

# GENOMA HUMANO

## O QUE É GENOMA

- Ø Conjunto de todo o material genético que estabelece a identidade de uma espécie.
- Ø **Genoma = DNA nuclear + DNA citoplasmático:**
  - Genoma nuclear
  - Genoma mitocondrial
  - Genoma cloroplástico (vegetais)
- Ø Pode ser formado por:
  - DNA: a maior parte dos organismos
  - RNA: vírus



## GENOMAS PROCARIÓTICOS X EUCARIÓTICOS

PROCARIOTES	EUCARIONTES
Genoma compacto e econômico.	Grandes blocos de DNA que nada codificam.
Seqüências únicas.	Múltiplas cópias de seqüências particulares.
Genes não-interrompidos.	Genes com íntrons.

## PROJETO GENOMA HUMANO (PGH)

- Ø Projeto coordenado pelo Departamento de Energia e Institutos Nacionais de Saúde dos E.U.A.
- Ø *International Human Genome Sequencing Consortium:*
  - EUA, Reino Unido, Japão, França, Alemanha e China (principalmente).
- Ø 1986 – início do projeto piloto.
- Ø 1990 – início oficial.



## OBJETIVOS DO PGH

### Ø INICIAIS:

- construir detalhadamente os mapas genético e físico do genoma humano;
- determinar a seqüência completa dos ~ 3 bilhões de pares de nucleotídeos do DNA humano;



- localizar os cerca de 25.000 genes dentro do genoma humano;

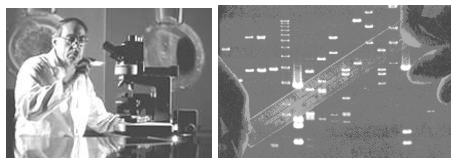


- executar análises semelhantes nos genomas de diversos organismos usados em laboratórios de pesquisas, como sistemas-modelo.
- estocar estas informações em banco de dados;
- melhorar a tecnologia para pesquisa biomédica (para o seqüenciamento, bioinformática e computação – ferramentas para análises dos dados);
- responder questões éticas, legais e sociais que surjam em decorrência das pesquisas do genoma humano.

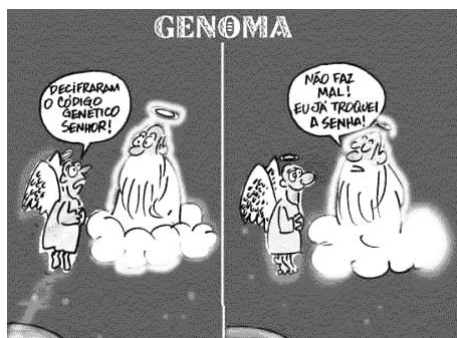


## Ø AINDA EM DESENVOLVIMENTO:

- desenvolver tecnologias para o seqüenciamento;



- identificar os genes no genoma humano;



- interpretar funções gênicas (genoma funcional);
- identificar variações de seqüências (Projeto de Diversidade Humana e SNPs);
- responder questões éticas, legais e sociais que surjam em decorrência das pesquisas do genoma humano.

## QUEM ESTÁ CONDUZINDO O PGH HOJE?



J. Craig Venter

- Ø The Human Genome Project, o consórcio inicial, coordenado por Francis Collins.
- Ø Celera Genomics, coordenado por Craig Venter.



Francis Collins

## DE QUEM É O GENOMA QUE ESTÁ SENDO SEQÜENCIADO?

- Ø O genoma-referência é uma composição de genomas de diferentes povos:
  - amostras de doadores anônimos de diferentes grupos étnicos.



## QUANTOS GENOMAS COMPLETOS EXISTEM NO MOMENTO?

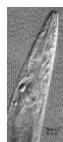
- Ø O número altera a cada dia:
  - Vírus: > 1.000
  - Bactéria: > 400
  - Archaea: > 35
  - Eukaryotes: > 85
- Ø Primeiro genoma completo Æ DNA mitocondrial humano (1981) com 16.159 pb.



