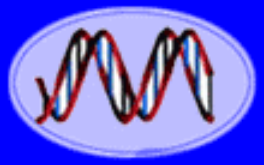
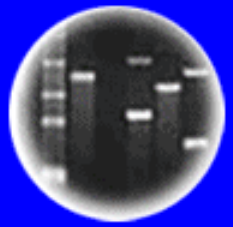


DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues
and Future Consequences

Plants of Tomorrow

# HC70A & SAS70A Winter 2011
# Genetic Engineering in Medicine, Agriculture, and Law

Professors Bob Goldberg, Channapatna Prakash & John Harada

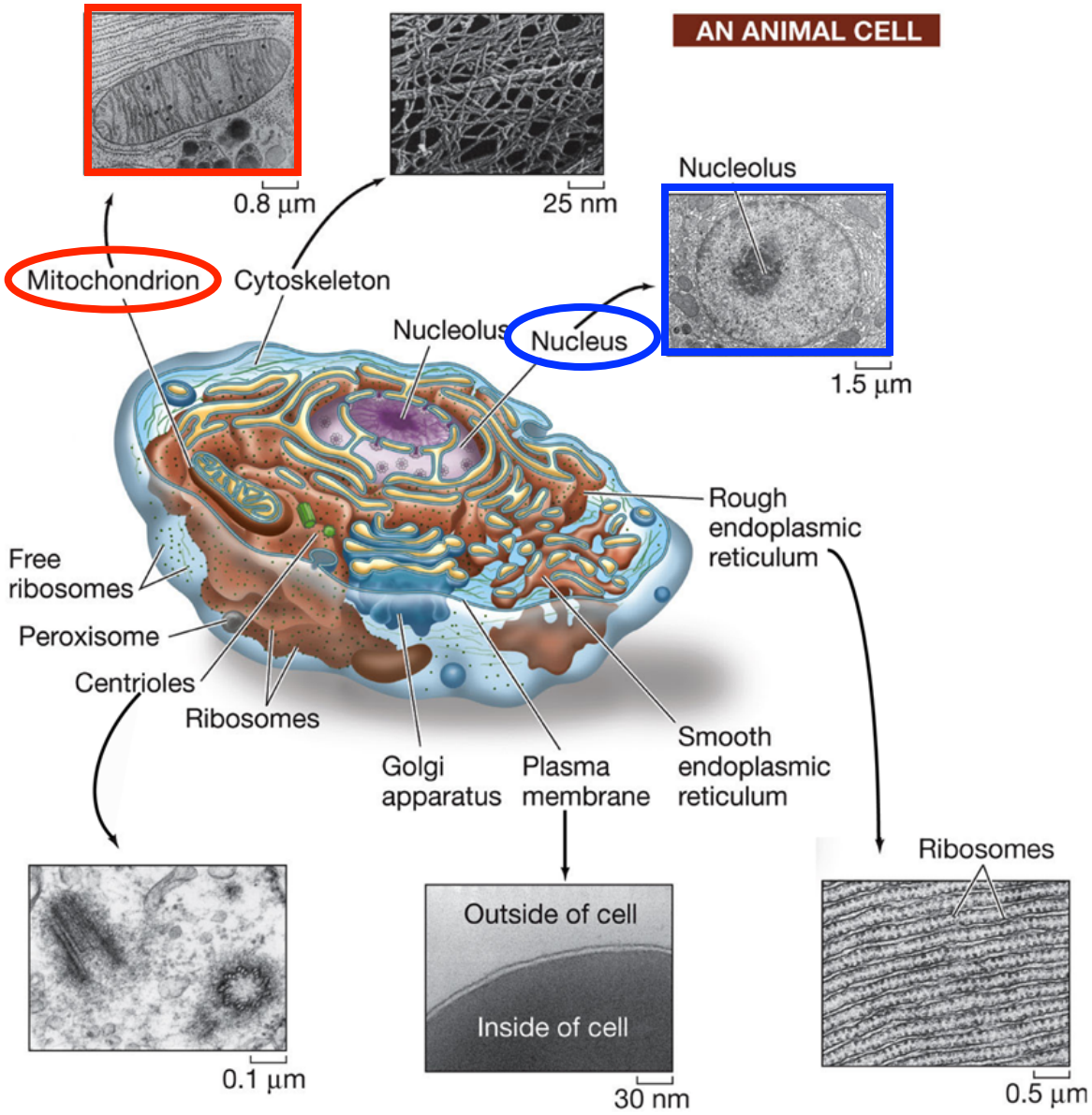
## Lecture 7
## Your Personal Genome & Tracing Your Ancestry

UCLA

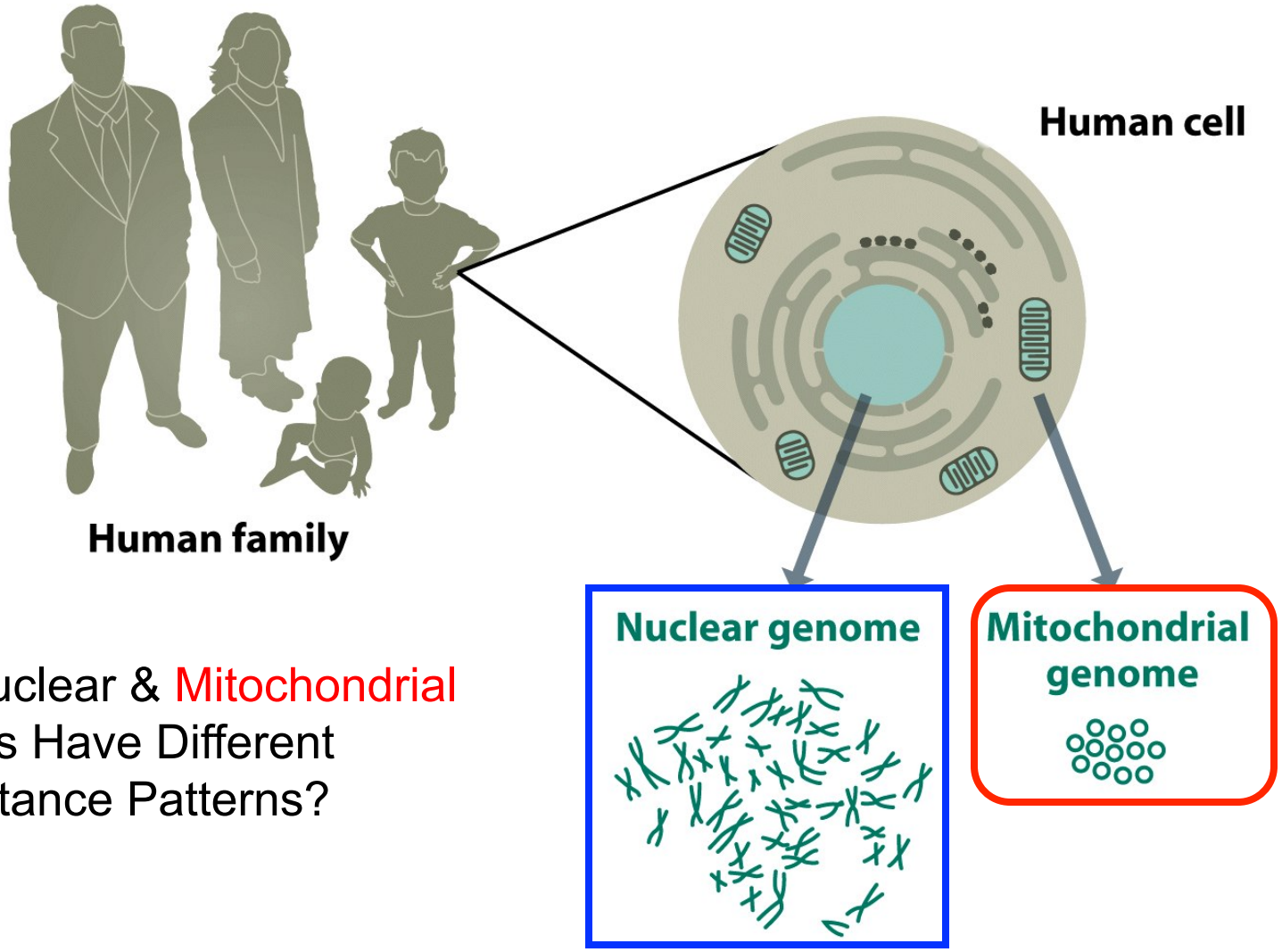
TUSKEGEE UNIVERSITY

UCDAVIS
UNIVERSITY OF CALIFORNIA

# Human Cells Have Two Genomes



# ……One in the Nucleus and One in the Mitochondria




Do Nuclear & Mitochondrial Genes Have Different Inheritance Patterns?

Figure 1-1 Genomes 3 (© Garland Science 2007)

# The Nuclear and Mitochondrial Genomes Differ in Size & Shape

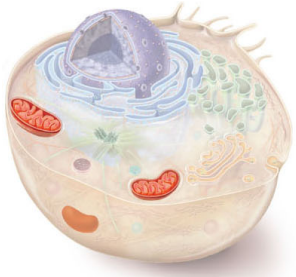
| Nuclear | Mitochondrial |
|---|---|
| 3.2 Mb | 17 kb |
| 25,000 Genes | 30 Genes |
| 24 Linear Pieces | 1 Circle |

**Table 9.1: The human nuclear and mitochondrial genomes**

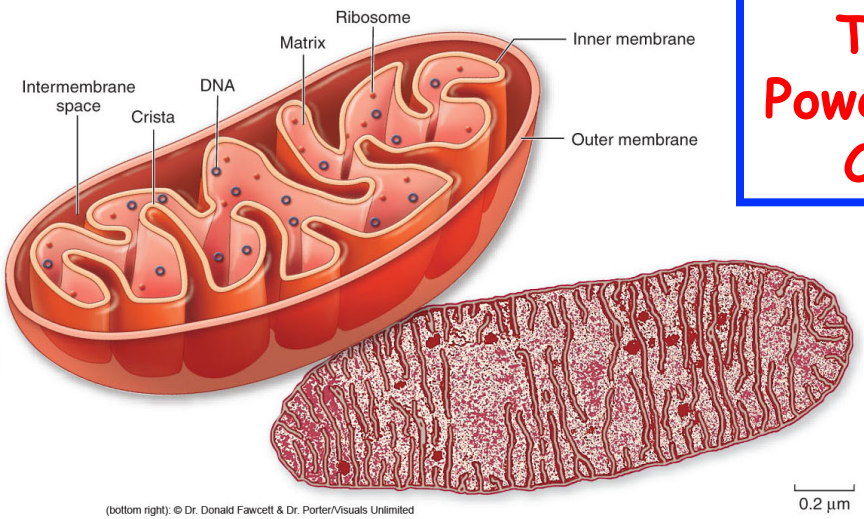
| | Nuclear genome | Mitochondrial genome |
|---|---|---|
| Size | 3200 Mb | 16.6 kb |
| No. of different DNA molecules | 23 (in XX cells) or 24 (in XY cells); all linear | One circular DNA molecule |
| Total no. of DNA molecules per cell | 46 in diploid cells, but varies according to ploidy | Often several thousands (but variable – see *Box 9.1*) |
| Associated protein | Several classes of histone and nonhistone protein | Largely free of protein |
| No. of genes | ~ 30 000–35 000 | 37 |
| Gene density | ~ 1/100 kb | 1/0.45 kb |
| Repetitive DNA | Over 50% of genome, see *Figure 9.1* | Very little |
| Transcription | The great bulk of genes are transcribed individually (*monocistronic transcription units*) | Co-transcription of multiple genes from both the heavy and the light strands (*polycistronic transcription units*) |
| Introns | Found in most genes | Absent |
| % of coding DNA | ~ 1.5% | ~ 93% |
| Codon usage | See *Figure 1.22* | See *Figure 1.22* |
| Recombination | At least once for each pair of homologs at meiosis | Not evident |
| Inheritance | Mendelian for sequences on X and autosomes; paternal for sequences on Y | Exclusively maternal |

# Mitochondria Power Human Cells and Contain a Circular Genome

**Each Cell Has Many Mitochondria**








(bottom right): © Dr. Donald Fawcett & Dr. Porter/Visuals Unlimited

**Makes ATP Energy That Powers All Cells!**

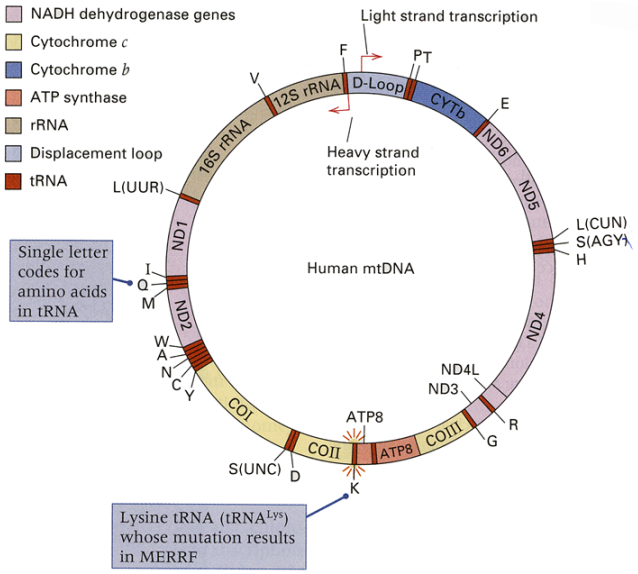



**Figure 16.3** Genes in human mitochondrial DNA. The tRNA genes are indicated by the one-letter amino acid symbols; hence $tRNA^{Lys}$ is denoted K. The positions of these and other genes in the mitochondrial DNA are indicated by color according to the key at the upper left. The arrows indicate the promoters for transcription of the heavy and light strands. [Courtesy of N-G. Larsson and D. A. Clayton. With permission, from the *Annual Review of Genetics* 29: 151. Copyright 1995 by Annual Reviews, www.AnnualReviews.org.]

**Semi-Autonomous Genome**

**DNA**

**Divides**
**Transcription**
**Translation**

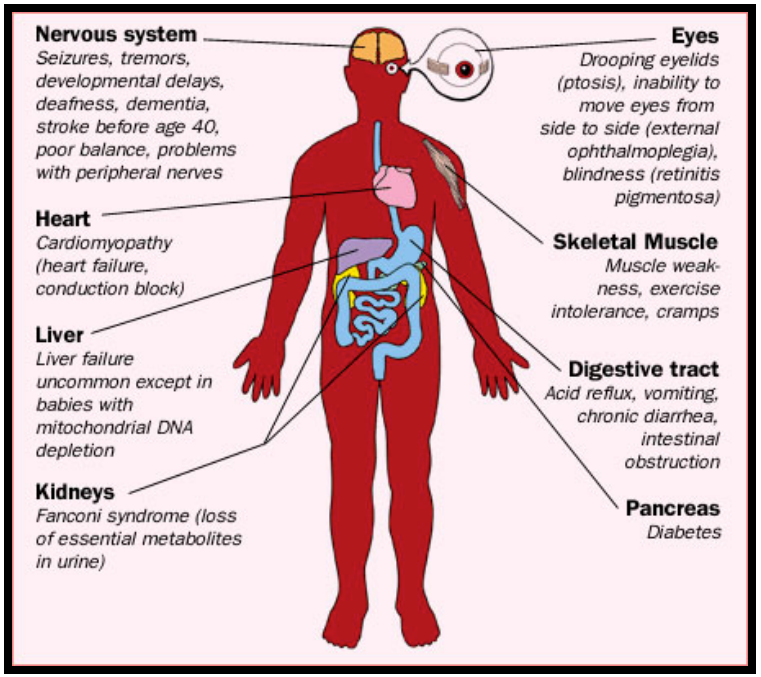
→ **Mitochondrial Proteins**

UNITED MITOCHONDRIAL DISEASE FOUNDATION.
HOPE. ENERGY. LIFE.

# Mitochondrial DNA Diseases

## Affect 1/400 People

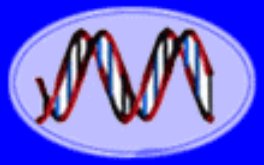
- Alpers Disease
- Barth syndrome
- Beta-oxidation Defects
- Carnitine-Acyl-Carnitine Deficiency
- Carnitine Deficiency
- Creatine Deficiency Syndromes
- Co-Enzyme Q10 Deficiency
- Complex I Deficiency
- Complex II Deficiency
- Complex III Deficiency
- Complex IV Deficiency
- Complex V Deficiency
- COX Deficiency
- CPEO
- CPT I Deficiency
- CPT II Deficiency
- Glutaric Aciduria Type II
- KSS
- Lactic Acidosis
- LCAD
- LCHAD
- Leigh Disease or Syndrome
- LHON
- LIC (Lethal Infantile Cardiomyopathy)
- Luft Disease
- MAD
- MCAD
- MELAS
- MERRF
- MIRAS
- Mitochondrial Cytopathy
- Mitochondrial DNA Depletion
- Mitochondrial Encephalopathy
- Mitochondrial Myopathy
- MNGIE
- NARP
- Pearson Syndrome
- Pyruvate Carboxylase Deficiency
- Pyruvate Dehydrogenase Deficiency
- POLG Mutations
- Respiratory Chain
- SCAD
- SCHAD
- VLCAD




**Treatment**

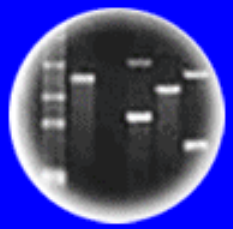
***At this time, there are no cures for these disorders.***

*MERRF Is Rare – Affecting 1/400,000 People*



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

# MERRF: A Mitochondrial Disease Example

## *Myoclonic Epilepsy With Ragged-Red Fiber Syndrome*

**MERRF**

Long Name: Myoclonic Epilepsy and Ragged-Red Fiber Disease.

Symptoms: Myoclonus, epilepsy, progressive ataxia, muscle weakness and degeneration, deafness, and dementia.

Cause: Mitochondrial DNA point mutations: A8344G, T8356C **Serine tRNA**

MERRF is a progressive multi-system syndrome usually beginning in childhood, but onset may occur in adulthood. The rate of progression varies widely. Onset and extent of symptoms can differ among affected siblings.

