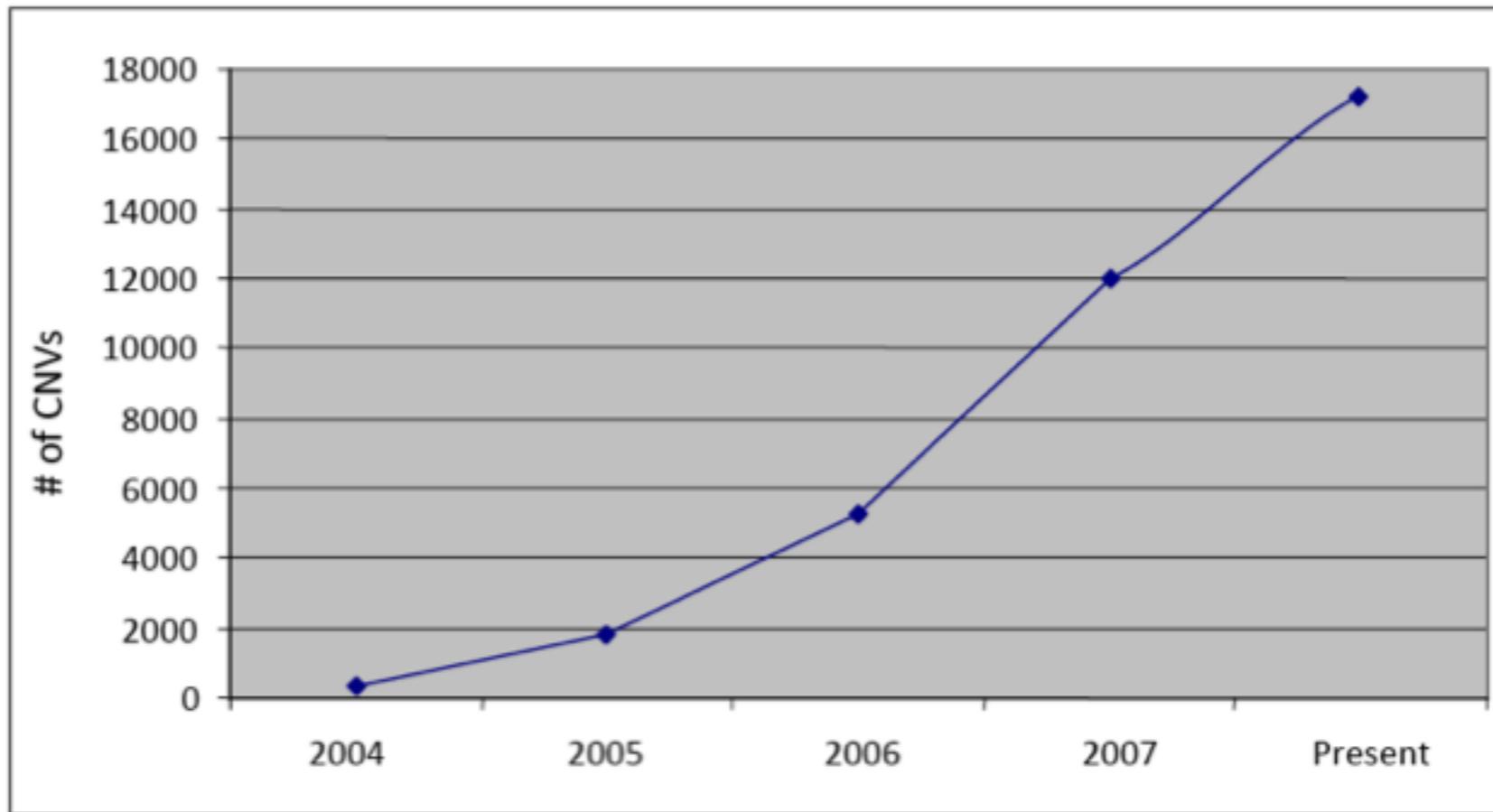
Last updated: Jun 26, 2008

Join our [mailing list](#)



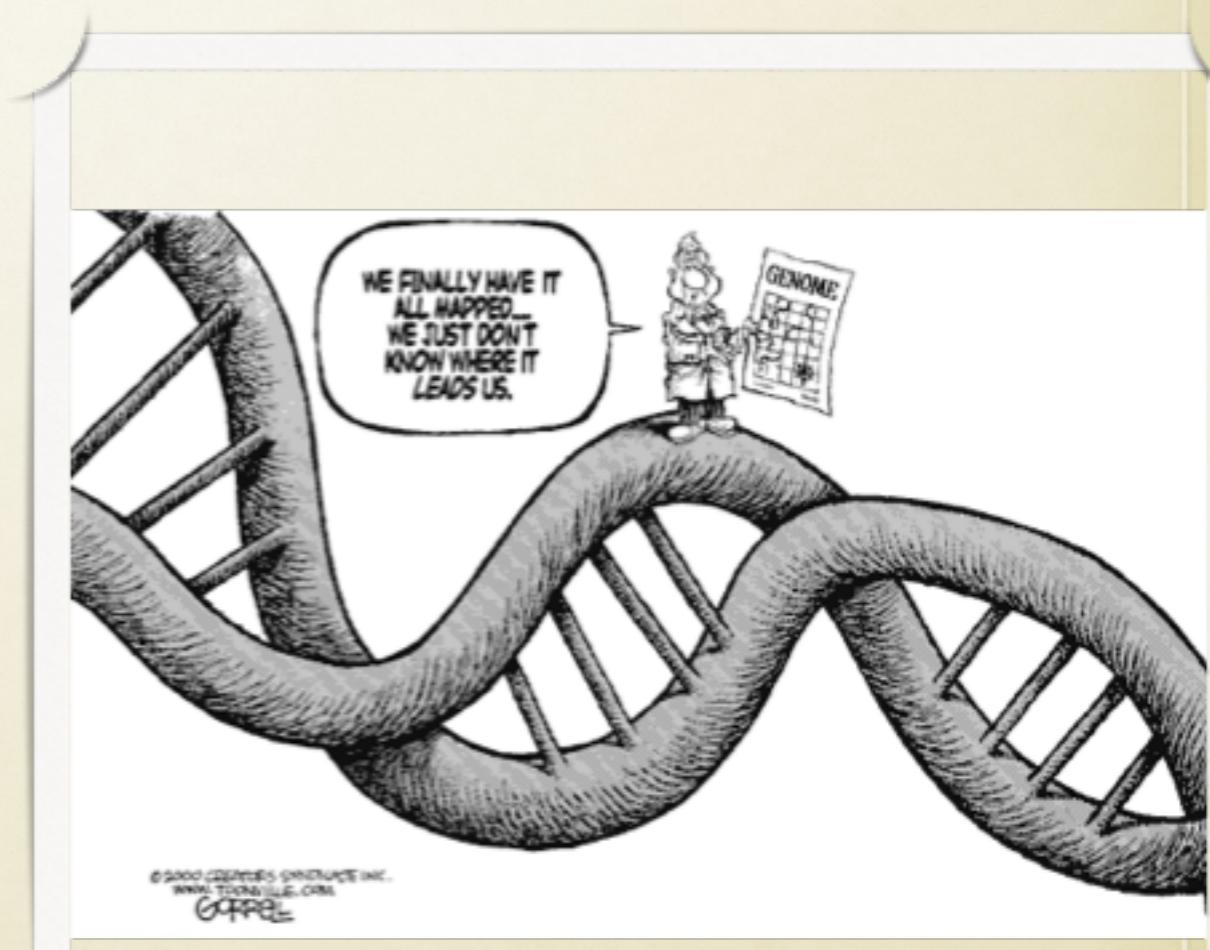
<http://projects.tcag.ca/variation/>

# AUMENTO DE CNVs NOS BANCOS DE DADOS



# LIMITAÇÃO DO ESTUDO DE CNVs

- Todos métodos usam comparação com um genoma “referência”, um fato que tem implicações no desenho e interpretação dos resultados
- Não foi adotado ainda um DNA como controle padronizado
  - pools de DNA muitas vezes são usados
- Interpretação: eis a questão