## ALGUNS GENOMAS FINALIZADOS



*S. cerevisiae*



*C. elegans*



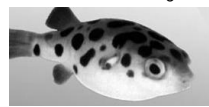
*D. melanogaster*



*A. thaliana*



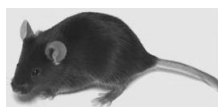
*E. coli*



*Tetraodon nigroviridis*



*Xenopus laevis*



*Mus musculus*



*Canis familiaris*



*Homo sapiens*

Catálogo de genomas já finalizados:  
[http://www.genome.jp/kegg/catalog/org\\_list.html](http://www.genome.jp/kegg/catalog/org_list.html)

## TAMANHO DE GENOMAS

Ø Valor C = tamanho do genoma haplóide

- SV40  $\approx 5 \times 10^3$  pb
- *Mycoplasma genitalium*  $\approx 6 \times 10^5$  pb (menor procarioto)
- *Escherichia coli* (bactéria)  $\approx 5 \times 10^6$  pb
- *Sacharomyces cerevisiae* (levedura)  $\approx 2 \times 10^7$  pb
- *Arabidopsis thaliana* (vegetal)  $\approx 10^8$  pb
- Carvalho  $\approx 5 \times 10^8$  pb
- Humanos  $\approx 3 \times 10^9$  pb
- Soja  $\approx 3 \times 10^9$  pb
- *Xenopus laevis* (rã)  $\approx 3 \times 10^9$  pb
- Salamandra  $\approx 8 \times 10^{10}$  pb
- Lúrio  $\approx 10^{11}$  pb

## PRIMEIRO RASCUNHO

Ø Jun 2000/publicado em fev.2001:

- PGH: Nature, vol. 409 (2001);
- Celera: Science, vol. 291 (2001).

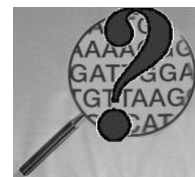


## SEQÜÊNCIA COMPLETA

Ø Abril de 2003  $\approx$  durante as comemorações de 50 anos da molécula de DNA.

Ø Onde encontrar a seqüência:

- University of California:  
<http://genome.ucsc.edu>
- The National Center for Biotechnology Information: <http://www.ncbi.nih.gov>



## GENOMA HUMANO EM NÚMEROS

Ø 3.000.000.000 pb

Ø Analogia: a enciclopédia genômica:

- média 5 letras por palavra  $\approx$  600.000.000 palavras;
- média 12 palavras por linha  $\approx$  50.000.000 linhas;
- média 70 linhas por página  $\approx$  700.000 páginas.



"Enciclopédia Genômica"

~700 volumes  
(1.000 páginas cada)

## GENOMA HUMANO EM NÚMEROS

Ø Estima-se que 99.9% da seqüência seja exatamente a mesma entre todos os seres humanos.

Ø Tamanho dos genes  $\approx$  varia enormemente:

- gene médio: 3.000 bases;
- maior gene conhecido: distrofina (2.4 milhões de bases).



Ø Número total de genes: estimado em ~25.000  $\approx$  muito menor que as estimativas prévias de 80.000 a 120.000.

Ø Mais de 50% dos genes encontrados ainda não têm função definida.

## COMO O GENOMA HUMANO ESTÁ ORGANIZADO

Ø Os genes estão distribuídos ao longo do genoma, linearmente.



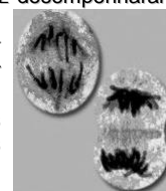
Ø Cromossomo 1 tem a maioria dos genes (2.968), e o Y a menor parte (231).

Ø Menos de 2% do genoma codifica proteínas  $\approx$  maioria do genoma é não codificador.

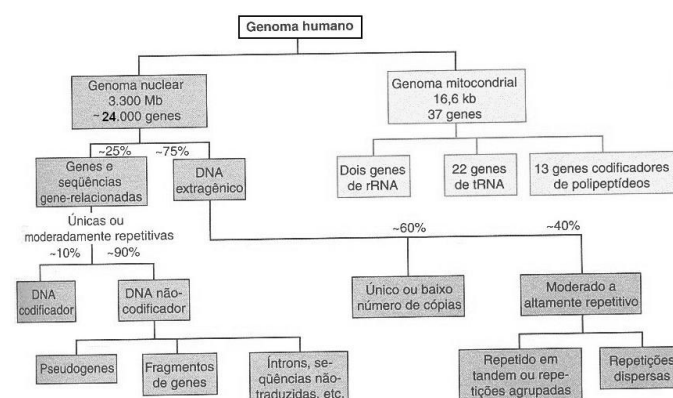
## COMO O GENOMA HUMANO ESTÁ ORGANIZADO

Ø Seqüências repetitivas não codificadoras ("junk DNA")  $\approx$  compõem pelo menos 50% do genoma:

- aparentemente não tem função direta  $\approx$  podem estar envolvidas com a estrutura e a dinâmica dos cromossomos;
- podem apresentar alguns trechos correspondentes a genes que ao longo da evolução deixaram de ter função  $\approx$  desempenharam papel importante na evolução;
- ao longo do tempo, remodelam o genoma por rearranjo, podendo criar genes novos, modificar e reembaralhar genes existentes;
- formam os centrômeros  $\approx$  estruturas fundamentais na correta distribuição dos cromossomos durante a divisão celular.
- formam os telômeros  $\approx$  estruturas que marcam o início e o fim dos cromossomos e estão relacionadas com o envelhecimento e a manutenção da integridade estrutural do cromossomo, garantindo a replicação completa das extremidades codificadoras.