The classic features of MERRF include:

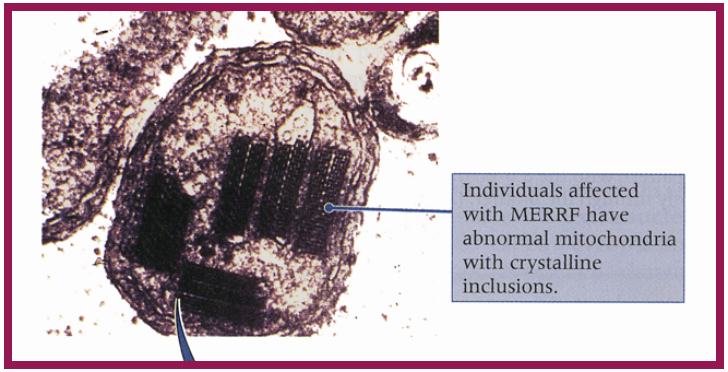
- Myoclonus (brief, sudden, twitching muscle spasms) – the most characteristic symptom
- Epileptic seizures
- Ataxia (impaired coordination)
- Ragged-red fibers (a characteristic microscopic abnormality observed in muscle biopsy of patients with MERRF and other mitochondrial disorders) Additional symptoms may include: hearing loss, lactic acidosis (elevated lactic acid level in the blood), short stature, exercise intolerance, dementia, cardiac defects, eye abnormalities, and speech impairment.

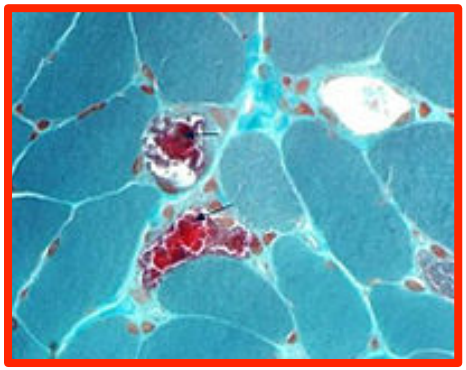
Although a few cases of MERRF are sporadic, most cases are maternally inherited due to a mutation within the mitochondria. The most common MERRF mutation is A8344G, which accounted for over 80% of the cases (GeneReview article). Four other mitochondrial DNA mutations have been reported to cause MERRF. While a mother will transmit her MERRF mutation to all of her offspring, some may never display symptoms.

As with all mitochondrial disorders, there is no cure for MERRF. Therapies may include coenzyme Q10, L-carnitine, and various vitamins, often in a "cocktail" combination. Management of seizures usually requires anticonvulsant drugs. Medications for control of other symptoms may also be necessary.

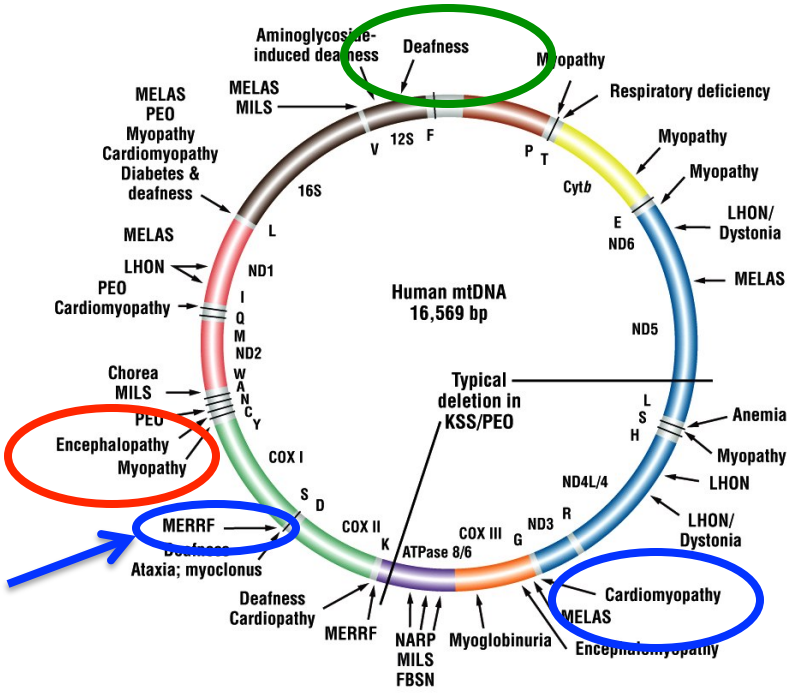
The prognosis for MERRF varies widely depending on age of onset, type and severity of symptoms, organs involved, and other factors.

Sources: Dr. Rolf Luft; The development of mitochondrial medicine. [Review] ; *Proceedings of the National Academy of Sciences of the United States of America* ; 1994 ; 91(19) ; 8731-8 & DiMauro



# The Circular Mitochondrial Genome is Inherited Maternally

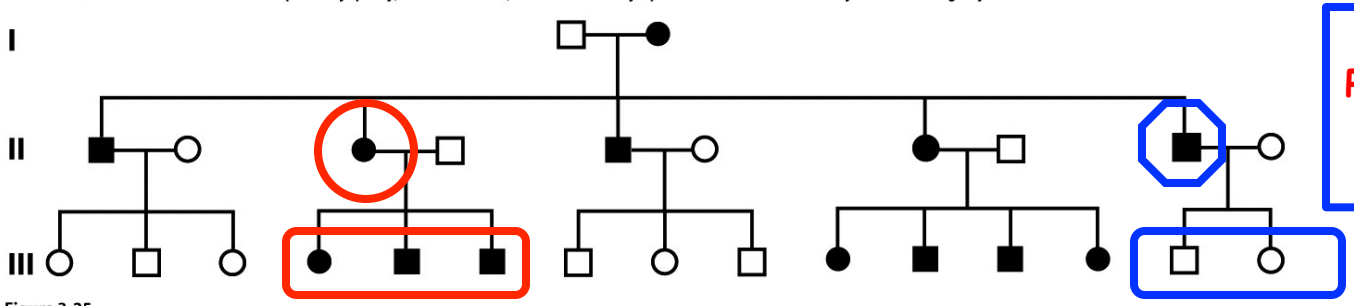



## Disease Genes Present on the Mitochondrial Genome

**Many Affect Muscles Because Mitochondria Produce Energy Needed For Muscle Activity**

**Diseases:**

| | | | |
|---|---|---|---|
| MERRF | Myoclonic epilepsy and ragged red fiber disease | MMC | Maternally inherited myopathy and cardiomyopathy |
| LHON | Leber hereditary optic neuropathy | PEO | Progressive external opthalmoplegia |
| NARP | Neurogenic muscle weakness, ataxia, and retinitis pigmentosum | KSS | Kearns–Sayre syndrome |
| MELAS | Mitochondrial encephalomyopathy, lactic acidosis, and strokelike symptoms | MILS | Maternally inherited Leigh syndrome |




**Note: Passed on From Mother to All Children**

Figure 3-25
*Introduction to Genetic Analysis, Ninth Edition*


***Provide a Hypothesis For the Variation in Disease Symptoms***

NUCLEAR TRANSPLANTATION

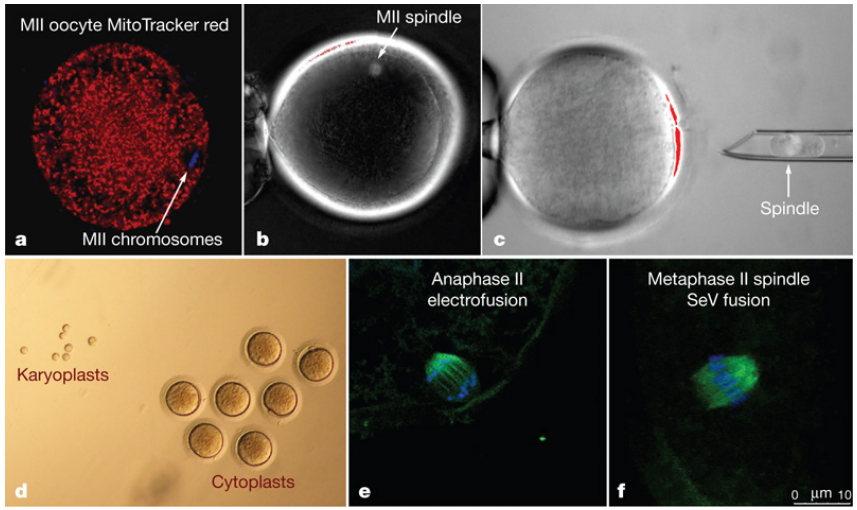
# Researchers Prevent Inheritance of Faulty Mitochondria in Monkeys

Vol 461|17 September 2009|doi:10.1038/nature08368 nature

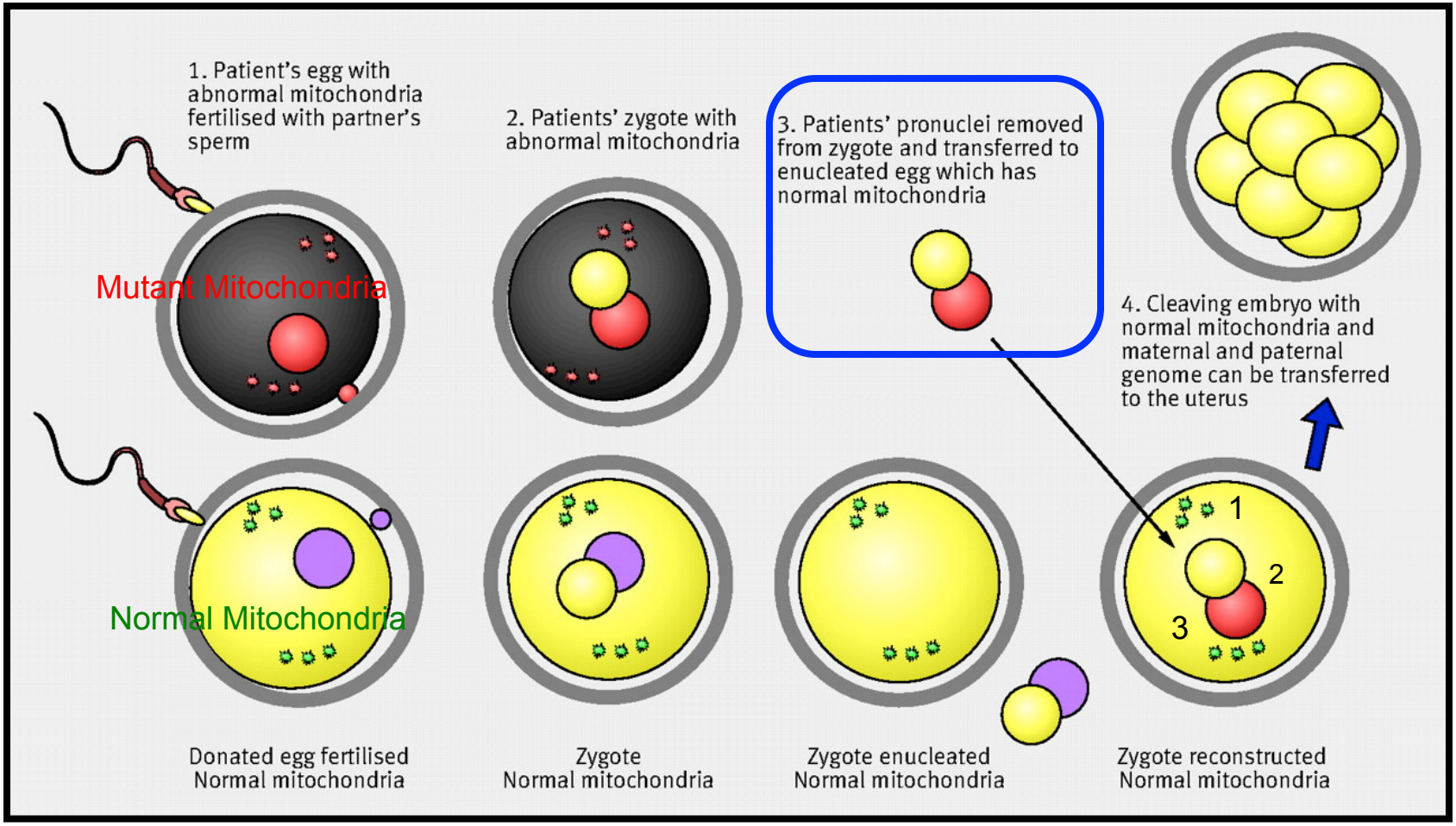
Nature 461, September 17, 2009

ARTICLES

## Mitochondrial gene replacement in primate offspring and embryonic stem cells




# Future Mitochondrial Gene Replacement Therapy




**Note: The Zygote Contains THREE Genomes --**
**One from Mother, One From Father, and One From Donor Mitochondria**

# Ethics of mitochondrial gene replacement: from bench to bedside

The prospect of using mitochondrial gene replacement to conceive children free of mitochondrial disease highlights the need for a sound ethical framework for reproductive genetic technology, say **Annelien Bredenoord** and **Peter Braude**

- How to Test Whether It Works?
- Testing on Live Human Embryos
- When To Test Its Effectiveness
- Safety & Long-Term Potential Problems
- Research Protocols & Oversight
- Informed Consent of Parents

British Medical Journal, January 8, 2011,342, 87-89

# RFLPs Can Be Used to Identify Individuals Using Mitochondrial DNAs

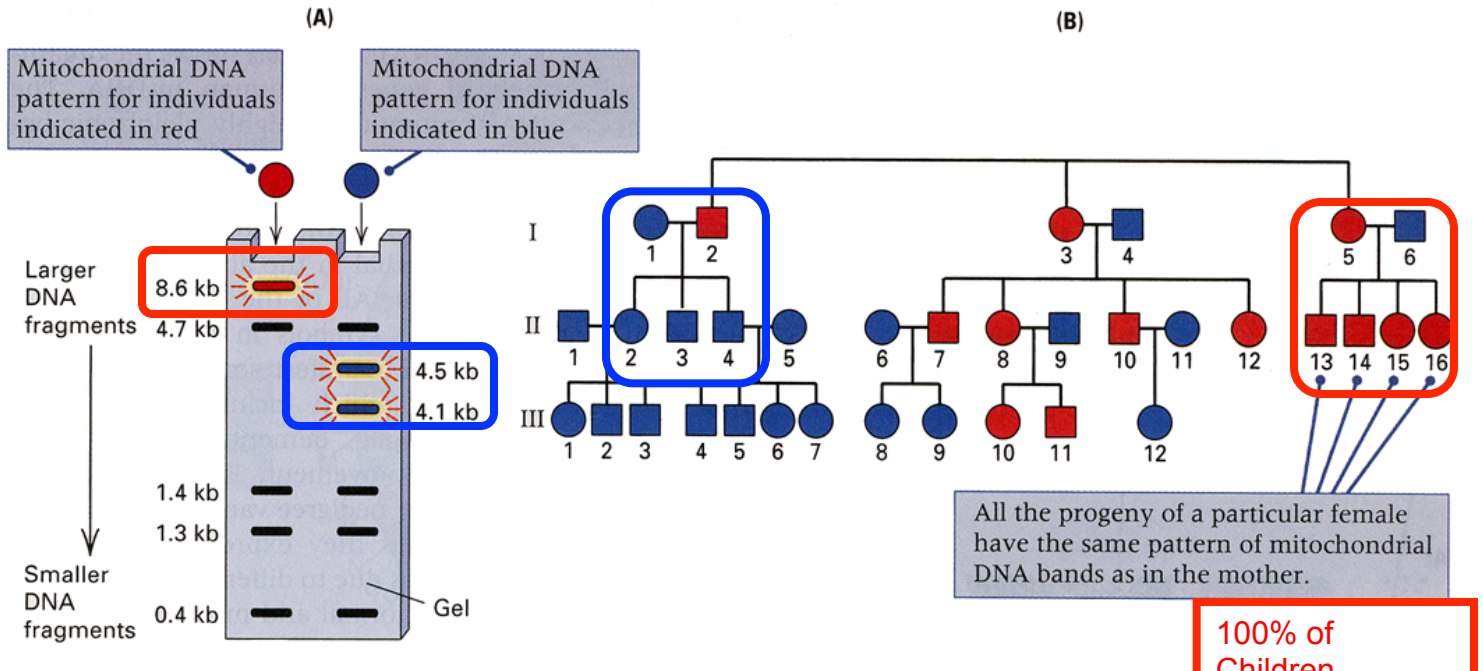
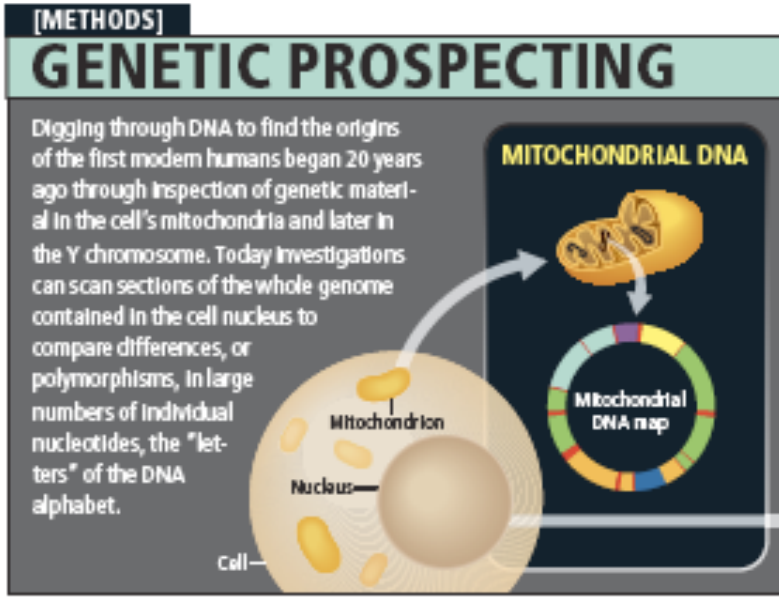


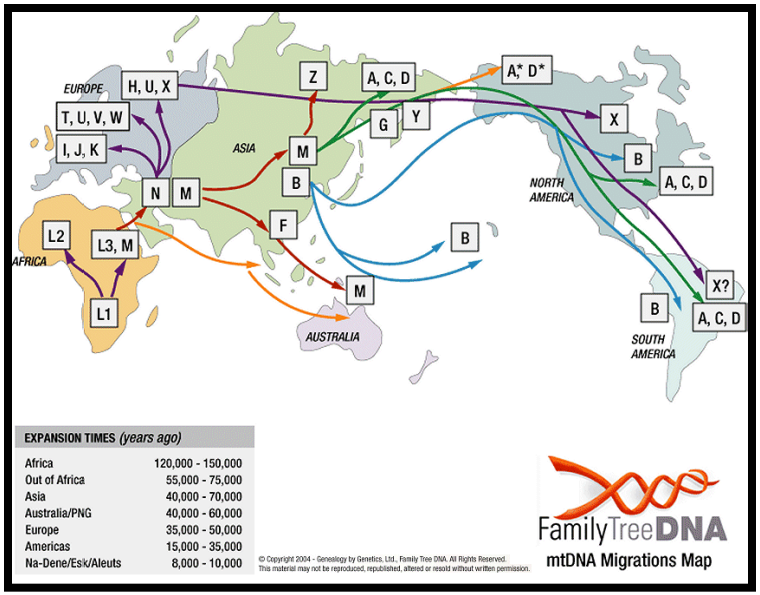

**Figure 16.1** Maternal inheritance of human mitochondrial DNA. (A) Pattern of DNA fragments obtained when mitochondrial DNA is digested with the restriction enzyme *Hae*II. The DNA type at the left includes a fragment of 8.6 kb (red). The DNA type at the right contains a cleavage site for *Hae*II within the 8.6-kb fragment, which results in smaller fragments of 4.5 kb and 4.1 kb (blue). (B) Pedigree showing maternal inheritance of the DNA pattern with the 8.6-kb fragment (red symbols). The mitochondrial DNA type is transmitted only through the mother. [After D. C. Wallace. 1989. *Trends in Genetics* 5: 9.]