# METODOLOGIAS PARA DETECÇÃO DE CNVs

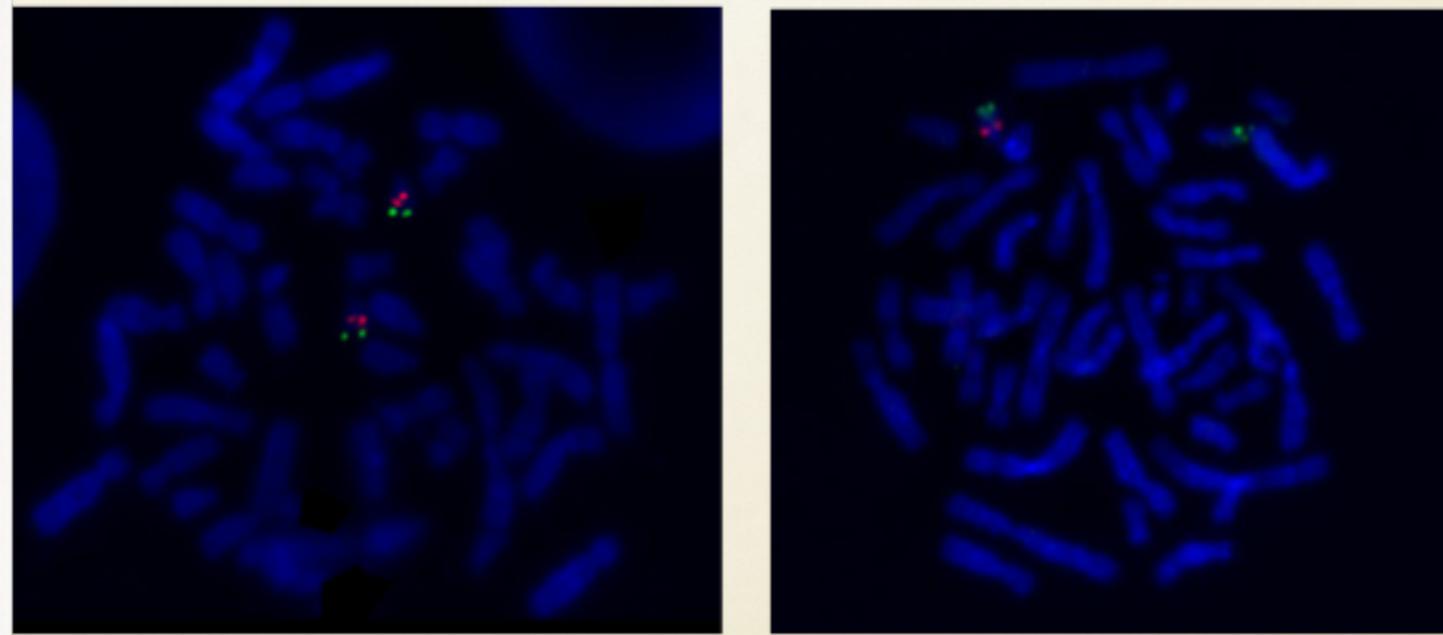
Table 1 | Methods for detecting structural variants in the human genome

Method	Translocation	Inversion	LCV (>50 kb)	CNV Indel (1–50 kb)	Small sequence variants (<1 kb)
<b>Genome-wide scans</b>					
Karyotyping	Yes (>3 Mb)	Yes (>3 Mb)	Yes (>3 Mb)	No	No
Clone-based array-CGH	No	No	Yes (>50 kb)	No	No
Oligonucleotide-based array-CGH	No	No	Yes (>35 kb)	Yes (>35 kb)	No
SNP array	No	No	Yes	Yes	Yes (SNPs)
Sequence-assembly comparison	Yes	Yes	Yes	Yes	Yes
Clone paired-end sequencing (fosmid)	Yes	Yes (breakpoints)	Yes (>8 kb of deletions)	Yes (>8 kb of deletions; <40 kb of insertions)	No
<b>Targeted scans</b>					
Microsatellite genotyping	No	No	Yes (deletions)	Yes (deletions)	Yes
MAPH	No	No	Yes	Yes	Yes
MLPA	No	No	Yes	Yes	Yes
QMPSF	No	No	Yes	Yes	Yes
Real-time qPCR	No	No	Yes	Yes	Yes
FISH	Yes	Yes	Yes	Yes	No
Southern blotting	Yes	Yes	Yes	Yes	Yes