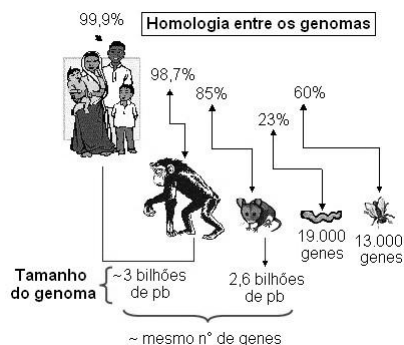
## COMPOSIÇÃO DO GENOMA HUMANO





## COMPARAÇÃO ENTRE GENOMAS

ORGANISMOS	TAMANHO (pb)	GENES ESTIMADOS
Humano ( <i>Homo sapiens</i> )	3 bilhões	~25.000
Camundongo ( <i>M. musculus</i> )	2,6 bilhões	~25.000
Chimpanzé ( <i>Pan troglodytes</i> )	3 bilhões	~25.000
Mostarda ( <i>A. thaliana</i> )	100 milhões	25.000
Verme ( <i>C. elegans</i> )	97 milhões	19.000
Mosca da fruta ( <i>D. melanogaster</i> )	137 milhões	13.000
Levedura ( <i>S. cerevisiae</i> )	12,1 milhões	6.000
<i>E. coli</i>	4,6 milhões	3.200
HIV	9,700	9



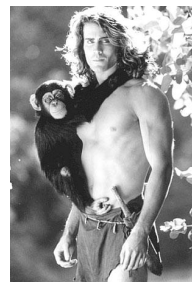
Ø O genoma do camundongo é cerca de 10% menor que o genoma humano:

- ~ mesmo número de genes;
- menor conteúdo de DNA repetitivo do que genoma humano.

## COMPARAÇÃO ENTRE GENOMAS

Ø Diferenças entre genomas de *Homo sapiens* (humanos) e *Pan troglodytes* (chimpanzé):

- mutações de um só nucleotídeo Æ 1, 23%
- duplicações Æ 2,7%
- 4 humanos Æ 33% correspondem a seqüências específicas da espécie (levando geralmente a aumento de expressão);
- 4 chimpanzés Æ cerca de 17% são de seqüências específicas da espécie.
- inserções/deleções de segmentos cromossômicos Æ 3%
- 4 há ~7.000 elementos Alu no genoma humano e ~2.300 no genoma do chimpanzé.



Ø O genoma humano tem uma porção bem maior de seqüências repetitivas (50%) do que a mostarda (11%), *C. elegans* (7%) e mosca doméstica (3%).

Ø O genoma humano tem áreas ricas em genes enquanto outros genomas são mais uniformes.

Ø Humanos tem em média três vezes mais produtos protéicos devido a *splicing* alternativo e modificações pós-traducionais (~25.000 genes e 400.000 proteínas) Æ **um gene = várias proteínas.**

Ø Humanos compartilham a maioria das mesmas famílias protéicas com nematelmintos, moscas e plantas, mas o número de genes destas famílias é maior em humanos, especialmente aqueles envolvidos com desenvolvimento e imunidade.

## O QUE AINDA NÃO SE SABE

- Ø Número, localização e função exata de todos os genes.
- Ø Regulação gênica.
- Ø Tipo, quantidade, distribuição, conteúdo informativo e função do DNA não-codificante.
- Ø Coordenação entre expressão gênica, síntese protéica e eventos pós-traducionais.
- Ø Interação das proteínas na maquinaria molecular.
- Ø Comparação entre a função gênica predita X a experimentalmente determinada.
- Ø Conservação evolutiva entre organismos.
- Ø Conservação protéica (estrutura e função).
- Ø Proteomas (conteúdo total das proteínas) dos organismos, bem como suas funções.
- Ø Correlação entre SNPs (variação de base simples entre indivíduos), saúde e doença.
- Ø Predição de suscetibilidade a doenças baseada na variação de seqüência gênica.
- Ø Envolvimento genético em características complexas e doenças multigênicas.
- Ø Genética do desenvolvimento.
- Ø Entre outras...



## PREVENDO O FUTURO

- Ø O que irá acontecer agora que os seqüenciamentos do genoma humano e de outros organismos estão completos?
- Ø Quanto tempo irá demorar os próximos passos?
- Ø Quais as dificuldades potenciais?