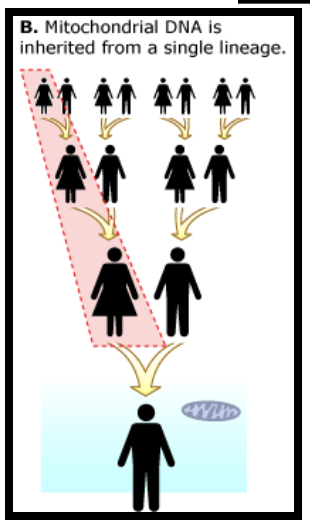
100% of Children Have 8.6 kb DNA Fragment From Mother

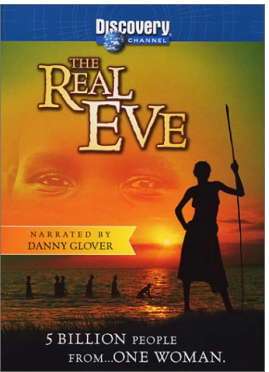
**Note How Mitochondrial RFLP Markers Are Inherited !!**

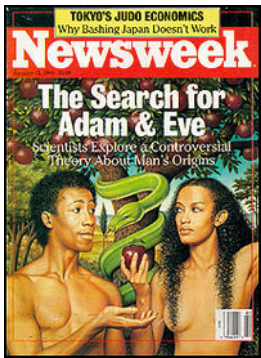
# Tracing Human Populations Using Mitochondrial DNA Polymorphisms - Back to Eve!



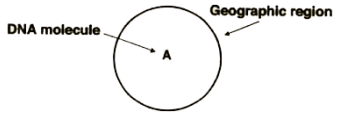
# Oldest Populations Contain the Most Diversity

**Analysis of human mtDNA led to the Mitochondrial Eve Hypothesis**

In the 1980s, Allan Wilson pioneered the use of mtDNA to study human evolution.

In two papers published in 1987 and 1991, he and his colleagues at Cal proposed that we all come from a population of humans that lived in Africa approximately 200,000 years ago.

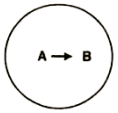
**Here's the logic behind the hypothesis.**



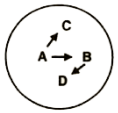

Start with A.

Time

Detected By Using Specific Markers (RFLPs or SNPS)



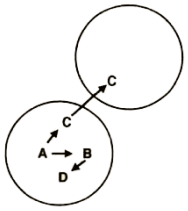

Mutation generates B from A;
now have individuals with both A and B DNAs in population.




Additional mutations generate diversity;
now have individuals with both A, B, C and D DNAs.

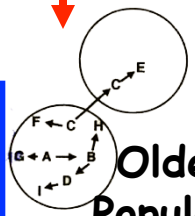
Newest Population

Subset of "Old" Variants + New Variants




C migrates to form separate population.

Lots of "Old" Variants




Oldest Population

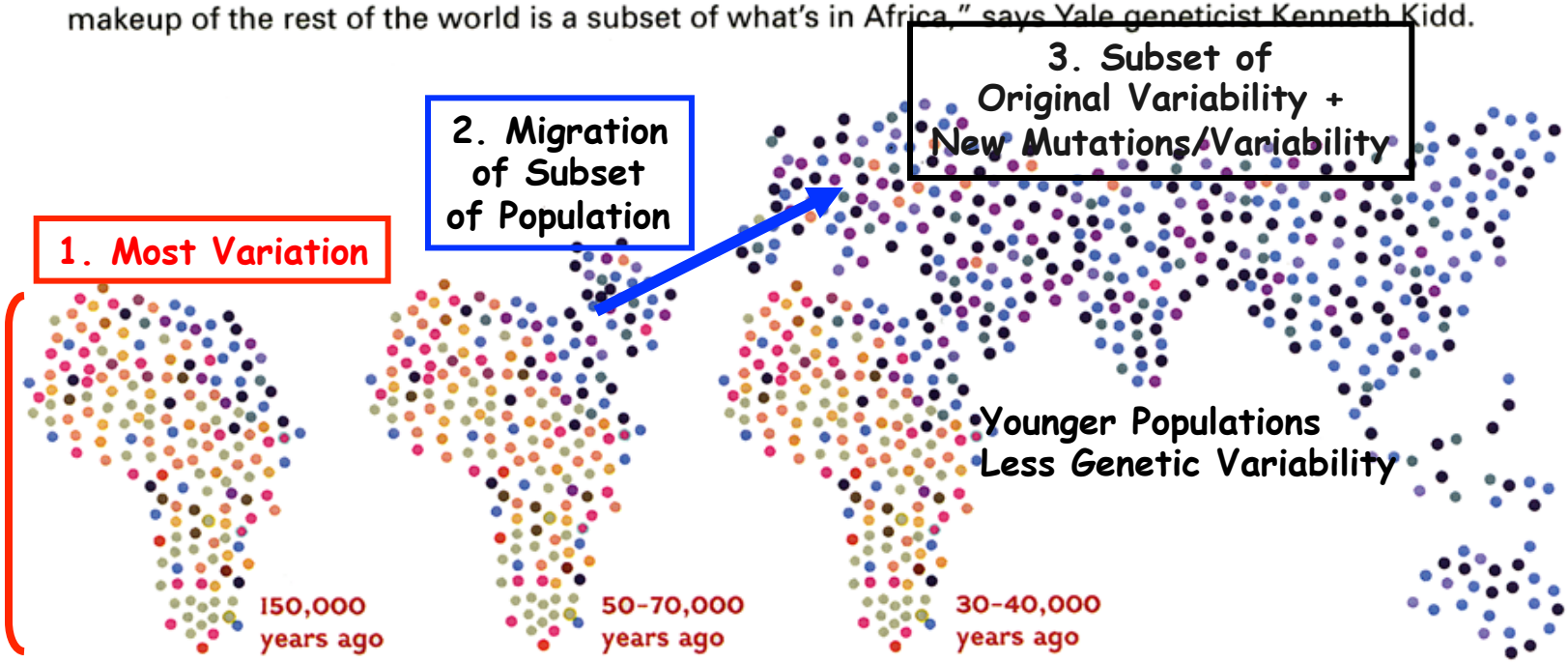
Additional mutations diversify DNAs in populations:
original population more diverse (A, B, C, D, F, G, H, I) than newer population (C, E).

Old Variants Trace Ancient Lineage. New Variants Mark New Populations SPECIFICALLY

# Most Genetic Diversity Originated in the Founder Populations to Modern Humans!

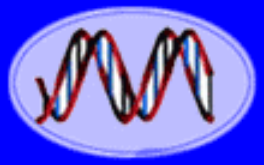
## Diverse From the Start

The diversity of genetic markers is greatest in Africa (multicolored dots in map), indicating it was the earliest home of modern humans. Only a handful of people, carrying a few of the markers, walked out of Africa (center) and, over tens of thousands of years, seeded other lands (right). "The genetic makeup of the rest of the world is a subset of what's in Africa," says Yale geneticist Kenneth Kidd.



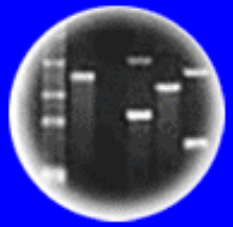

Genetic Variation Proportional to Population Age

Markers From Original Population + New Markers For "New" Population



DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria

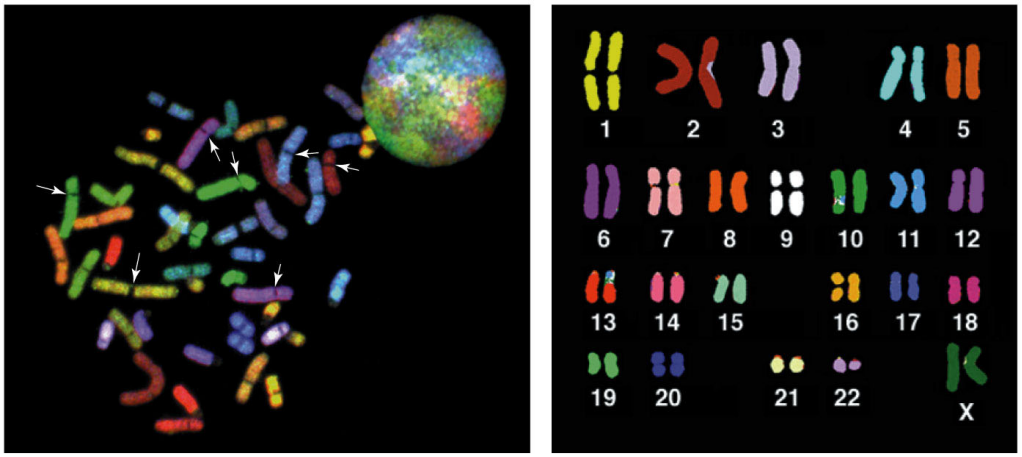


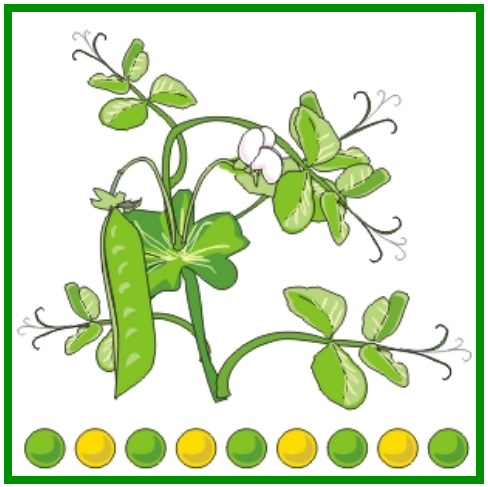
DNA Fingerprinting

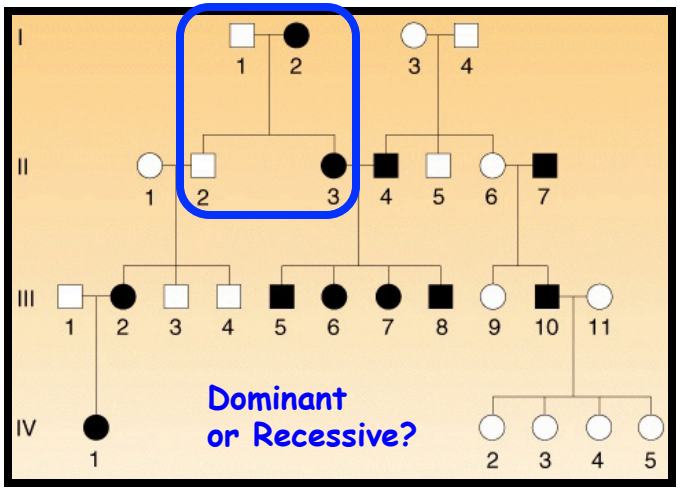
Cloning: Ethical Issues
and Future Consequences

Plants of Tomorrow

# The Nuclear Genome

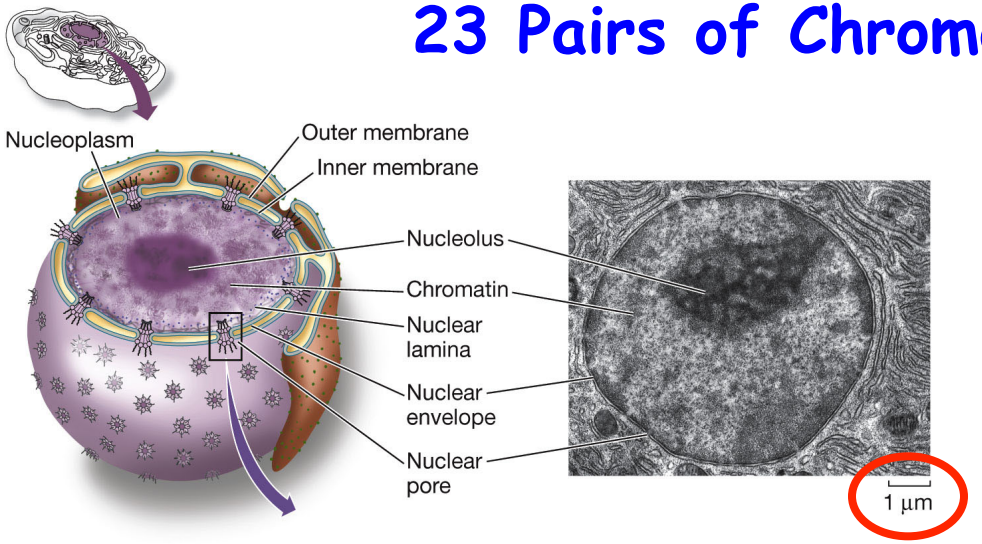


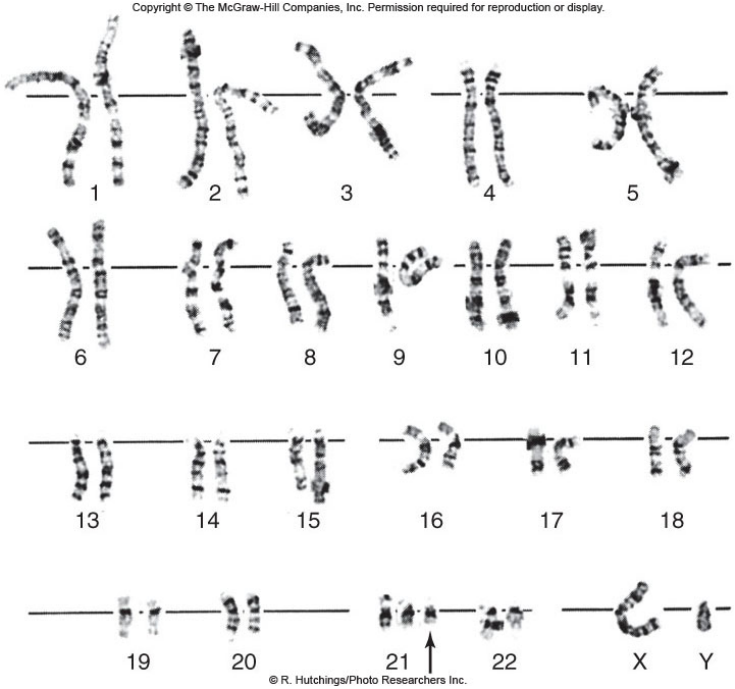



**Note: Gene is Inherited in a Mendelian Pattern**

# The Nucleus Is A Complex Organelle With 23 Pairs of Chromosomes (Humans)

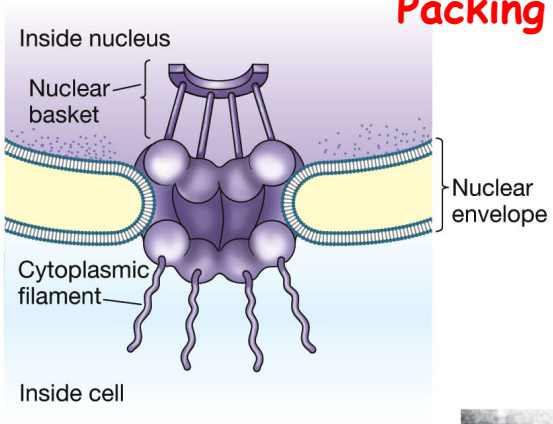


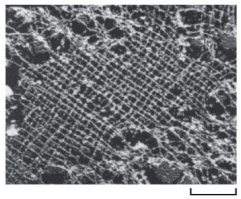


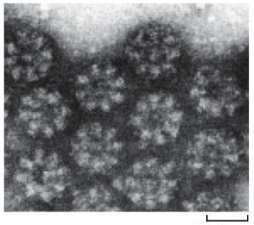


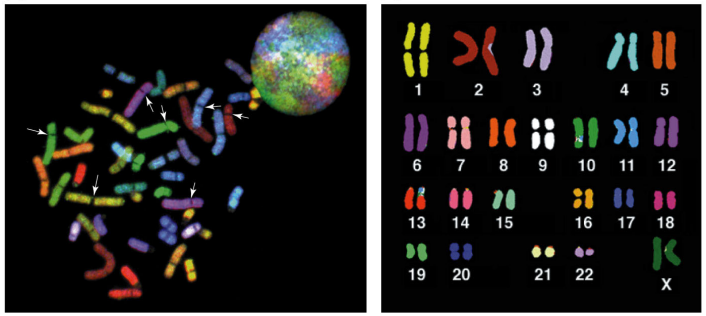

**Packing Problem?**




**RNA & Protein Transport**



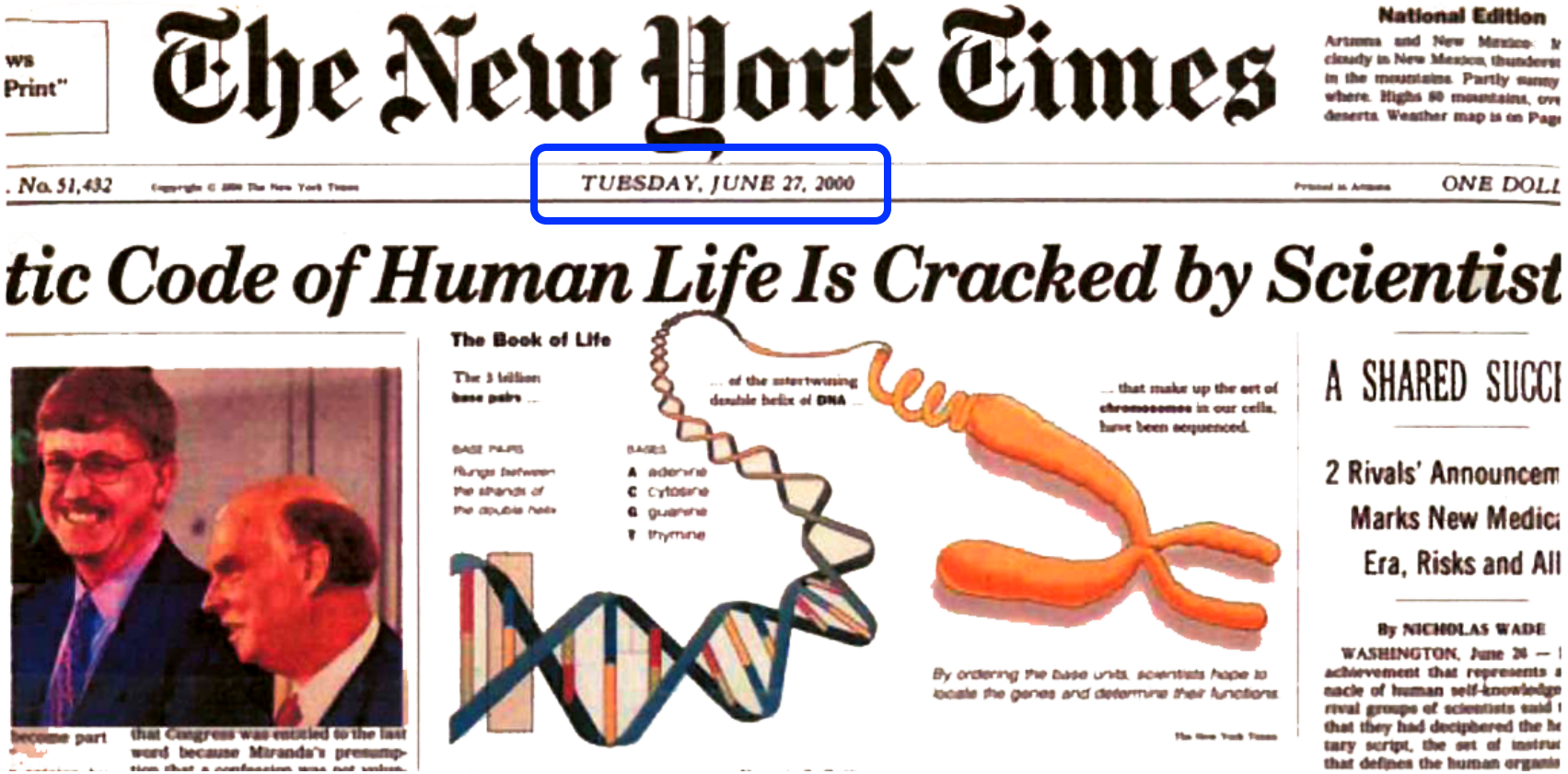

**The Human Genome End to End is 1.1 Meters in Length!!!!!!**

**Note: Chromosome Sizes & Bands = Markers**

*LIFE 8e*, Figure 4.8 (Part 2)