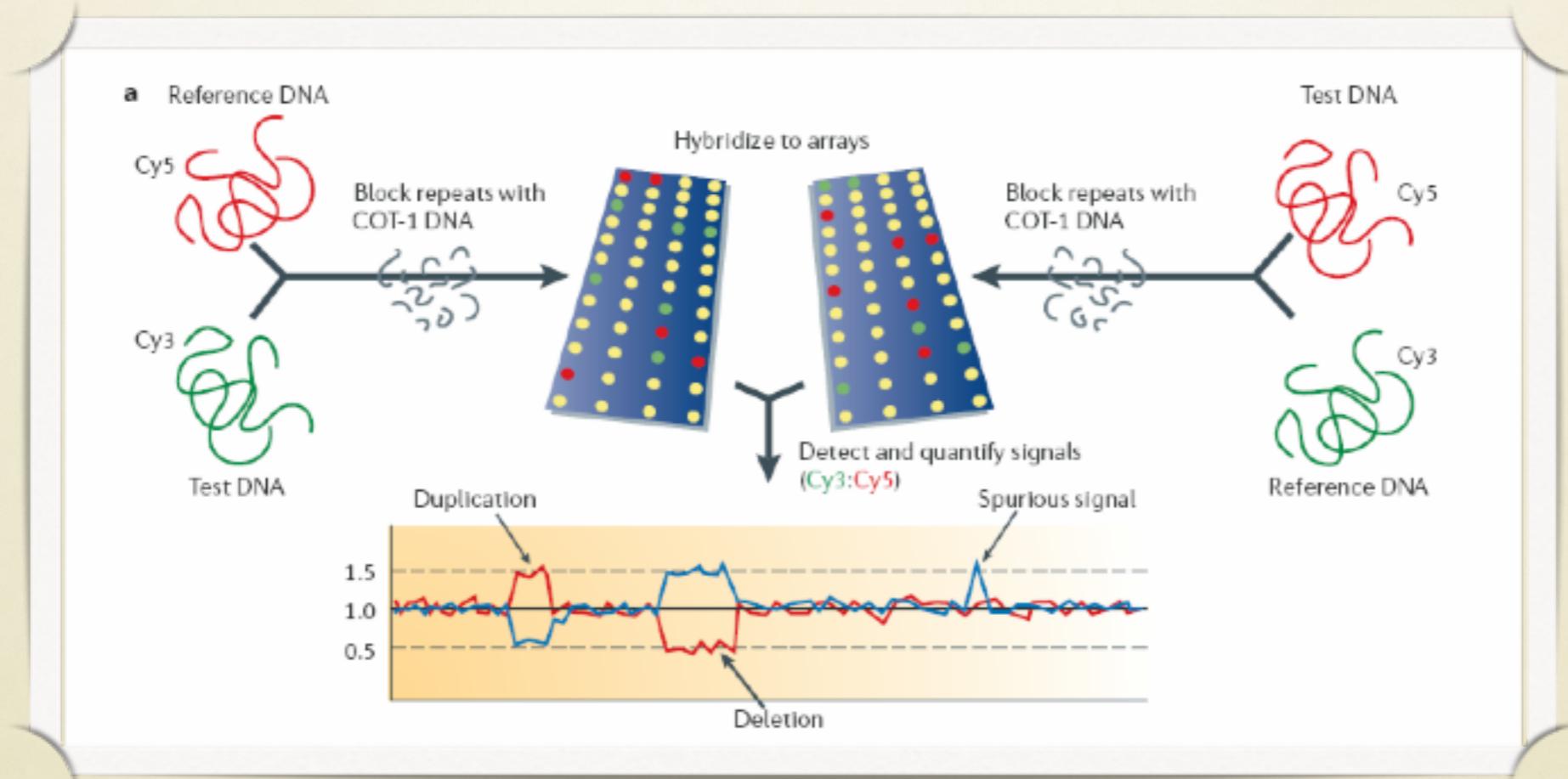
Detection limits are shown in parentheses where applicable. The emphasis is on those approaches that are used for the detection of submicroscopic variants, although karyotyping is also shown. For comparison, each technology's ability to detect smaller sequence variants (<1 kb) is also shown. CGH, comparative genome hybridization; CNV, copy-number variant; FISH, fluorescence *in situ* hybridization (including metaphase, Interphase and fibre FISH); Indel, insertion or deletion; LCV, large-scale CNV; MAPH, multiplex amplifiable probe hybridization; MLPA, multiplex ligation-dependent probe amplification; QMPSF, quantitative multiplex PCR of short fluorescent fragments; qPCR, quantitative PCR.

# FISH - HIBRIDAÇÃO IN SITU POR FLUORESCÊNCIA

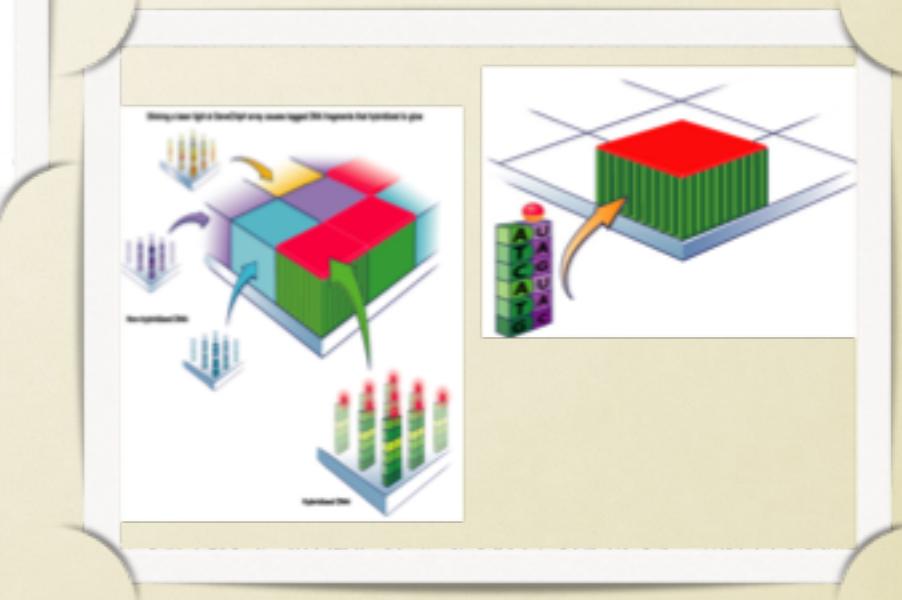
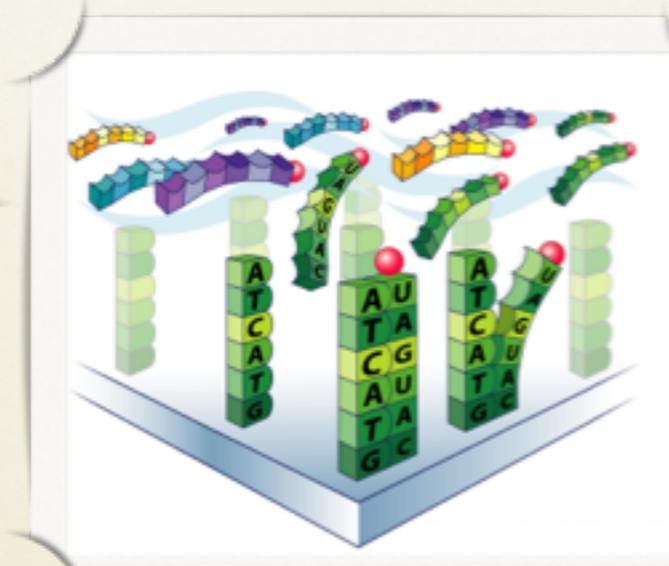
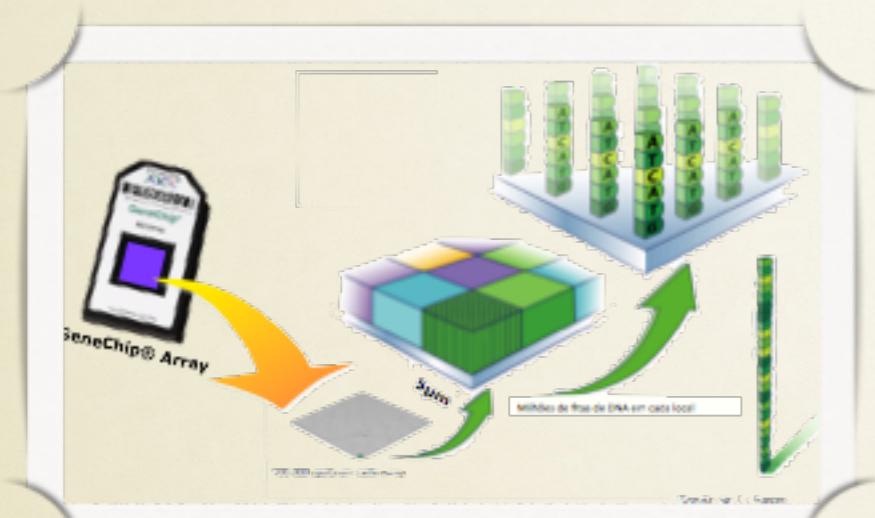
Del22q11.2