## O PLANO

- Encontrar todos os genes
- È
- Definir quais são os produtos gênicos
- È
- "Computar" funções por similaridade com proteínas conhecidas
- È
- "Computar" estrutura protéica
- È
- Verificar experimentalmente estrutura e função

## BENEFÍCIOS DO PROJETO GENOMA HUMANO

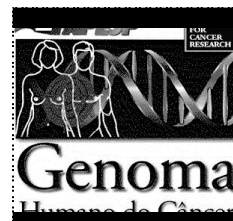


Ø Melhoria de exames laboratoriais e diagnósticos precoces de predisposição a doenças Æ medicina preventiva (modelada individualmente, de acordo com as suscetibilidades de cada um).



Ø Criação de drogas específicas para cada tipo de doença, com redução de efeitos colaterais (farmacogenômica).

Ø Diagnóstico de tumores em fase inicial Æ aumento da probabilidade de cura.



Ø Terapia Gênica (ainda em estudo laboratorial).

## BENEFÍCIOS DO PROJETO GENOMA HUMANO

Ø Uso de animais transgênicos e clonagem de embriões para transplante de órgãos.

Ø Entendimento dos mecanismos de envelhecimento, obesidade etc.

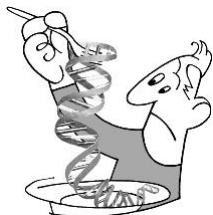
Ø Testes genéticos melhorados baseados em avanços tecnológicos e detecção mais precisa de assassinos e outros criminosos.

Ø Teste de paternidade e maternidade.

Ø Desenvolvimento de novos medicamentos através de organismos transgênicos.

Ø Melhoramento genético de plantas e animais para uso humano.

Ø Estudos de evolução e migração humana (uso de marcadores genéticos).



## IMPLICAÇÕES ÉTICAS, LEGAIS E SOCIAIS DO PROJETO GENOMA HUMANO

### IMPARCIALIDADE E PRIVACIDADE

Ø Quem pode ter acesso à informação genética de um indivíduo?

Ø Como esta informação pode ser usada?

Ø Quem pode deter e controlar essa informação?

Ø Considerando essas possibilidades:

- seguradoras, empregadores, cortes, escolas, agências de adoção de indivíduos, militares, governo?
- outras possibilidades?



### IMPACTO PSICOLÓGICO E ESTIGMATIZAÇÃO

Ø Como um indivíduo pode ser afetado pelo conhecimento de que tem uma predisposição a uma doença qualquer?

Ø Quanto esta informação afeta a percepção da família e amigos com relação a este indivíduo?

Ø Quanto esta informação afeta a percepção da sociedade com relação a este indivíduo?



## AVALIAÇÕES GENÉTICAS

Ø Avaliações genéticas para uma determinada doença ou predisposição genética só devem ser realizados com objetivo de decisão informada (pré-natal) quando há história na família?

Ø "Screening" populacionais devem ou não ser realizados (pré-matrimonial, pré-natal, recém-nascidos, ocupacional)?

Ø Avaliações genéticas devem ser realizadas quando não há tratamento disponível?

Ø Avaliações Genéticas devem ser feitas para genes de suscetibilidade?

Ø Os pais tem direito de avaliarem geneticamente seus filhos menores para doenças da idade adulta?

Ø As avaliações genéticas são viáveis e possíveis de interpretação pela comunidade médica?

Ø Quais profissionais estão qualificados para realizar o aconselhamento genético?



## QUESTÕES REPRODUTIVAS

Ø Consentimento informado para procedimentos.  
Ø Utilização de informação genética em tomadas de decisões.  
Ø Direito de reprodução.



## IMPLICAÇÕES FILOSÓFICAS E CONCEITUAIS



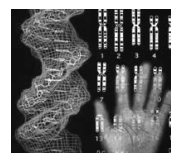
- Ø Responsabilidade humana.
- Ø Predisposição:
  - determinismo genético;
  - conceitos de saúde e doença.

### DETERMINISMO GENÉTICO



- Ø As pessoas teriam um programa embutido no "hardware" que não poderia ser alterado?
- Ø Até que ponto os genes determinam nosso destino?

### REDUCIONISMO



- Ø A sequência do genoma fornece informação suficiente que explica toda causalidade?

### COMERCIALIZAÇÃO



- Ø Direitos de propriedade.
- Ø Patentes, "copyrights" e "trade secrets".
- Ø Acessibilidade aos dados e materiais gerados.

## QUESTÕES CLÍNICAS

Ø Educação dos profissionais de saúde, pacientes e público em geral.

Ø Implementação de padrões e medidas de controle de qualidade nos procedimentos de avaliação genética.



## UM POUCO DE HUMOR