# The Human Genome Was Sequenced Ten Years Ago!

## The Human Genome Project



The New York Times

National Edition

TUESDAY, JUNE 27, 2000

ONE DOLL

tic Code of Human Life Is Cracked by Scientist

The Book of Life

A SHARED SUCCI

2 Rivals' Announcem

Marks New Medic

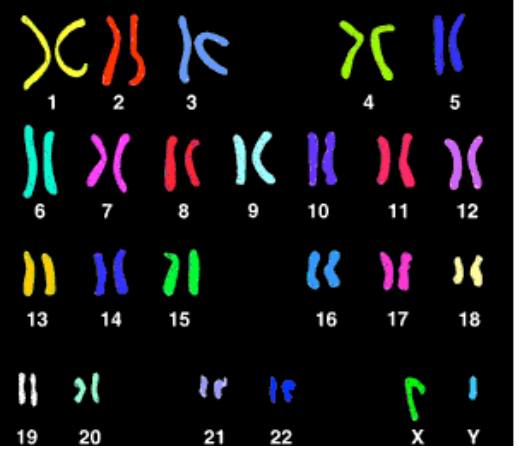
Era, Risks and All

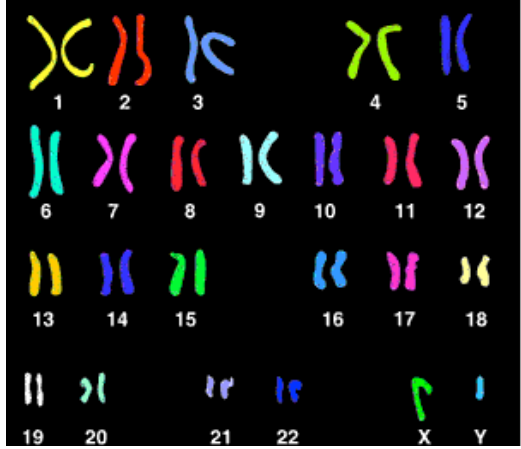
By NICHOLAS WADE

WASHINGTON, June 26 —

Public & Private Effort Using Different Strategies - A Race!

3 Billion Dollars & Took 15 Years



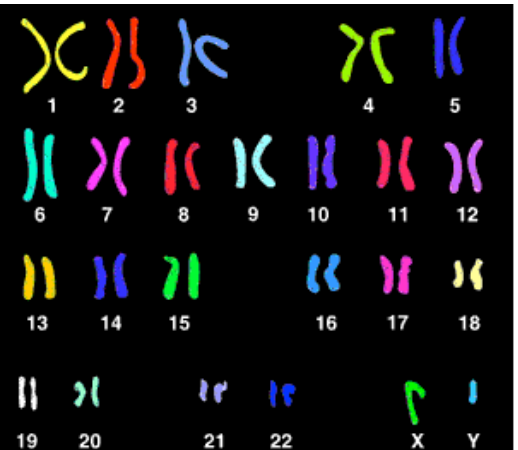


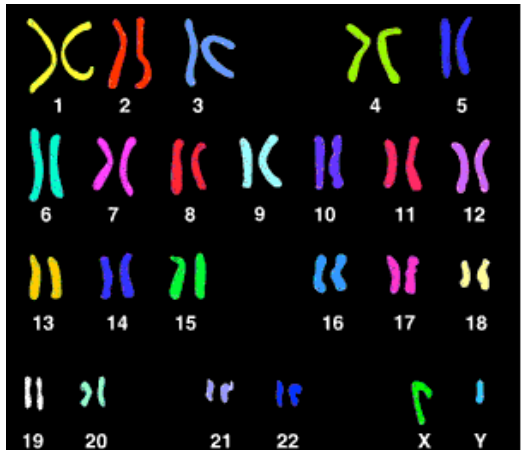
# The Human Genome

Large

Small

| Chromosome | Including gaps | Sequenced | Gaps |
|---|---|---|---|
| 1 | 247,249,719 | 224,999,719 | 22,250,000 |
| 2 | 242,951,149 | 237,712,649 | 5,238,500 |
| 3 | 199,501,827 | 194,704,827 | 4,797,000 |
| 4 | 191,273,063 | 187,297,063 | 3,976,000 |
| 5 | 180,857,866 | 177,702,766 | 3,155,100 |
| 6 | 170,899,992 | 167,273,992 | 3,626,000 |
| 7 | 158,821,424 | 154,952,424 | 3,869,000 |
| 8 | 146,274,826 | 142,612,826 | 3,662,000 |
| 9 | 140,273,252 | 120,143,252 | 20,130,000 |
| 10 | 135,374,737 | 131,624,737 | 3,750,000 |
| 11 | 134,452,384 | 131,130,853 | 3,321,531 |
| 12 | 132,349,534 | 130,303,534 | 2,046,000 |
| 13 | 114,142,980 | 95,559,980 | 18,583,000 |
| 14 | 106,368,585 | 88,290,585 | 18,078,000 |
| 15 | 100,338,915 | 81,341,915 | 18,997,000 |
| 16 | 88,827,254 | 78,884,754 | 9,942,500 |
| 17 | 78,774,742 | 77,800,220 | 974,522 |
| 18 | 76,117,153 | 74,656,155 | 1,460,998 |
| 19 | 63,811,651 | 55,785,651 | 8,026,000 |
| 20 | 62,435,964 | 59,505,253 | 2,930,711 |
| 21 | 46,944,323 | 34,171,998 | 12,772,325 |
| 22 | 49,691,432 | 34,851,332 | 14,840,100 |
| X | 154,913,754 | 151,058,754 | 3,855,000 |
| Y | 57,772,954 | 25,652,954 | 32,120,000 |
| M | 16,571 | 16,571 | 0 |
| **Total genome** | **3,080,436,051** | **2,858,034,764** | **222,401,287** |



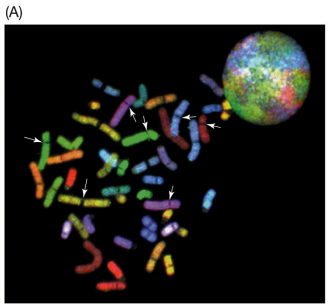


# Only A Small Fraction of the Human Genome Encodes Proteins



**TABLE 18.1 Classes of DNA Sequences Found in the Human Genome**

| Class | Frequency (%) | Description |
|---|---|---|
| Protein-encoding genes | 1.5 | Translated portions of the 25,000 genes scattered about the chromosomes |
| Introns | 24 | Noncoding DNA that constitutes the great majority of each human gene |
| Segmental duplications | 5 | Regions of the genome that have been duplicated |
| Pseudogenes (inactive genes) | 2 | Sequence that has characteristics of a gene but is not a functional gene |
| Structural DNA | 20 | Constitutive heterochromatin, localized near centromeres and telomeres |
| Simple sequence repeats | 3 | Stuttering repeats of a few nucleotides such as CGG, repeated thousands of times |
| Transposable elements | 45 | 21%: Long interspersed elements (LINEs), which are active transposons<br>13%: Short interspersed elements (SINEs), which are active transposons<br>8%: Retrotransposons, which contain long terminal repeats (LTRs) at each end<br>3%: DNA transposon fossils |





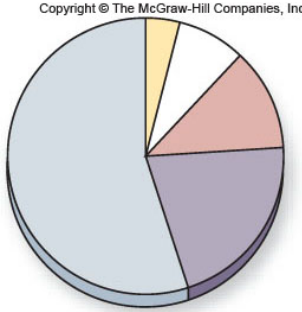
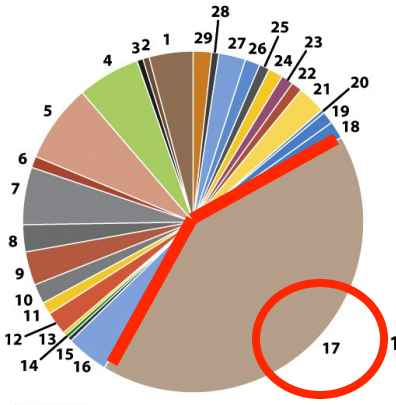




**Table 20.6 Average characteristics of genes in the human genome**

| Characteristic | Average |
|---|---|
| Number of exons | 8.8 |
| Size of internal exon | 145 bp |
| Size of intron | 3,365 bp |
| Size of 5′ untranslated region | 300 bp |
| Size of 3′ untranslated region | 770 bp |
| Size of coding region | 1,340 bp |
| Total length of gene | 27,000 bp |

Table 20-6

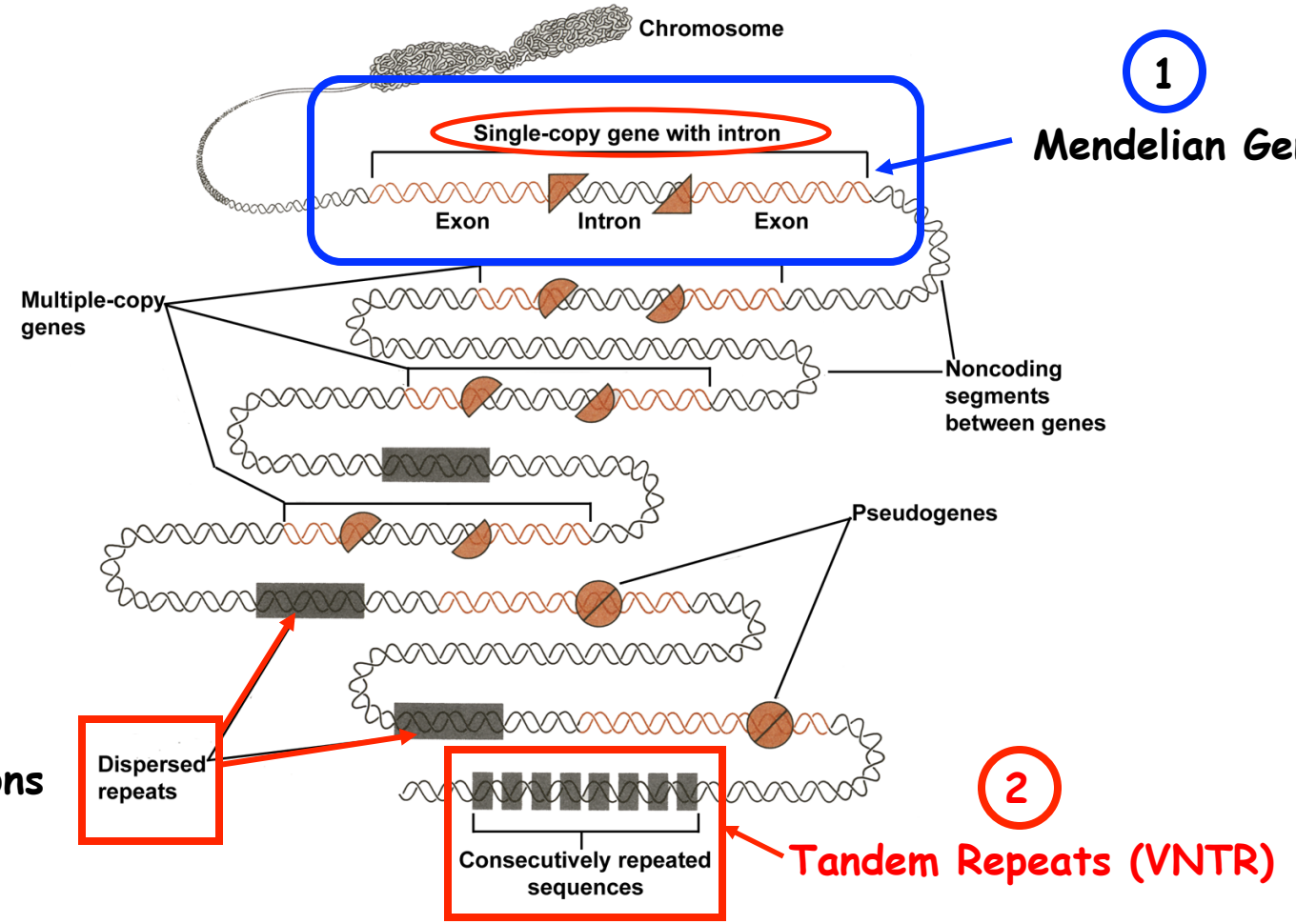
**The Human Genome Has DNA Sequences Present Once As Well as Repeated Many Times**




1. Miscellaneous
2. Viral protein
3. Transfer or carrier protein
4. Transcription factor
5. Nucleic acid enzyme
6. Signaling molecule
7. Receptor
8. Kinase
9. Select regulatory molecule
10. Transferase
11. Synthase and synthetase
12. Oxidoreductase
13. Lyase
14. Ligase
15. Isomerase
16. Hydrolase
17. Molecular function unknown
18. Transporter
19. Intracellular transporter
20. Select calcium-binding protein
21. Protooncogene
22. Structural protein of muscle
23. Motor
24. Ion channel
25. Immunoglobulin
26. Extracellular matrix
27. Cytoskeletal structural protein
28. Chaperone
29. Cell adhesion

**Human Genes are Large but Contain Mostly Introns**

# The Human Genome Landscape

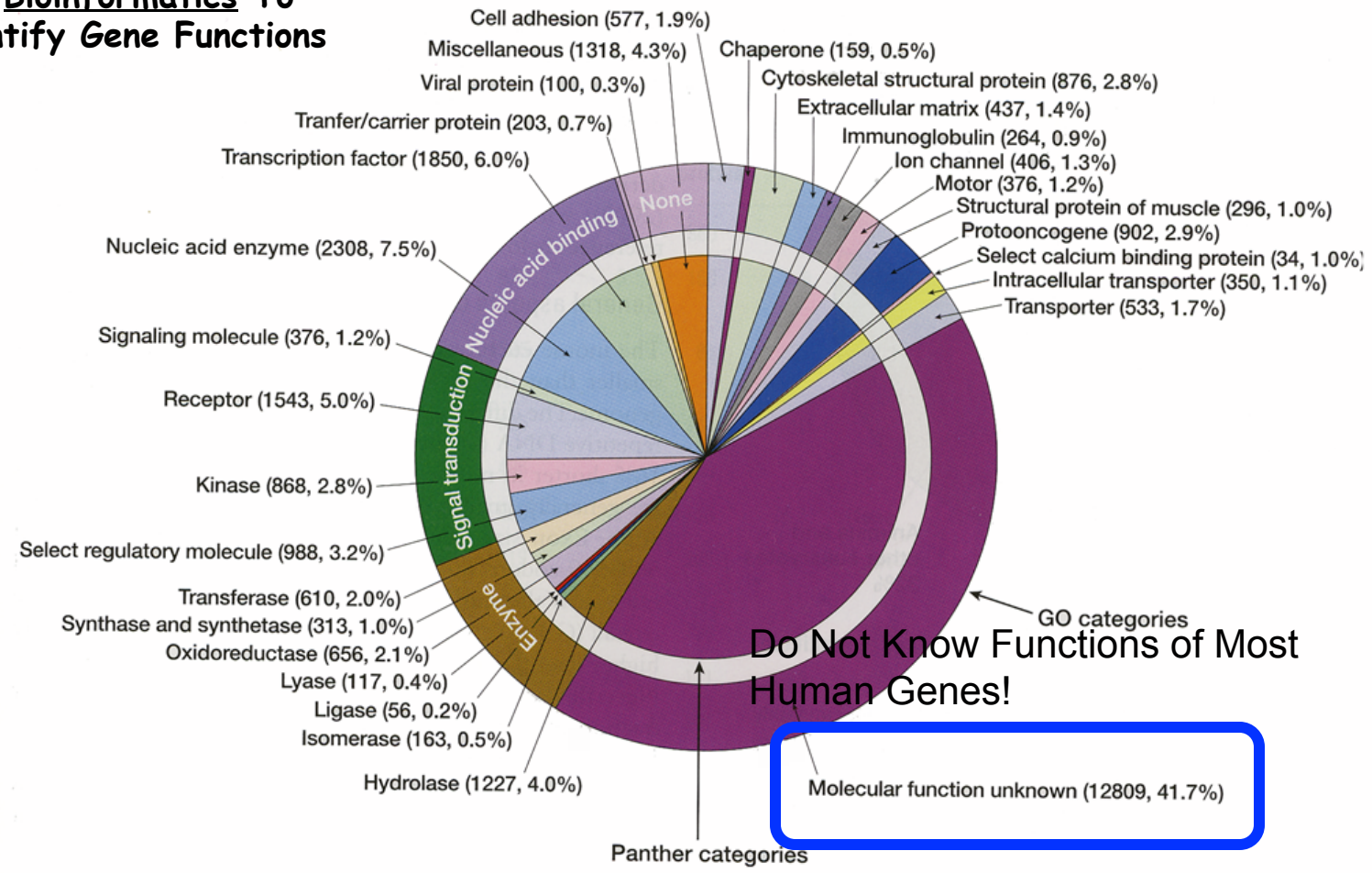



**Tandem Repeats or VNTRs are Useful for DNA Fingerprinting Studies!**

e.g., DIS80 Locus For Class DNA Fingerprint on Chromosome 4 Core = 16bp

# The Human Genome Contains ~25,000 Different Genes

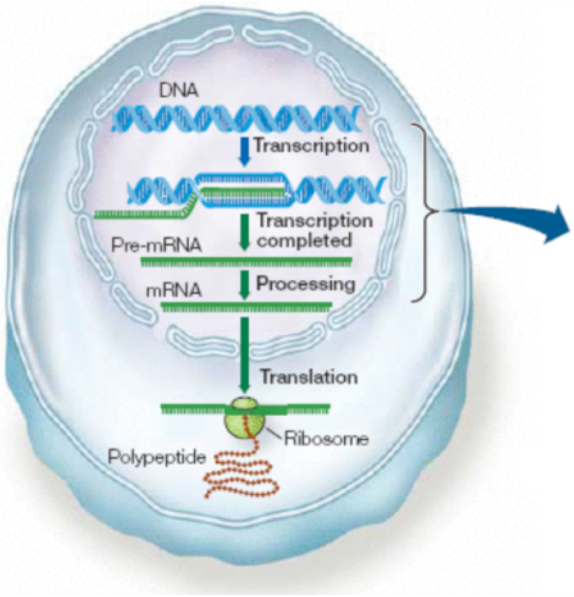
**Use Bioinformatics To Identify Gene Functions**