# CGH ARRAY

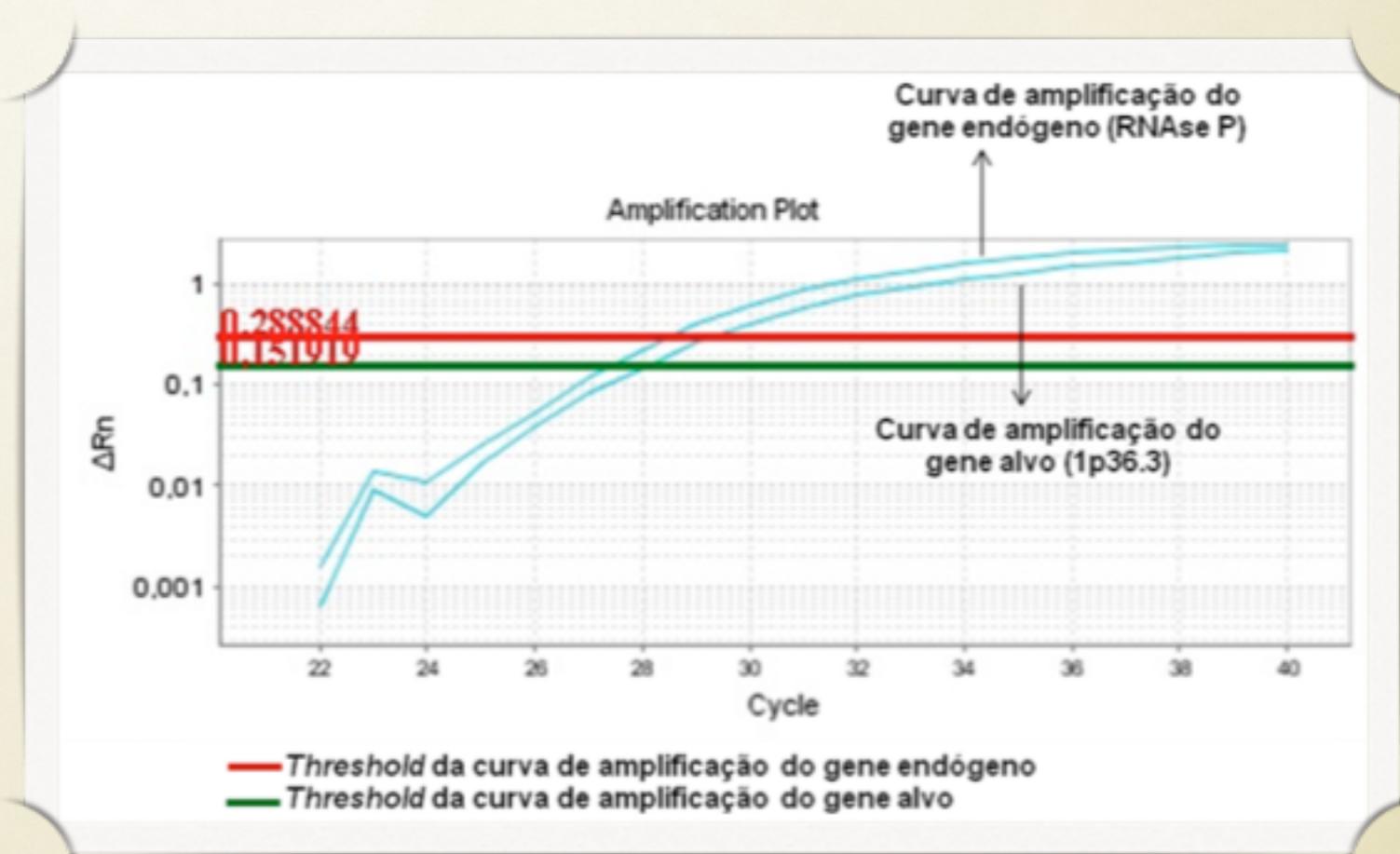


# MICROARRAY/SNP-ARRAY



# QPCR

- Validação de Microarray
- Detecção CNV



# **QPCR - $\Delta\Delta Ct$**

## **Método de Ct comparativo ( $\Delta\Delta Ct$ )**

- **$\Delta Ct$  de cada amostra:**
  - $Ct$  gene de interesse -  $Ct$  gene endógeno (cada poço)
  - Média aritmética das replicatas
- **$\Delta\Delta Ct$ :**
  - Média  $\Delta Ct$  da amostra alvo - Média  $\Delta Ct$  amostra calibradora
- $RQ = 2^{-\Delta\Delta Ct}$
- $RQ \times 2 =$  número de cópias

# QPCR - $\Delta\Delta Ct$

## Exemplo

**Reference assay**  
(VIC,  
RNase P)

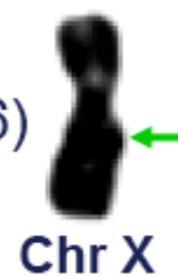


**Control**  
Sample (♀)

Ct (VIC)  
= 27.5

Ct (FAM)  
= 27.0

**Test assay**  
(FAM, YIPF6)



**Test**  
Sample (♂)

Ct (VIC)  
= 27.0

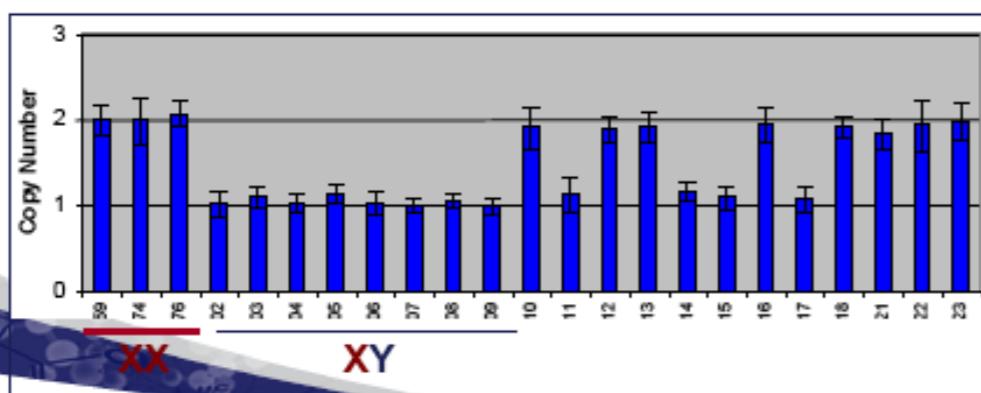
Ct (FAM)  
= 28.5

$\Delta Ct = 0.5$

$\Delta Ct = 1.5$

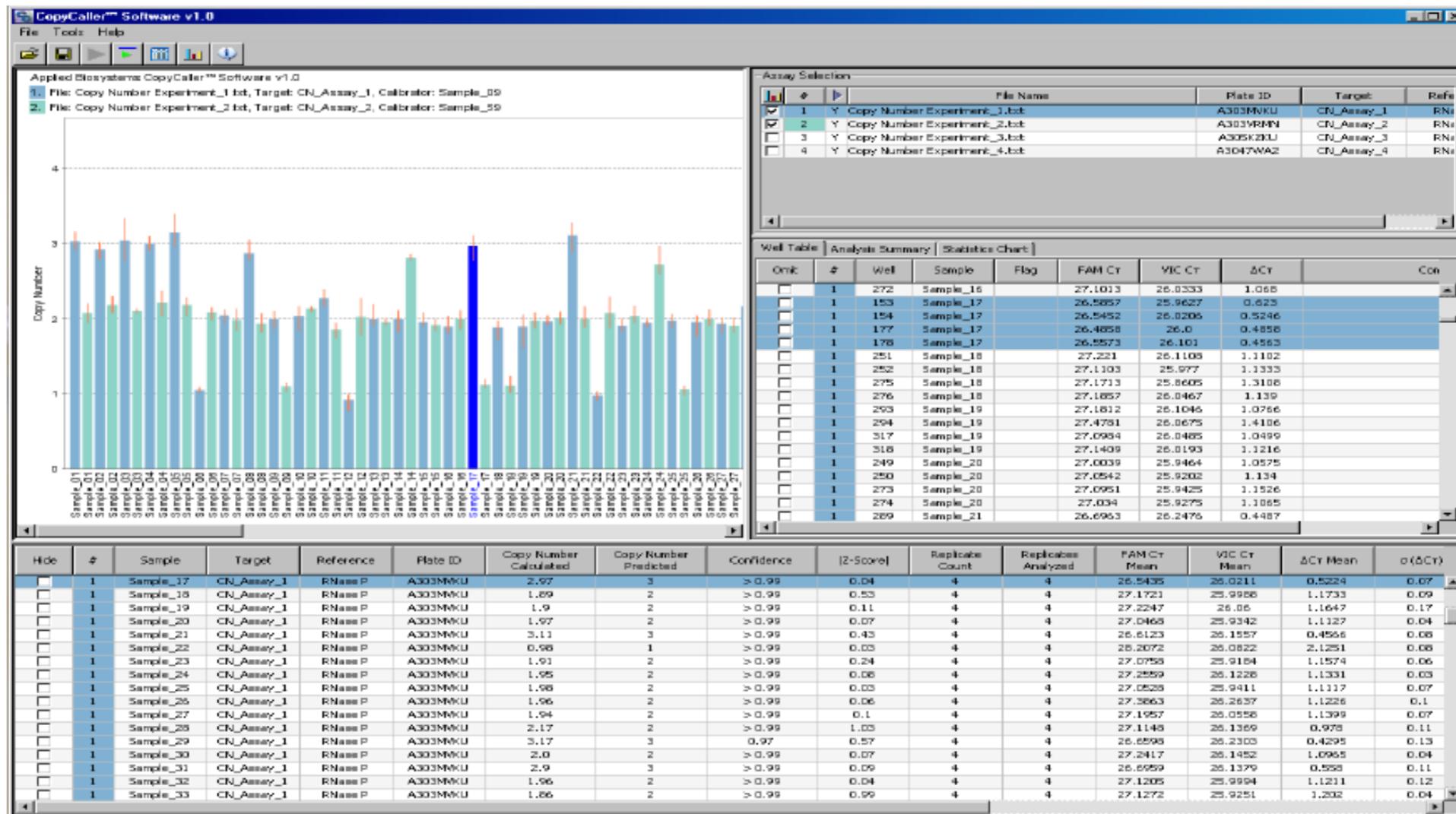
$\Delta\Delta Ct = 1$

Copy number  
 $2^{-\Delta\Delta Ct} \times 2 = 1$



# QPCR - $\Delta\Delta CT$

## Análise pelo Software CopyCaller™



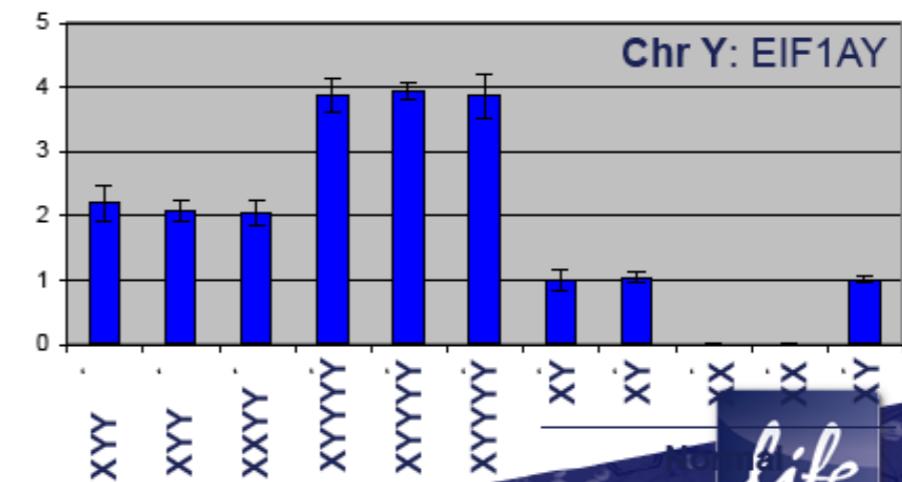
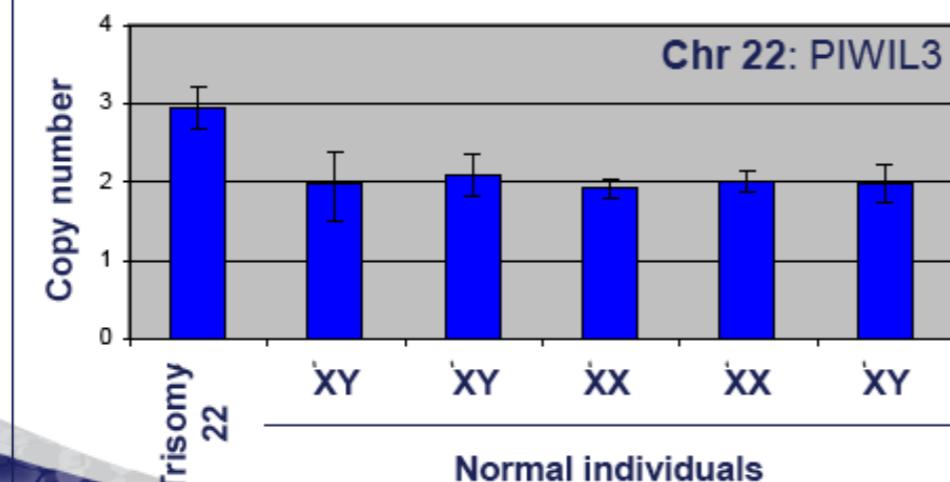
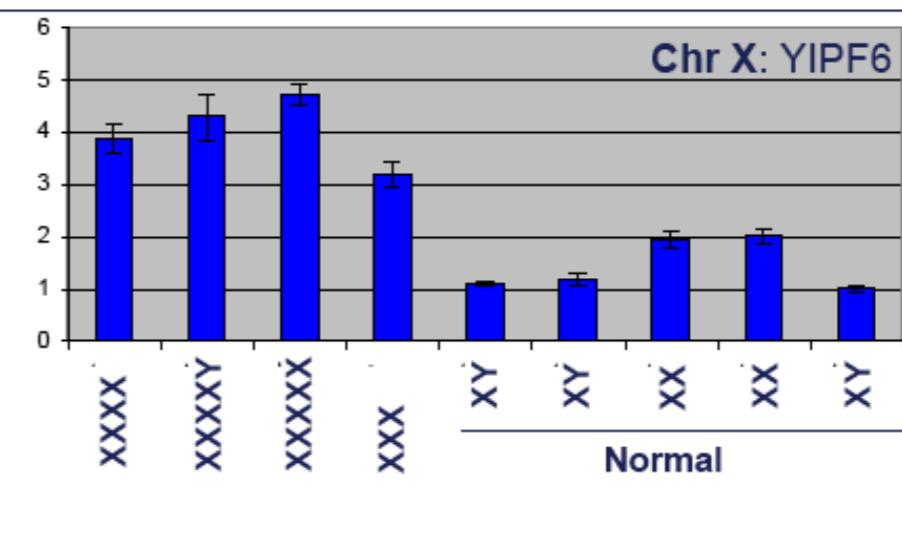
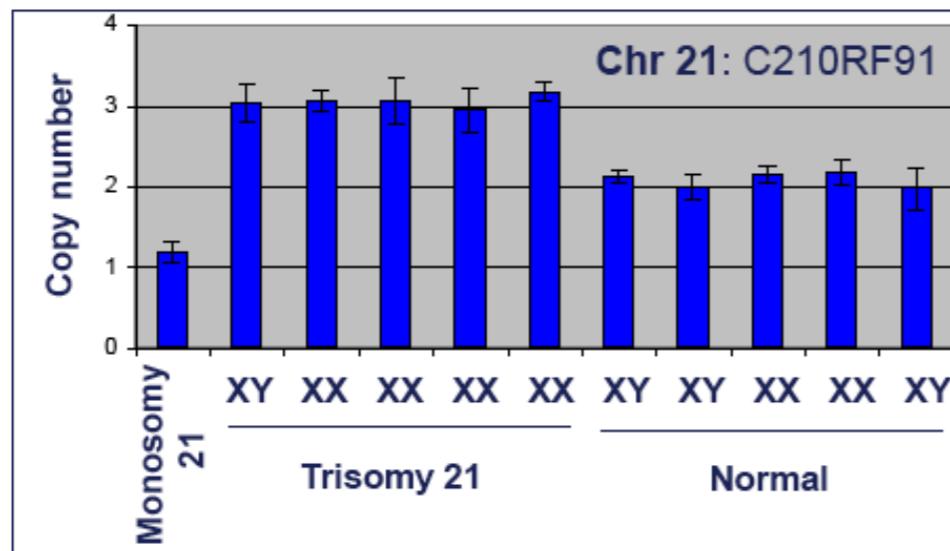
- Software gratuito para a análise dos dados de TaqMan Copy Number Assays
- Fácil utilização da interface gráfica, com resultados em poucos segundos
- Algoritmos para a identificação do nº de cópias de amostras utilizando valores de confiança