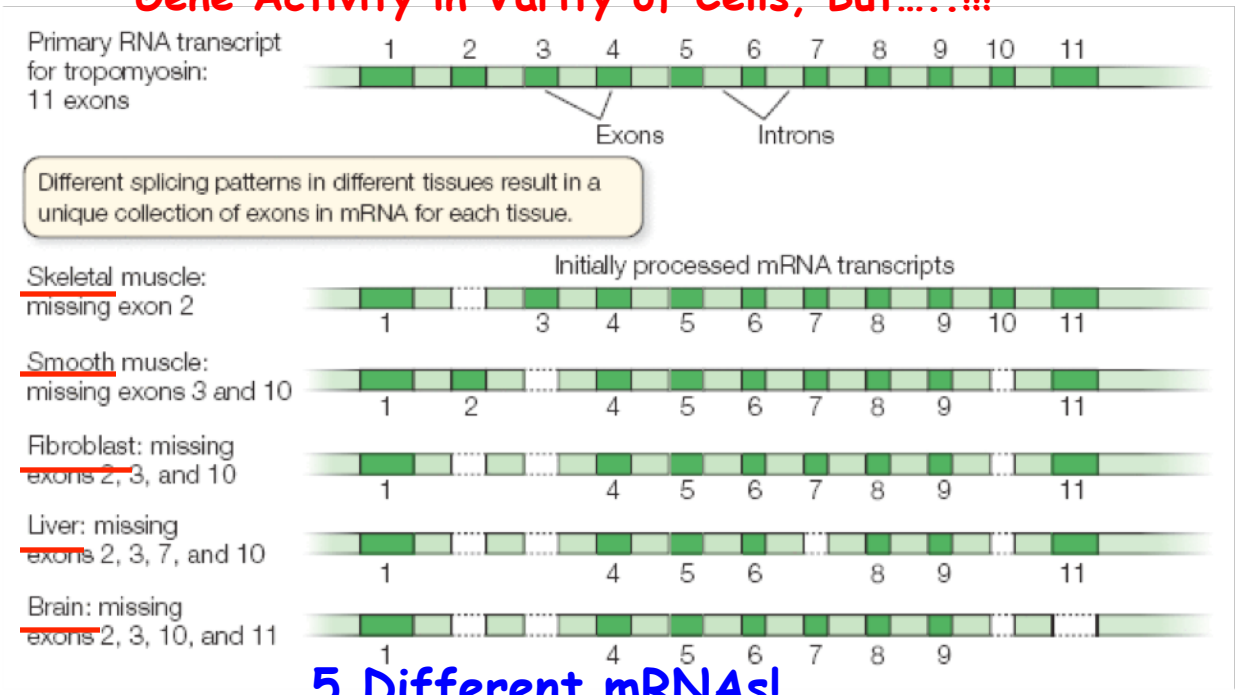

Do Not Know Functions of Most Human Genes!

**How Many Encoded Proteins? Alternative Splicing?**

# Alternative Splicing- One Gene → Several mRNAs & Proteins




## Gene Activity in Varity of Cells, But…..!!!



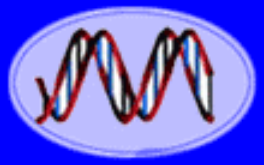

**5 Different mRNAs!**

**Different mRNA = Different Proteins = Different Functions!**

**Implication- Human Genome Has Only 25,000 Genes But Can Give Rise to Many More Proteins which Are Responsible For Producing the Phenotype**

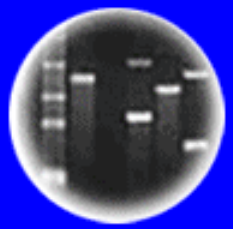
**Reason Why Human Genome Can Contain Same Number of Genes as Fly and Plant Genomes!!**

**Implications for Genetic Engineering? Use Specific cDNA!**



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

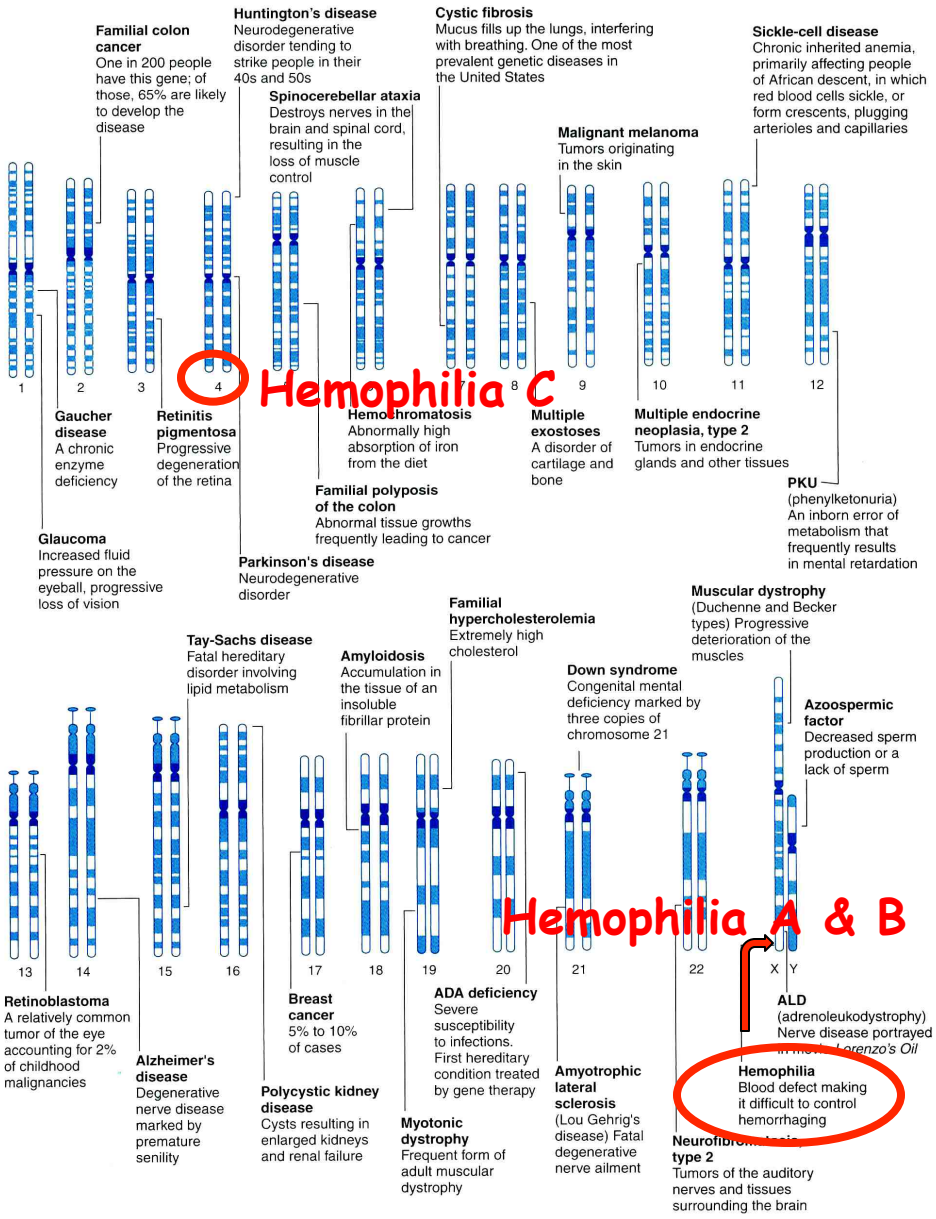


DNA Fingerprinting

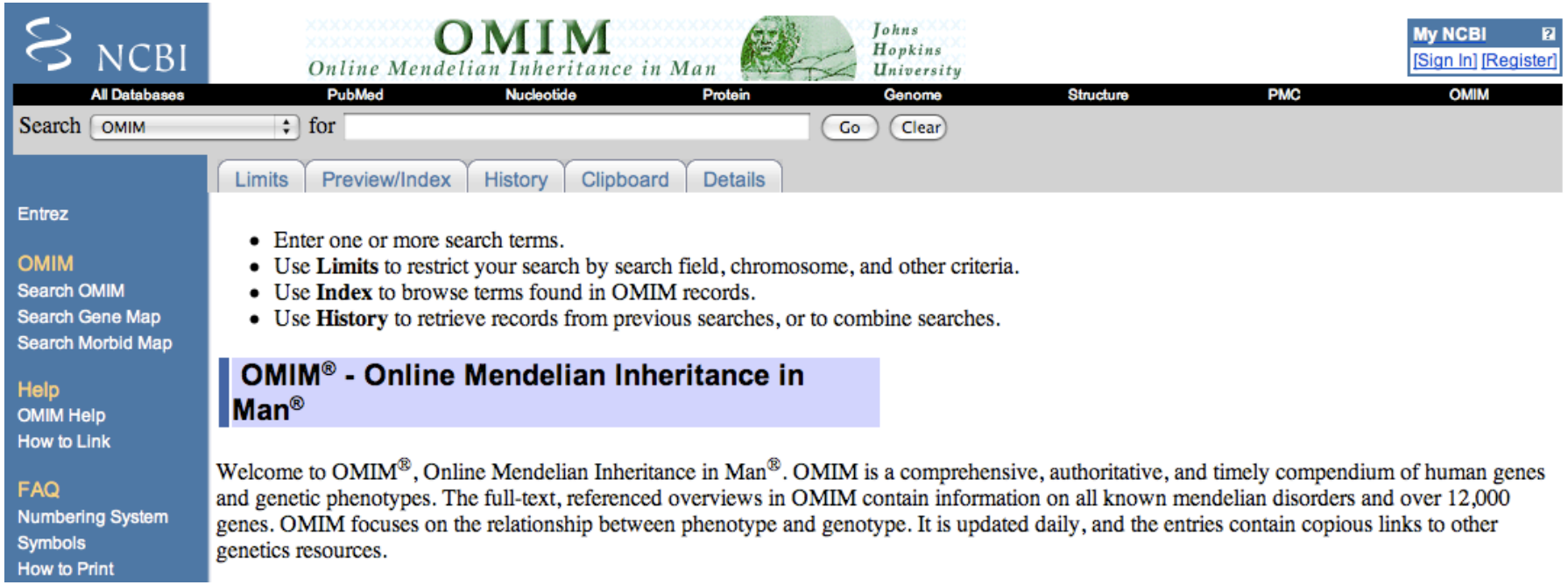
Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

# Human Disease Genes Are Present on All Chromosomes



# How Many Human Disease Genes Have Been Identified?




**There are ~25,000 Genes in The Human Genome**

1. ~3,407 Genes Correlate With a Disease Phenotype
2. The Molecular Basis of These Genetic Diseases Are Known (e.g., Sickle Cell Anemia, Hemophilia A, Tay-Sachs, Cystic Fibrosis, Duchene Muscular Dystrophy, Huntington Disease, etc,)

**Genetic Tests Exist For These Disease Genes**

OMIM, February 19, 2012

PERSPECTIVE

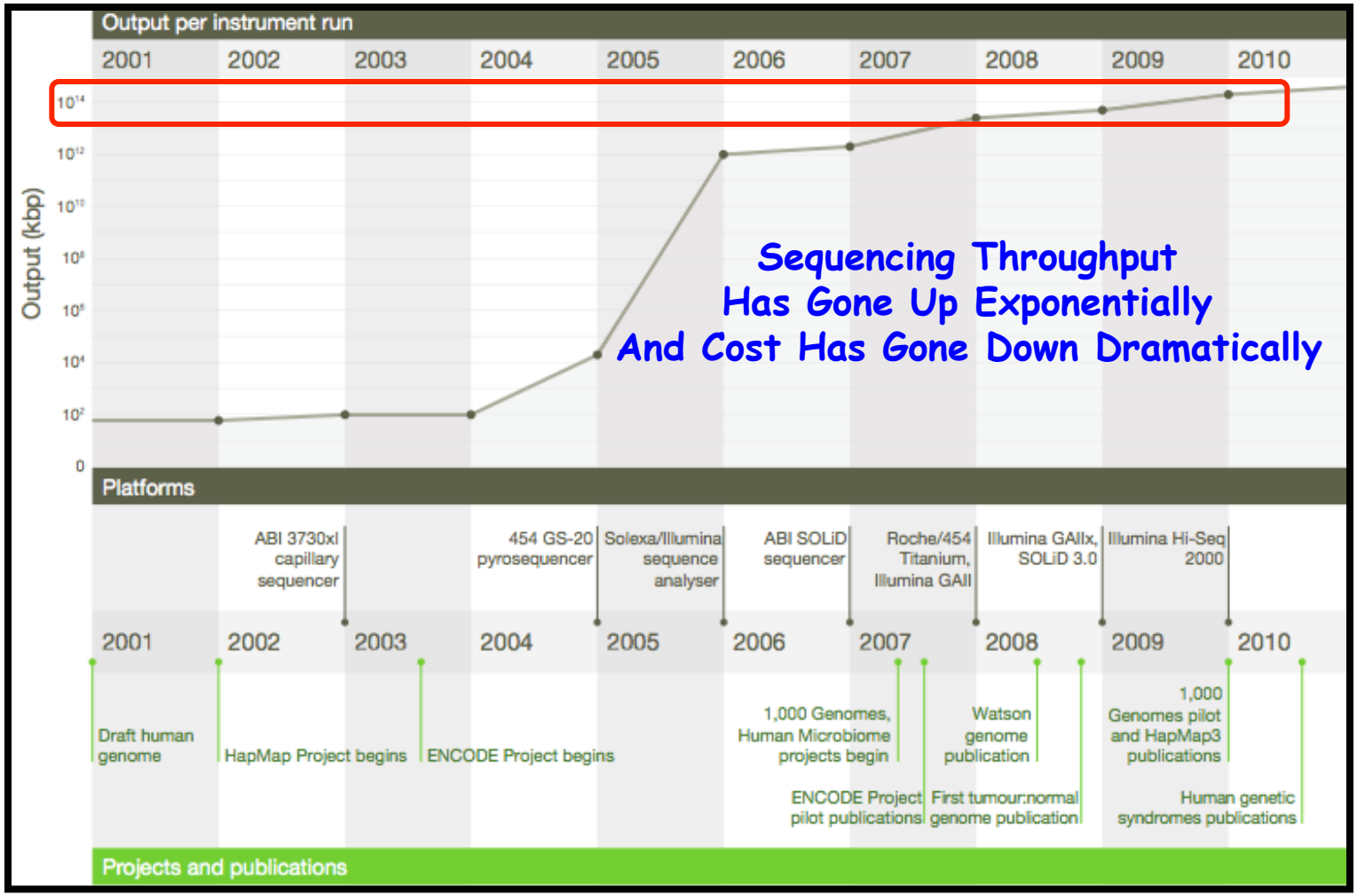
doi:10.1038/nature09796

# A decade's perspective on DNA sequencing technology

*Nature, February 10, 2011*

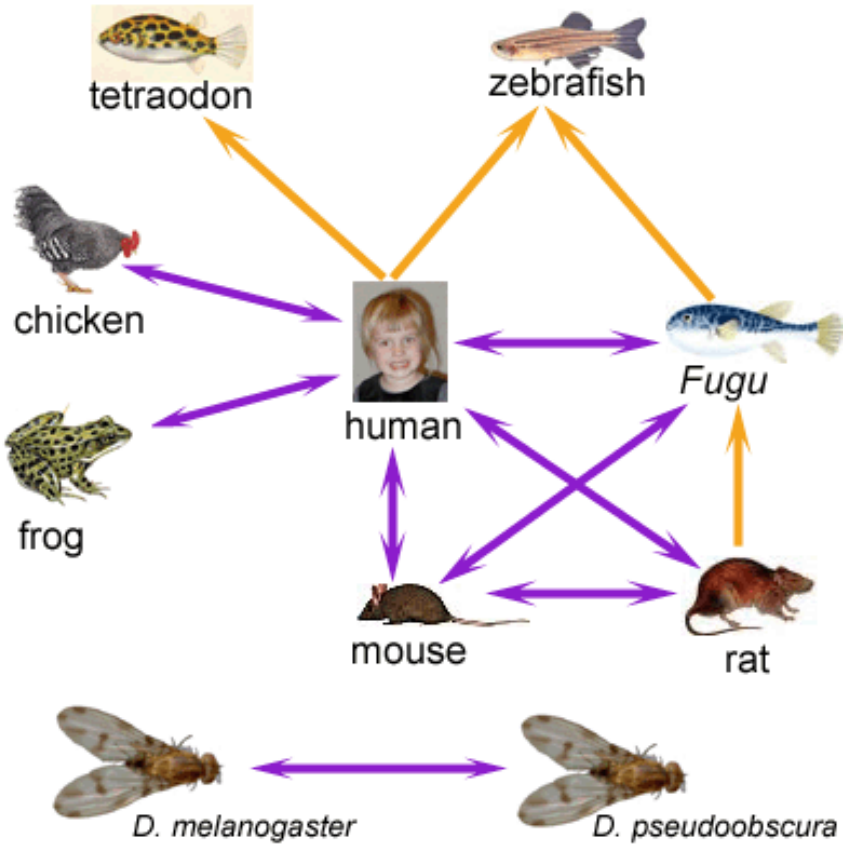
Elaine R. Mardis[1]

$10^{14}$ kbp



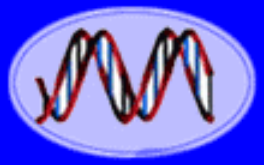
# The Genomes of Many Organisms Have Been Sequenced Providing New Knowledge About Our Origins and Cellular Functions

**The Future is Unlimited!**



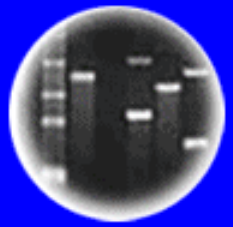

**173 Eukaryotic & 2,960 Prokaryotic Genomes Have Been Sequenced as of 2012 & 15,285 Genome Projects Ongoing!**
http://www.genomesonline.org/cgi-bin/GOLD/bin/gold.cgi

*Providing Thousands of New Genes and Proteins To Be Engineered For Practical Applications (e.g., cellulases in termite gut bacteria for biofuel production)*



DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria

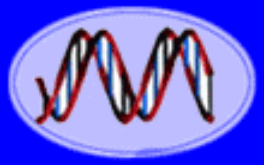


DNA Fingerprinting

Cloning: Ethical Issues
and Future Consequences

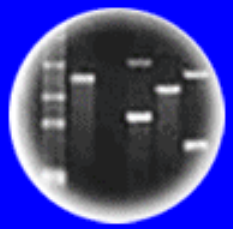
Plants of Tomorrow

# A New Comparative Genomics Field Has Emerged Allowing the Comparison of Entire Genomes!



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

# TABLE 17.1

## Representative Sequenced Genomes

| ORGANISM | HAPLOID GENOME SIZE (Mb) | NUMBER OF GENES | PROTEIN-CODING SEQUENCE |
|---|---|---|---|
| Bacteria | | | |
| *M. genitalium* | 0.58 | 485 | 88% |
| *H. influenzae* | 1.8 | 1,738 | 89% |
| *E. coli* | 4.6 | 4,377 | 88% |
| Yeasts | | | |
| *S. cerevisiae* | 12.5 | 5,770 | 70% |
| *S. pombe* | 12.5 | 4,929 | 60% |
| Plants | | | |
| *A. thaliana* | 115 | 28,000 | 25% |
| Rice | 390 | 37,544 | 12% |
| Animals | | | |
| *C. elegans* | 100 | 19,427 | 25% |
| *D. melanogaster* | 123 | 13,379 | 13% |
| Pufferfish | 342 | 27,918 | 10% |
| Chicken | 1,130 | 25,000 | 3% |
| Human | 3,300 | 24,000 | 1.2% |

Mb = millions of base pairs

# Learning About "Life" By Peering Into Whole Genomes

## TABLE 17.2

### Comparison of the Genomes of *E. coli* and Yeast

| | *E. COLI* | YEAST |
|---|---|---|
| Genome length (base pairs) | 4,640,000 | 12,068,000 |
| Number of protein-coding genes | 4,290 | 5,770 |
| Proteins with roles in: | | |
| Metabolism | 650 | 650 |
| Energy production/storage | 240 | 175 |
| Membrane transport | 280 | 250 |
| DNA replication/repair/ recombination | 120 | 175 |
| Transcription | 230 | 400 |
| Translation | 180 | 350 |
| Protein targeting/secretion | 35 | 430 |
| Cell structure | 180 | 250 |

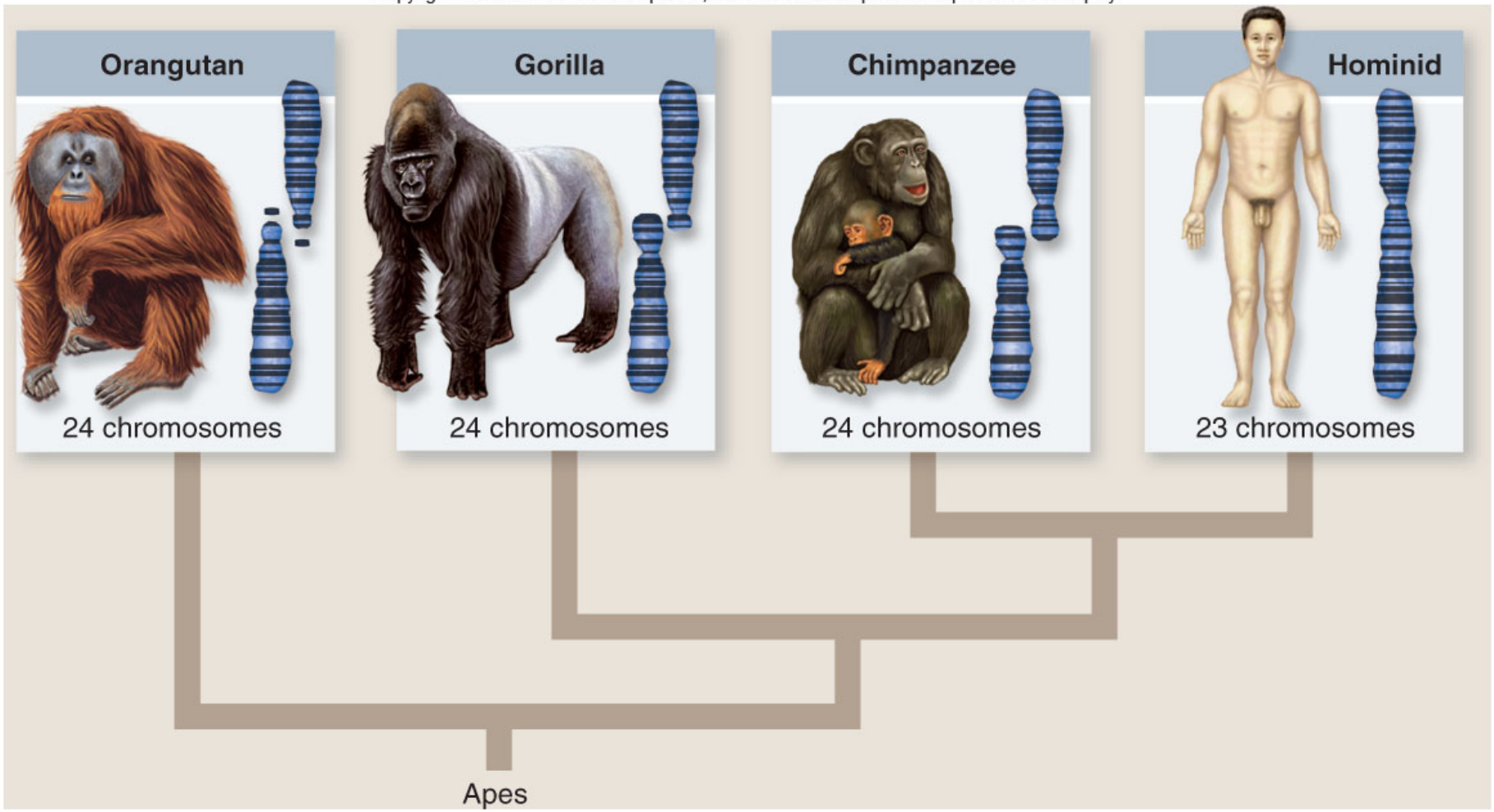
## TABLE 17.4

### *Arabidopsis* Genes Unique to Plants

| FUNCTION | NUMBER OF GENES |
|---|---|
| Cell wall and growth | 42 |
| Water channels | 300 |
| Photosynthesis | 139 |
| Defense and metabolism | 94 |

# Comparison of Mammalian Genomes Attempts To Determine "What Makes a Man, a Man and a Mouse a Mouse"



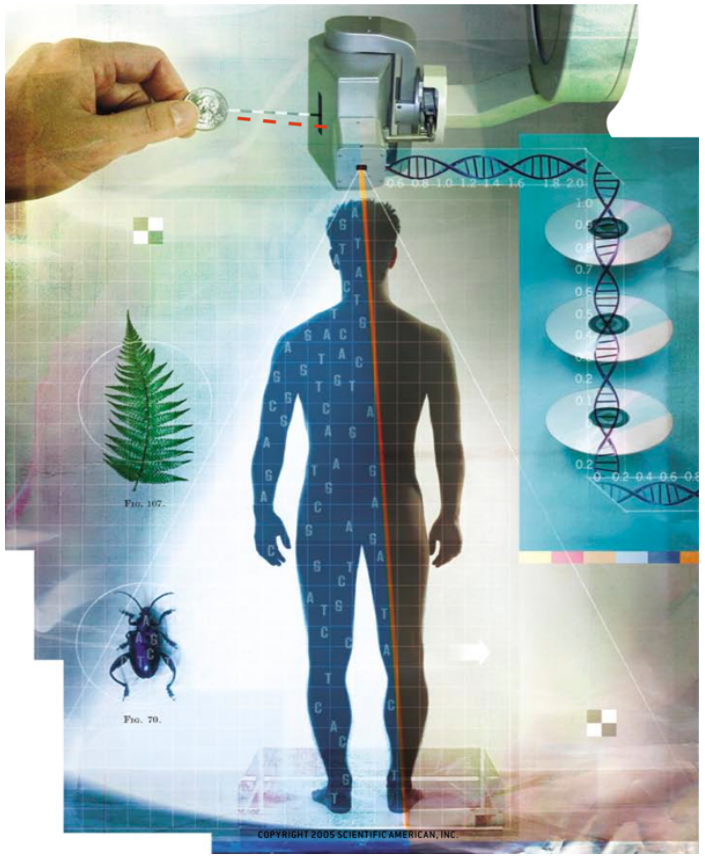



There is <1% Difference Between Human & Chimpanzee DNAs!

# The Ultimate Measure of Individuality Personal Genome Sequence & Comparing Individual Human Genomes

## Genomes for ALL

*Next-generation technologies that make reading DNA fast, cheap and widely accessible are coming in less than a decade. Their potential to revolutionize research and bring about the era of truly personalized medicine means the time to start preparing is now*

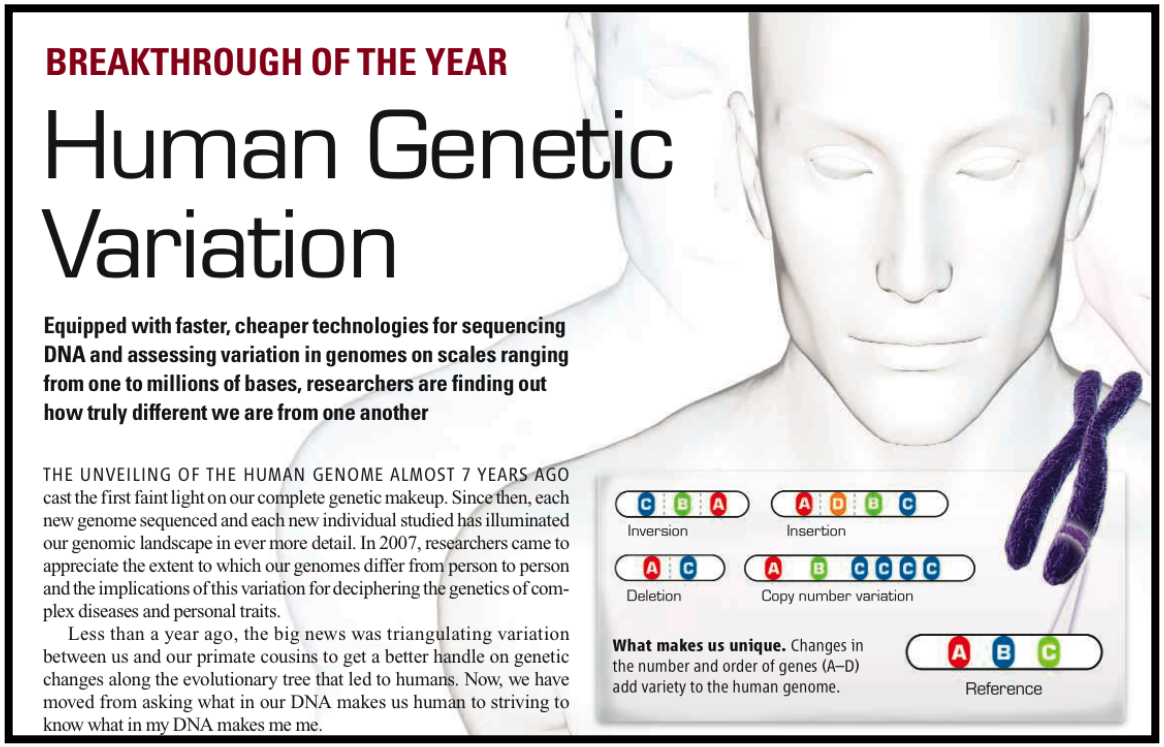



**Find DNA Variability in <u>All</u> Genes & Associate with Specific Traits!**

**Ultimately-You <u>Are</u> What Is In Your Genome**

# Identifying DNA Variations Between Individuals Many Uses

1. Marking and Identifying Disease Genes
2. Paternity, Individual Identification, Forensics
3. Human Population History and Origins



BREAKTHROUGH OF THE YEAR

## Human Genetic Variation

**Equipped with faster, cheaper technologies for sequencing DNA and assessing variation in genomes on scales ranging from one to millions of bases, researchers are finding out how truly different we are from one another**

THE UNVEILING OF THE HUMAN GENOME ALMOST 7 YEARS AGO cast the first faint light on our complete genetic makeup. Since then, each new genome sequenced and each new individual studied has illuminated our genomic landscape in ever more detail. In 2007, researchers came to appreciate the extent to which our genomes differ from person to person and the implications of this variation for deciphering the genetics of complex diseases and personal traits.

Less than a year ago, the big news was triangulating variation between us and our primate cousins to get a better handle on genetic changes along the evolutionary tree that led to humans. Now, we have moved from asking what in our DNA makes us human to striving to know what in my DNA makes me me.



**What makes us unique.** Changes in the number and order of genes (A–D) add variety to the human genome.

# Your Complete Genome Can Now Be Decoded and Sequenced Very Inexpensively ($1,000)!!

## Genome of DNA Pioneer Is Deciphered

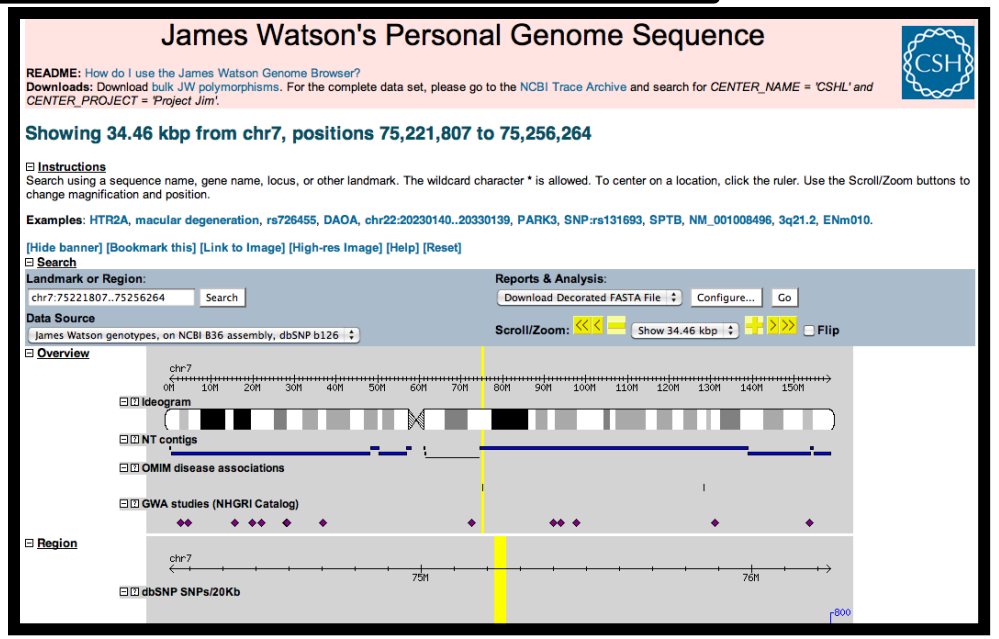
By NICHOLAS WADE
Published: May 31, 2007

## A map of human genome variation from population-scale sequencing

~200 Individual Genomes

The 1000 Genomes Project Consortium*

Nature, October 28, 2010


## *The Era of Personalized Genomes is Here!*

RESEARCH ARTICLE

PRENATAL DIAGNOSIS

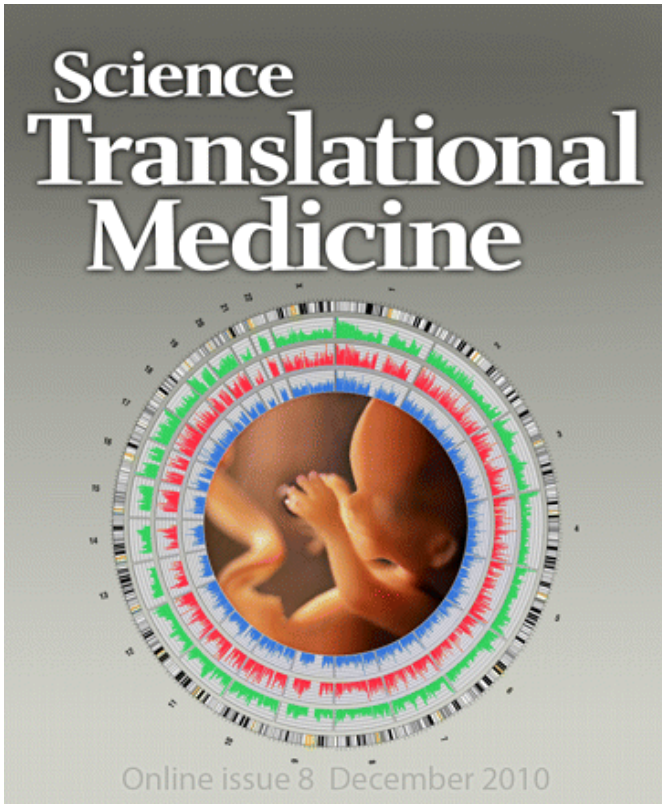
# Maternal Plasma DNA Sequencing Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus

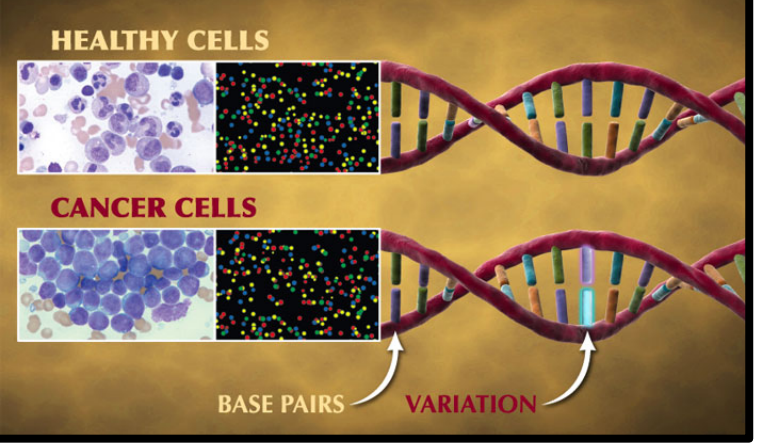
Science Translational Medicine, December 8, 2010 (61,1-12)

## Sequencing DNA From the Blood of a Pregnant Woman Allows the Complete Genome Of the Fetus to Be Decoded!

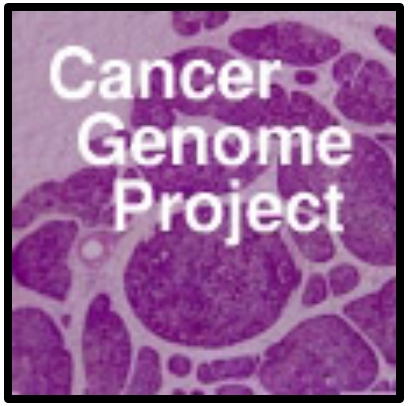
*~10% of DNA in Maternal Plasma is From the Fetus*

## A New Non-Invasive Era in DNA Testing!!










# Mapping Cancer Genes


# The 1,000 Genomes Project Will Provide Novel Insight in Human Genomes, Ancestry, & Disease Genes

Only Possible Using New Sequencing Methods

## 1,000 Genomes

Gene-sequencing projects keep getting bigger.
Tuesday, January 22, 2008
By Emily Singer

In a testament to the steady plummet in sequencing costs, today the National Human Genome Research Institute (NHGRI) announced a massive international collaboration to sequence the genomes of 1,000 people from around the world.