# QPCR - $\Delta\Delta CT$

## TaqMan Copy Number Assays

Acurácia para determinação de 0 a 5 cópias de gDNA



# QPCR

## COPY CALLER

Hide	#	Sample	Target	Reference	Plate ID	Copy Number Calculated	Copy Number Predicted	Confidence	Z-Score	Replicate Count	Replicates Analyzed	FAM Ct Mean
<input checked="" type="checkbox"/>	2		CNV3	RNAseP			undetermined			3	0	
<input type="checkbox"/>	2	BFS	CNV3	RNAseP		2.36	2	0.62	1.07	4	4	27.9944
<input type="checkbox"/>	2	CF	CNV3	RNAseP		1.67	2	> 0.99	0.57	4	4	27.9413
<input type="checkbox"/>	2	CM	CNV3	RNAseP		2.0	2	0.96	0.02	4	4	27.7674
<input type="checkbox"/>	2	ECF	CNV3	RNAseP		2.55	3	< 0.50	0.6	4	4	27.9813
<input type="checkbox"/>	2	GBL	CNV3	RNAseP		2.06	2	0.96	0.05	4	4	27.7597
<input type="checkbox"/>	2	JJB	CNV3	RNAseP		2.19	2	0.86	0.41	4	4	28.5934
<input type="checkbox"/>	2	LE	CNV3	RNAseP		1.88	2	0.99	0.03	4	4	29.6313
<input type="checkbox"/>	2	MJS	CNV3	RNAseP		2.9	3	0.73	0.0	4	4	29.6882
<input type="checkbox"/>	2	SCB	CNV3	RNAseP		3.79	4	0.53	0.03	4	4	28.1919
<input type="checkbox"/>	2	SJS	CNV3	RNAseP		1.62	2	> 0.99	0.95	4	4	28.1747

**Tabela 1 – Valores da quantificação relativa (RQ) de cada amostra em relação à amostra referência.**

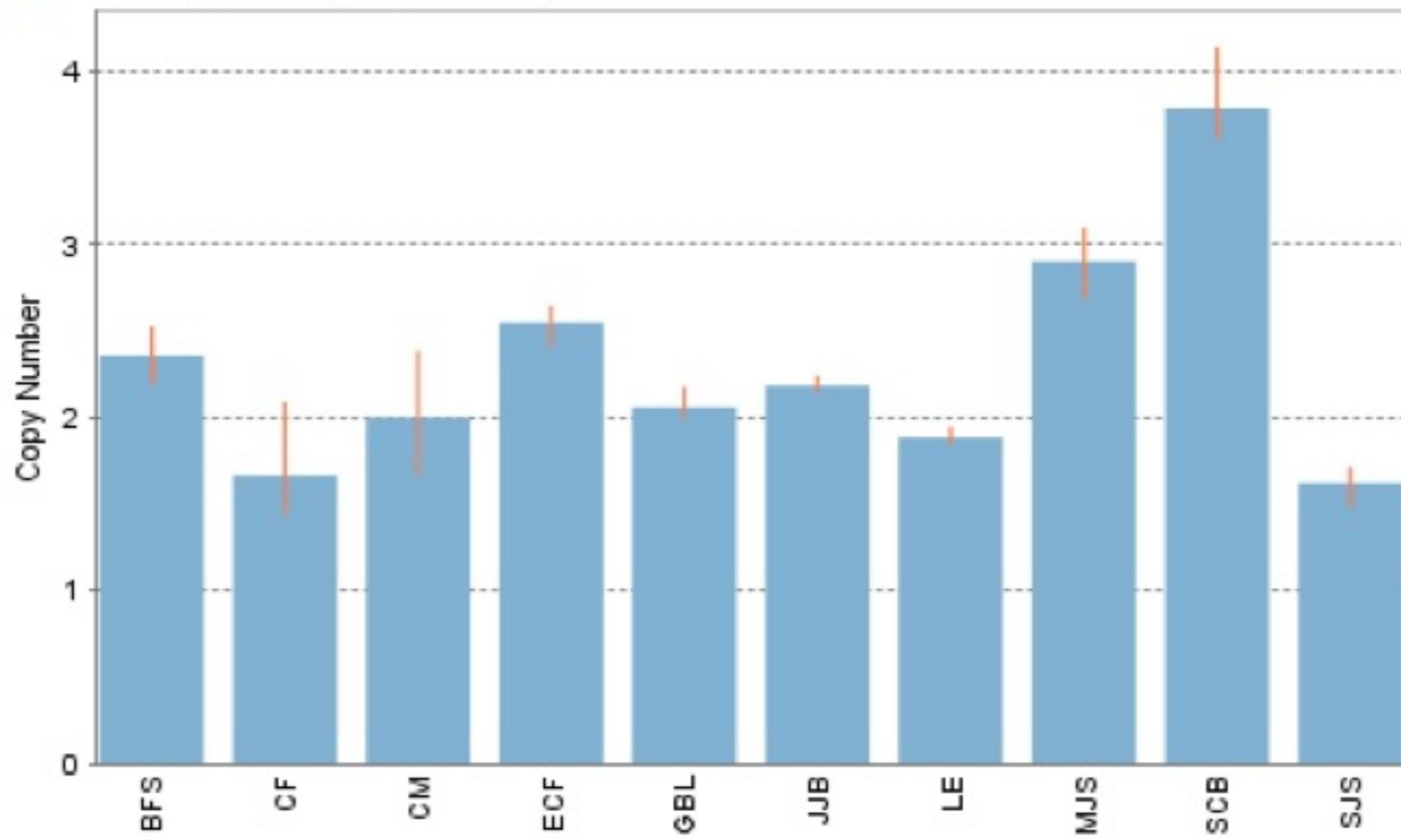
Amostra	RQ	RQ Min	RQ Max	RQ arredondado
GBL	0,527875	0,427823	0,651325107	0,5
ECF	0,821889	0,599272	1,127204418	1
SJS	0,28728	0,246363	0,334992647	0,5
LE	1,831166	1,637094	2,048244476	2
BFS	1,746405	1,687752	1,807095528	2
JCB	3,960221	3,800852	4,126272678	4
MJS	2,877312	2,610138	3,171834946	3
JJB	1,04678	0,898521	1,219502211	1
DNA calibrador (CM)	1	0,812303	1,231068254	

# QPCR

# COPY CALLER

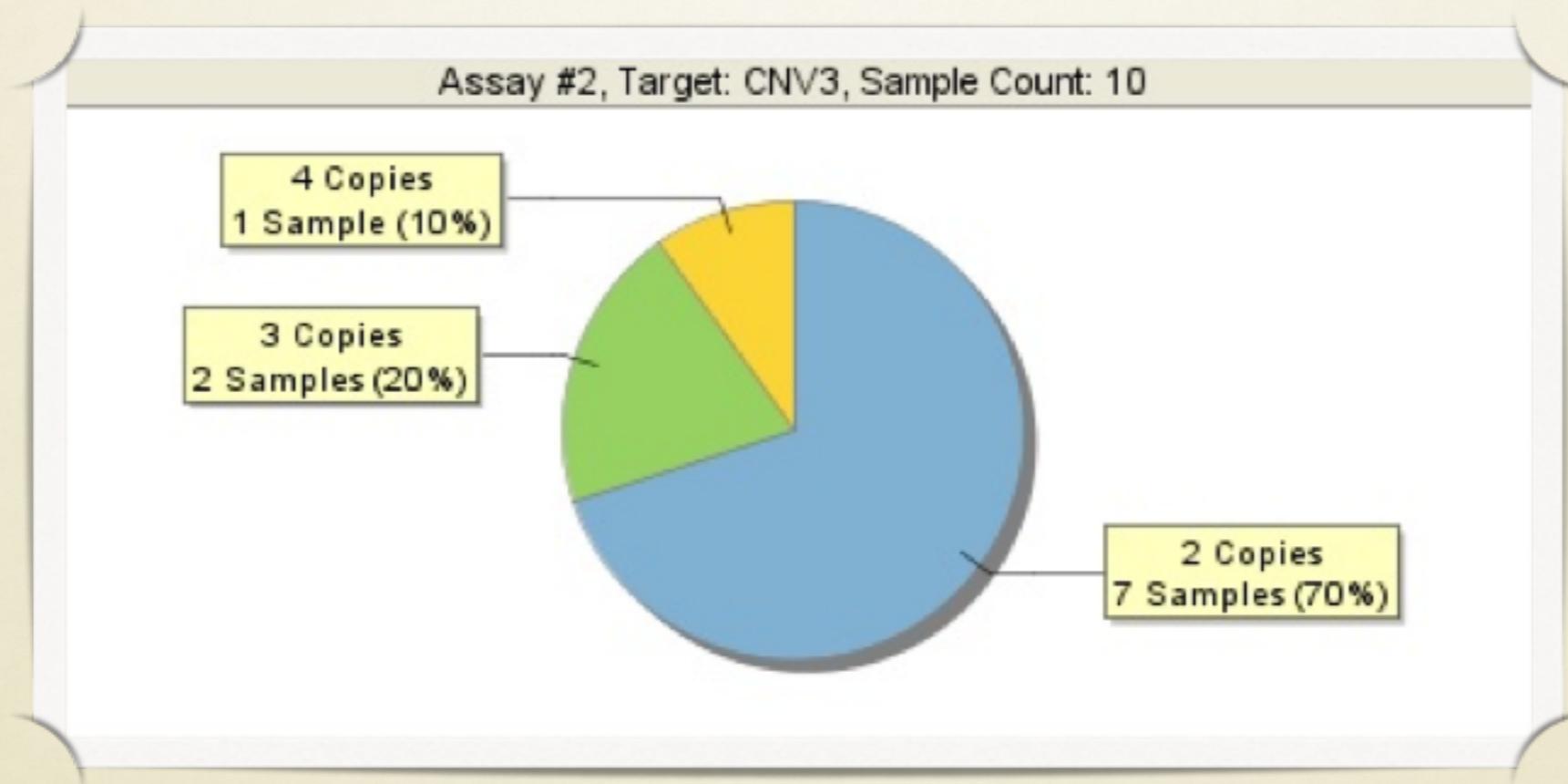
Applied Biosystems CopyCaller™ Software v1.0

2. File: CNV3\_MOD.txt, Target: CNV3, Calibrator: CM



# QPCR

## COPY CALLER

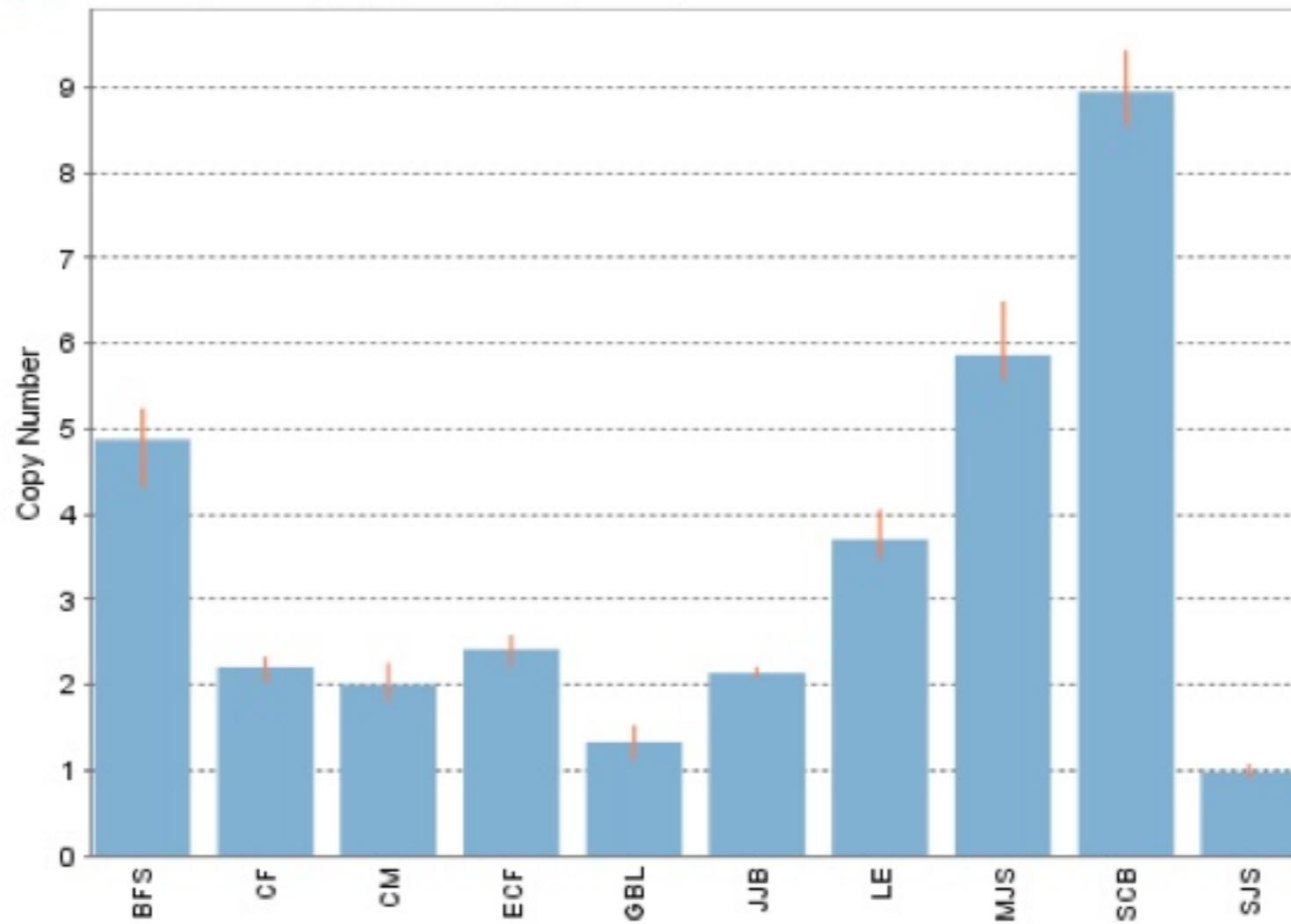


# QPCR

# COPY CALLER

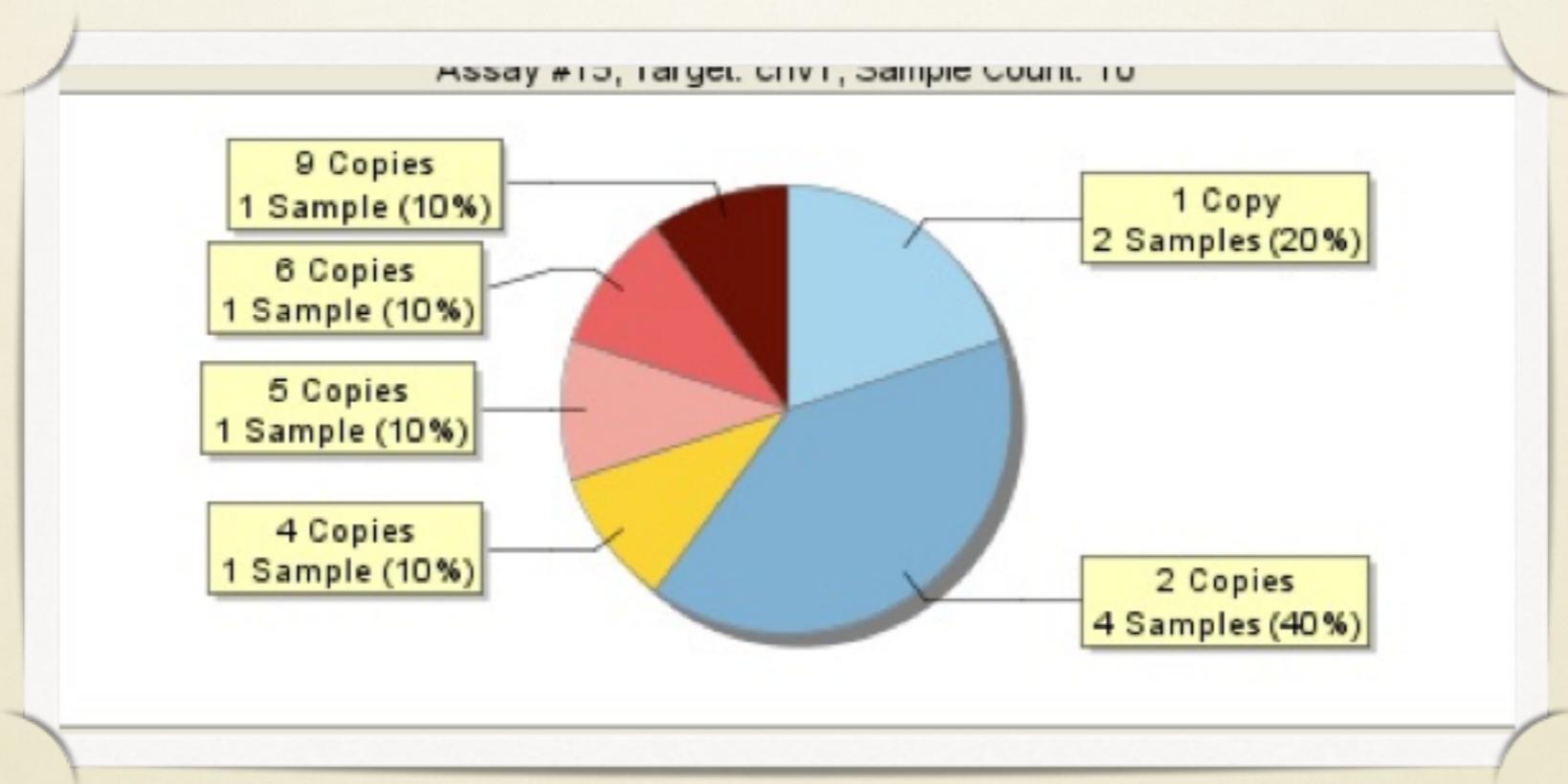
Applied Biosystems CopyCaller™ Software v1.0

15. File: cnv1\_marcos\_22\_3\_data.txt, Target: cnv1, Calibrator: CM



# QPCR

## COPY CALLER



**FIM**

**- SINTIA IOLE BELANGERO**