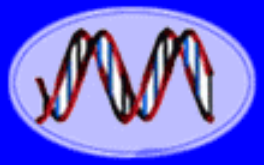
According to the NHGRI statement,

> "The 1000 Genomes Project will examine the human genome at a level of detail that no one has done before," said Richard Durbin, Ph.D., of the Wellcome Trust Sanger Institute, who is co-chair of the consortium. "Such a project would have been unthinkable only two years ago. Today, thanks to amazing strides in sequencing technology, bioinformatics and population genomics, it is now within our grasp. So we are moving forward to build a tool that will greatly expand and further accelerate efforts to find more of the genetic factors involved in human health and disease."
>
> During its two-year production phase, the 1000 Genomes Project will deliver sequence data at an average rate of about 8.2 billion bases per day, the equivalent of more than two human genomes every 24 hours. The volume of data--and the interpretation of those data--will pose a major challenge for leading experts in the fields of bioinformatics and statistical genetics.

2 Human Genomes Every 24 hrs !

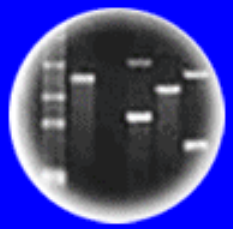
The 1,000 volunteers will be selected from those who participated in the HapMap project, a map of common genetic variation (see "A New Map for Health"), and will include:

Yoruba in Ibadan, Nigeria; Japanese in Tokyo; Chinese in Beijing; Utah residents with ancestry from northern and western Europe; Luhya in Webuye, Kenya; Maasai in Kinyawa, Kenya; Toscani in Italy; Gujarati Indians in Houston; Chinese in metropolitan Denver; people of Mexican ancestry in Los Angeles; and people of African ancestry in the southwestern United States.



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

ARTICLE

Nature, October 28, 2010

doi:10.1038/nature09534

# A map of human genome variation from population-scale sequencing

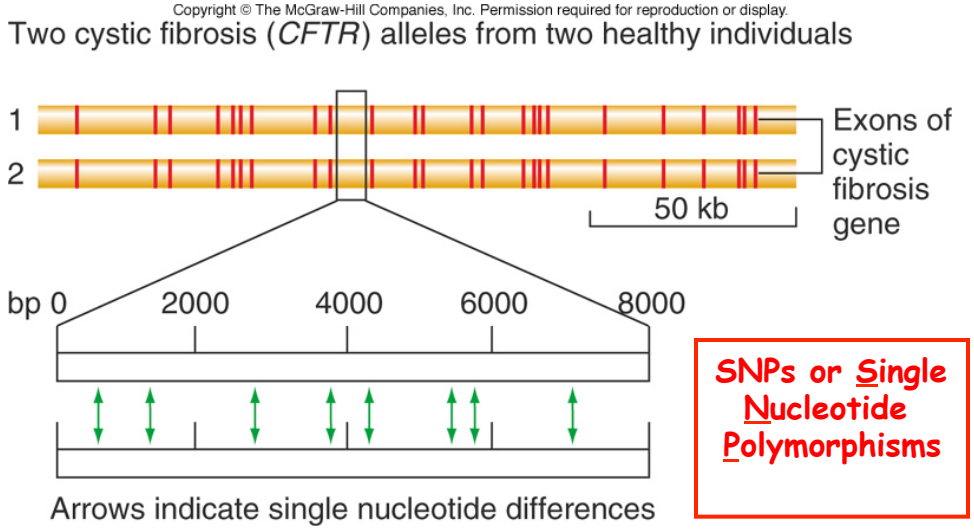
The 1000 Genomes Project Consortium*

The 1000 Genomes Project aims to provide a deep characterization of human genome sequence variation as a foundation for investigating the relationship between genotype and phenotype. Here we present results of the pilot phase of the project, designed to develop and compare different strategies for genome-wide sequencing with high-throughput platforms. We undertook three projects: low-coverage whole-genome sequencing of 179 individuals from four populations; high-coverage sequencing of two mother–father–child trios; and exon-targeted sequencing of 697 individuals from seven populations. We describe the location, allele frequency and local haplotype structure of approximately 15 million single nucleotide polymorphisms, 1 million short insertions and deletions, and 20,000 structural variants, most of which were previously undescribed. We show that, because we have catalogued the vast majority of common variation, over 95% of the currently accessible variants found in any individual are present in this data set. On average, each person is found to carry approximately 250 to 300 loss-of-function variants in annotated genes and 50 to 100 variants previously implicated in inherited disorders. We demonstrate how these results can be used to inform association and functional studies. From the two trios, we directly estimate the rate of *de novo* germline base substitution mutations to be approximately $10^{-8}$ per base pair per generation. We explore the data with regard to signatures of natural selection, and identify a marked reduction of genetic variation in the neighbourhood of genes, due to selection at linked sites. These methods and public data will support the next phase of human genetic research.

- Sequenced Genomes of ~900 individuals
- From Seven Different Global Populations
- Identified 15,000,000 SNPs
- 50-100 Variants in Disease Genes Per Person
- $10^{-8}$ Mutations Per bp Per Generation (~30 per Genome)
- 3,000,000 Unique SNPs Per Person
- 750,000 Unique Indels Per Person

# Most DNA Variations Between Individuals Occur Because of Base-Pair Changes in Non-Coding Regions of the Genome






**SNPs or Single Nucleotide Polymorphisms**

**To Be on the safe side, suppose you assume that only 80% (0.8) of the 3 billion base pairs in the genome are noncoding, and on average only 1 base pair in 700 is polymorphic. With these assumptions, you can determine the frequency of polymorphism within a single individual by multiplying 3 billion by 0.8 and then multiplying that amount by 1/700:**

**(3x109) x 0.8 = 2.4 x 109, (2.4 x 109) x 1/700 = 3.4 million.**

**The result of 3.4 million is astonishing: It means that there are millions of differences between any two haploid sets of human chromosomes. Combined with differences in coding and regulatory sequences (which occur much less frequently), the millions of polymorphisms at anonymous loci contribute to an enormous pool of potential DNA markers.**

## Types of DNA Polymorphisms

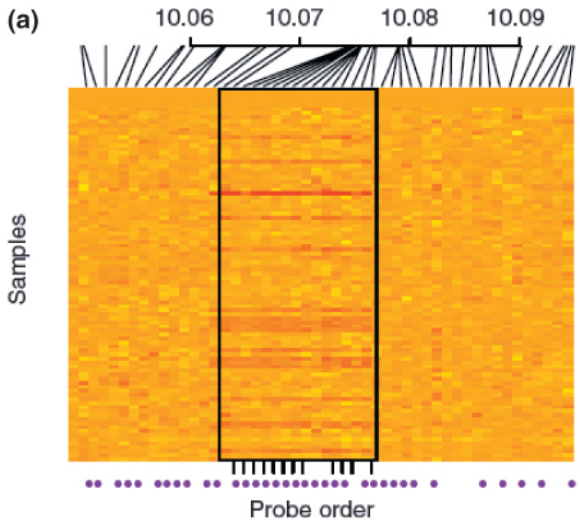
**TABLE 11.1 Classes of DNA Polymorphisms**

| Class | Size of Locus | Number of Alleles | Number of Loci in Population | Rate of Mutation | Use | Method of Detection |
|---|---|---|---|---|---|---|
| *SNP* | Single base pair | 2 | 100 million | $10^{-9}$ | Linkage and association mapping | PCR followed by ASO hybridization or primer extension |
| *Microsatellite* | 30–300 bp | 2–10 | 200,000 | $10^{-3}$ | Linkage and association mapping | PCR and gel electrophoresis |
| *Multilocus minisatellite* | 1–20 kb | 2–10 | 30,000 | $10^{-3}$ | DNA fingerprinting | Southern blot and hybridization |
| *Indels (deletions and duplications)* | 1–100 bp | 2 | N/A | $<10^{-9}$ | Linkage and association mapping | PCR and gel electrophoresis |




**This is What Makes Us Unique Individuals!**

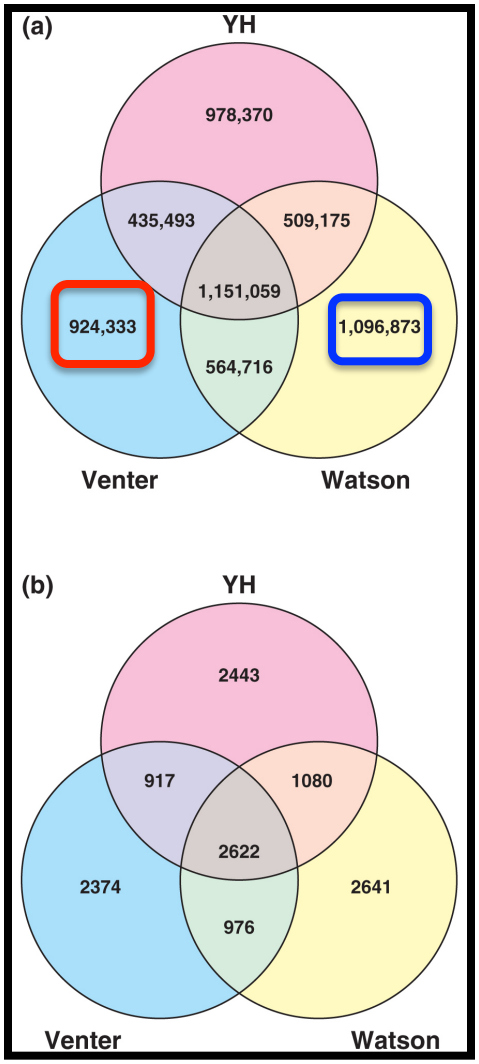
**There is ~1bp Change per 700bp in Human Genomes or ~3.4 Million bp Differences Between Individuals ~0.1% of Genome**




**Copy Number Variants Also Occur in the Human Genome and Can Vary From Individual to Individual**

**And Can Be Correlated With Diseases (Cause or Effect?)**

# Comparison of Watson's and Ventner's Genomes Indicates That Individuals Have Unique SNPs!




YH= Anonymous Chinese Man



(a)

| Allele | Genome | Single-Strand Notation |
|---|---|---|
| 1 | AGCTCTGATAC<br>TCGAGACTATG | AGCTCTGATAC |
| 2 | AGCTCCGATAC<br>TCGAGGCTATG | AGCTCCGATAC |

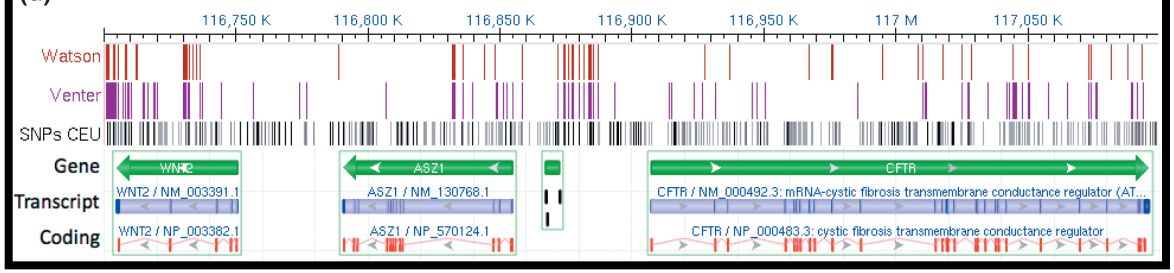
(b)

| Genotype | Notation |
|---|---|
| Homozygous T | TT |
| Heterozygous T | TC |
| Homozygous C | CC |

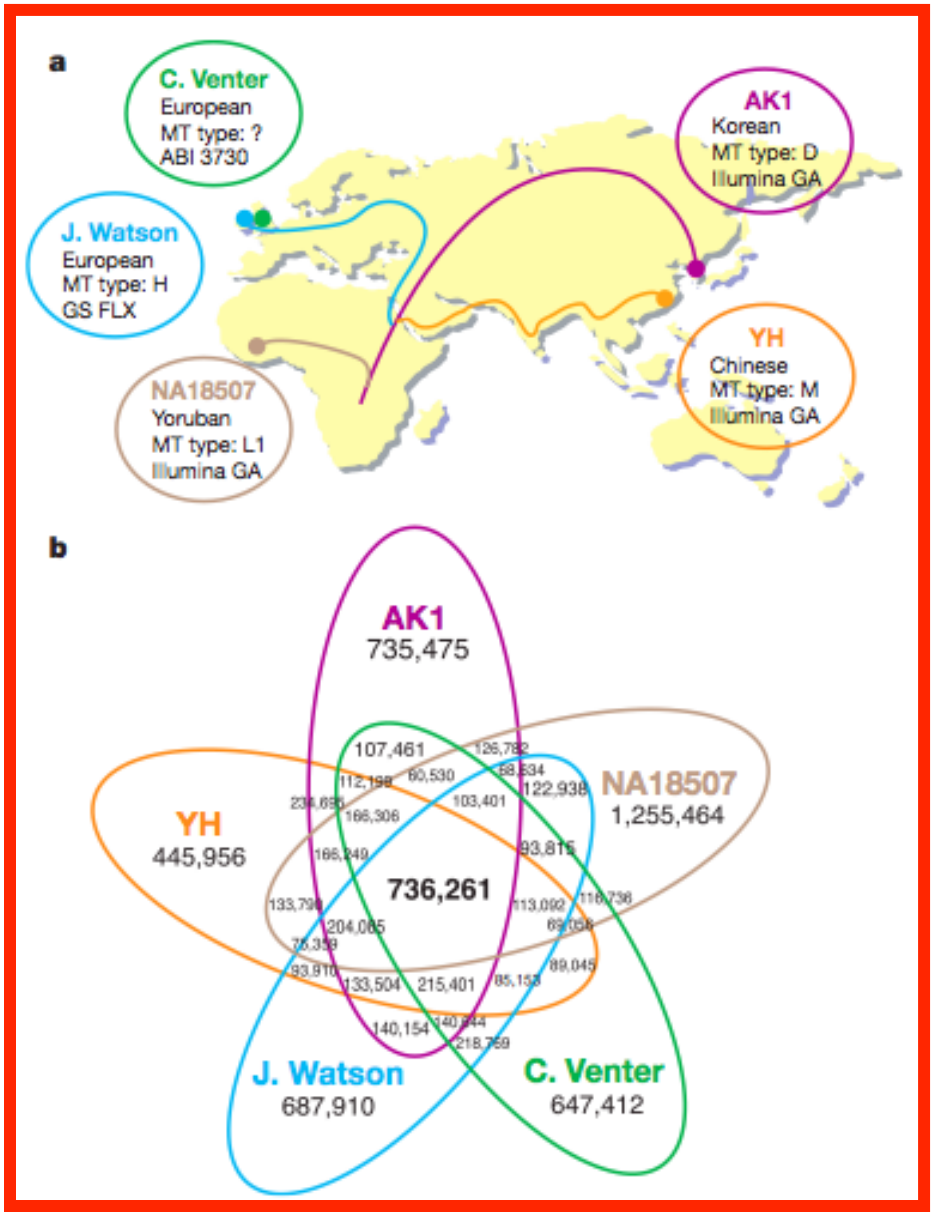
(c)

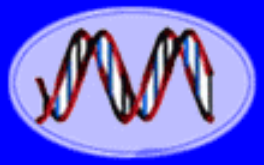
Human 1 TTGACGTATAAATGATCTTTATATTTTCAGAAGTC
Human 2 TTGACGTATAAATGATCTTTATATCTTCAGAAGTC
Chimp TTGACATATAAATGATCTTTATATCTTCAGAAGTC

(d)



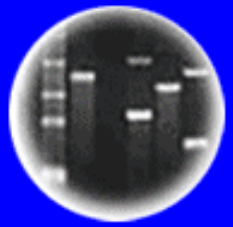
# In Fact, Everyone Has a Large Number of Unique SNPs!





DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria



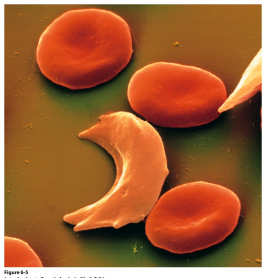
DNA Fingerprinting

Cloning: Ethical Issues
and Future Consequences

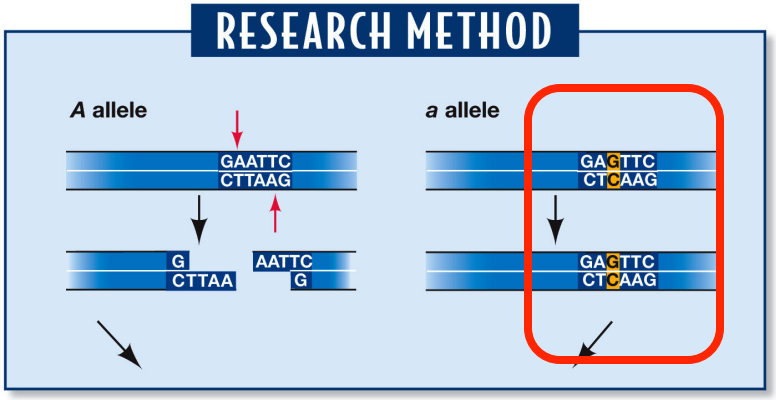
Plants of Tomorrow

# Using SNPs or DNA Sequence Variation As Markers For Disease Genes

**Remember**: Only a Small Fraction of Human Genes Are Known To Cause Diseases

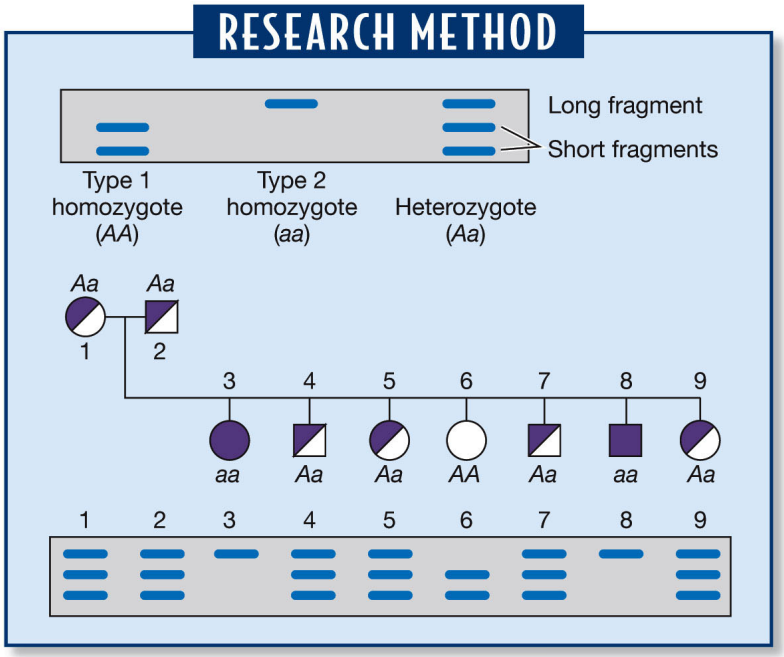


# Using RFLPs + Markers to Identify the Sickle Cell Allele (Single Gene Test)



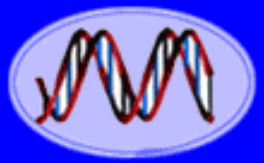

Loss of Restriction Site in a Allele (in gene)

Detected By Blots Or PCR



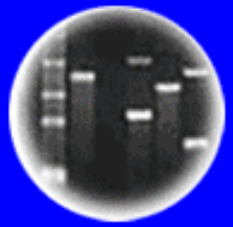

SNP Leads to RFLP!!!

*LIFE 8e*, Figure 17.8 (Part 2)



DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



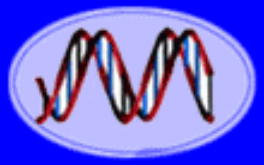
DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

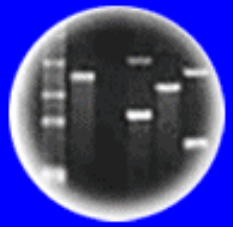
## Should DNA Testing Should Be Carried Out On Every Individual Born in the US?

a. Yes
b. No



DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria



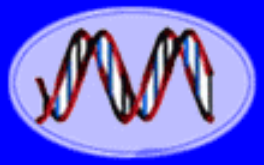
DNA Fingerprinting

Cloning: Ethical Issues
and Future Consequences

Plants of Tomorrow

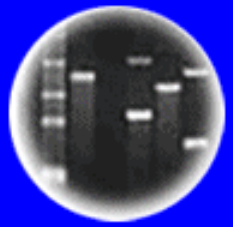
## Are DNA Tests e Carried Out On Every New Baby Born in the US?

a. Yes
b. No



DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria



DNA Fingerprinting

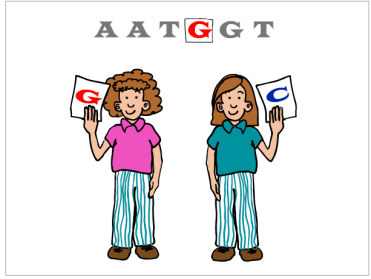
Cloning: Ethical Issues
and Future Consequences

Plants of Tomorrow

## Should Prenatal DNA Testing Be Covered By Insurance?

a. Yes
b. No

# SNPs Can Be Associated/Linked With Specific Physical Traits




## OCA2

**From SNPedia**

**OCA2**, the oculocutaneous albinism gene (also known as the human P protein gene, or, DN10), is a gene associated with albinism and certain pigmentation effects in general such as eye color, skin color, and hair color.

A large (>3,000 individuals) study of Caucasians indicates that the following **OCA2** variants, all located in the first intron of the gene, are preferentially linked to blue eye color inheritance; together, they form haplotypes that (in some cases at least) predict eye color with greater than 50:50 odds. [PMID 17236130; OMIM 203200.0013 (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=203200&a=203200_AllelicVariant0013) ]

- rs7495174
- rs6497268
- rs11855019

The haplotypes are defined in order as listed above for these 3 SNPs, so, for example, the TGT haplotype refers to rs745174(T)-rs6497268(G)-rs11855019(T). The correspondence between diplotypes (the two haplotypes in one individual) and the % of individuals with blue/gray, green/hazel/ and brown eye color , respectively, was reported as follows for the most common diplotypes[PMID 17236130]:

- TGT/TGT: 62.5, 28.0, 9.5
- TGT/***TTC***: 47.1, 20.3, 32.6
- TGT/***CGT***: 28.6, 14.3, 57.1
- TGT/***TGC***: 27.9, 22.1, 50.0
- ***TGC***/***TTC***: 25.0, 8.3, 66.7
- ***TTT***/***TGC***: 20.7, 31.0, 48.3
- TGT/***TTT***: 17.6, 38.5, 44.0
- TGT/***CTC***: 7.9, 23.3, 68.8

The haplotypes shown in ***bold italics*** represent the ones reported by the authors of this study to be most associated with brown eye color. Furthermore, the haplotypes shown above are as published, and the associated SNPs - which have since changed # as well - are not in the orientation shown in dbSNP.

More recently, a study of a large Danish family led to associations with 2 SNPs in a different region of **OCA2** as linked to blue or brown eye color:

- rs12913832
- rs1129038

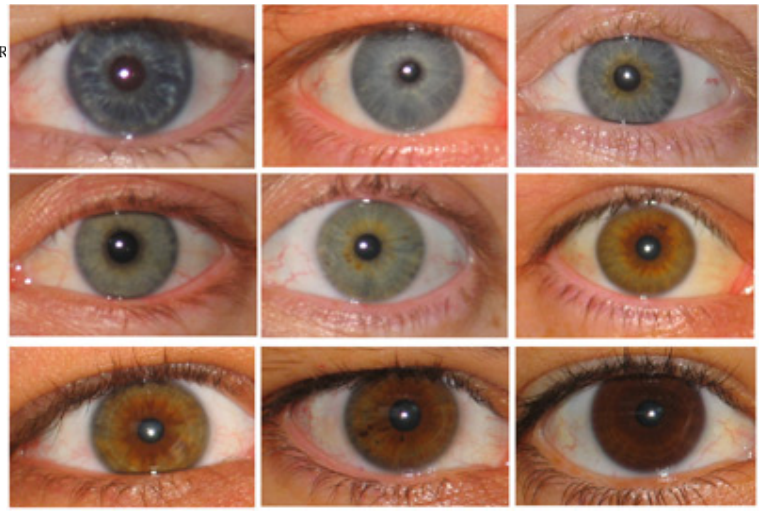
Earlier studies found different regions of the **OCA2** gene to also be predictive of eye color;

- **OCA2** SNP rs1800401 helps predict brown eye color. [PMID 12163334, PMID 15889046; OMIM 203200.0011 (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=203200&a=203200_AllelicVariant0011) ]
- **OCA2** SNP rs1800407 may be associated with green/hazel eye color in some populations, but not others. [PMID 12163334, PMID 15889046; OMIM 203200.0012 (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=203200&a=203200_AllelicVariant0012) ]

| | |
|---|---|
| **is a** | gene |
| **is** | mentioned by |
| **wikipedia** | OCA2 (http://en.wikipedia.org/wiki/OCA2) |
| **google** | OCA2 (http://www.google.com/search?hl=en&q=OCA2) |
| **gopubmed** | OCA2 (http://www.gopubmed.org/search?q=OCA2) |
| **23andMe** | OCA2 (https://www.23andme.com/you/explorer/gene/?gene_name=OCA2) |
| **GeneRIF** | 4948 (http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=4948&ordinalpos=1&itool=EntrezSystem2.PEntrez.Gene.Gene_ResultsPanel.Gene_R |
| **dbSNP** | 4948 (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=4948&chooseRs=all) |
| **PubMed** | 4948 (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=Link&LinkName=gene_pubmed&from_uid=4948) |
| **HugeNav** | 4948 (http://hugenavigator.net/HuGENavigator/huGEPedia.do?firstQuery=OCA2}&geneID=4948&typeSubmit=GO&check=y&typeOption=gene&which=2&pubOrderType=pubD) |

| | **Chromosome position** |
|---|---|
| Rs1129038 | 26,030,454 |
| Rs11631797 | 26,175,874 |
| Rs12593929 | 26,032,853 |
| Rs1800401 | 25,933,648 |
| Rs1800407 | 25,903,913 |
| Rs2238289 | 26,126,810 |
| Rs2240203 | 26,167,797 |
| Rs28934272 | 25,903,842 |
| Rs3935591 | 26,047,607 |
| Rs3940272 | 26,142,318 |
| Rs4778241 | 26,012,308 |
| Rs7170852 | 26,101,581 |
| Rs7183877 | 26,039,328 |
| Rs7495174 | 26,017,833 |
| Rs8028689 | 26,162,483 |
| Rs916977 | 26,186,959 |

## SNPs in Human P Protein (OCA2) Gene Lead To Different Eye Colors (Physical & Molecular Markers)



Human Eye Color

# Constructing Portraits From DNA

Research & Discovery

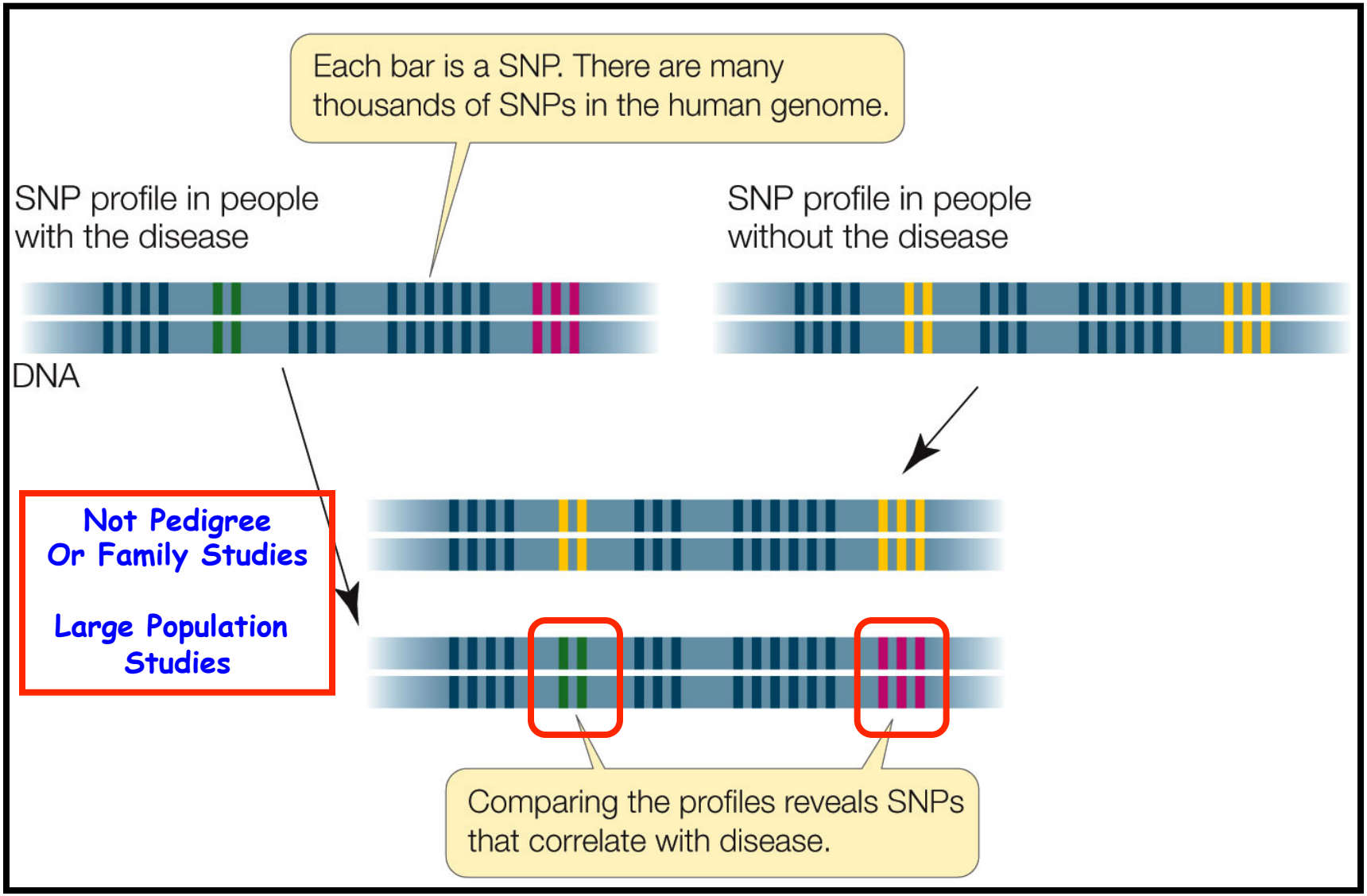
## Portrait in DNA

**Can forensic analysis yield police-style sketches of suspects?** BY CHRISTINE SOARES

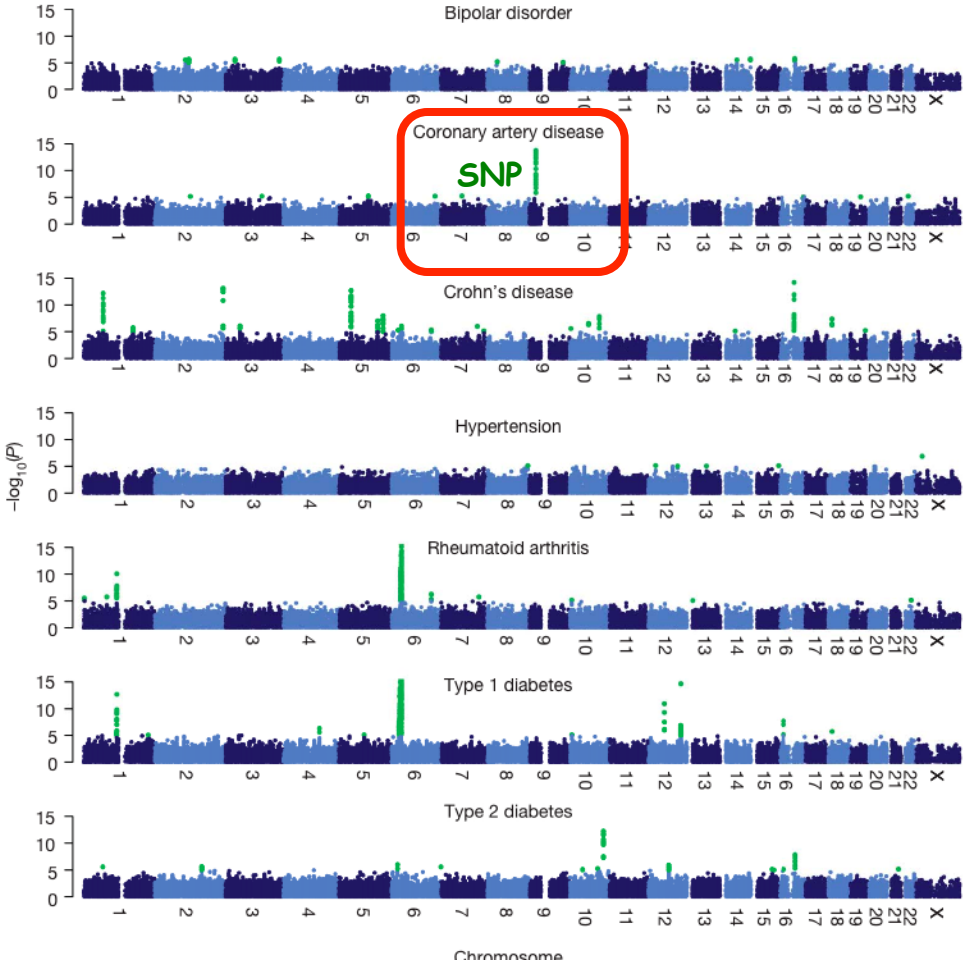
MALE, SHORT AND STOUT, WITH DARK SKIN, BROWN EYES, shovel-shaped teeth, type A+ blood and coarse, dark brown hair giving way to pattern baldness. He would have a high tolerance for alcohol and a higher-than-average risk of nicotine dependence—fortunately, he lived thousands of years before humans discovered smoking. The description of a Stone Age Greenland resident published in February paints an extraordinary portrait of a man who vanished more than 4,000 years ago, drawn almost solely from his DNA remains.

RECONSTRUCTED: Ancient DNA provided details about the looks of a man who lived in Greenland more than 4,000 years ago.

# Using SNPs and Population Association Studies to Find Disease Markers and Genes



# Correlating SNPs With Specific Diseases Using SNP Chips & Association Studies




Note: Probability

SNPs May Be Near Or In Relevant Genes

# TABLE 17.5

## SNP Human Genome Scans and Diseases

| DISEASE | LOCATION OF SNP (CHROMOSOME NUMBER) | % INCREASED RISK | |
|---|---|---|---|
| | | HETEROZYGOTES | HOMOZYGOTES |
| Breast cancer | 8 | 20 | 63 |
| Coronary heart disease | 9 | 20 | 56 |
| Heart attack | 9 | 25 | 64 |
| Obesity | 16 | 32 | 67 |
| Diabetes | 10 | 65 | 277 |
| Prostate cancer | 8 | 26 | 58 |

# Using Large Populations SNPs Can Be Used As Markers For Specific Genes/Traits

## SNPedia (http://www.snpedia.com/)

- New model for prostate cancer based on 5 SNPs
- rs1815739 sprinters vs endurance athletes
- rs4420638 and rs429358 can raise the risk of Alzheimer's disease by more than 10x
- rs6152 can prevent baldness
- rs9939609 triggers obesity
- rs662799 prevents weight gain from high fat diets
- rs7495174 green eye color
- rs7903146 in 3% of the population greatly increases the risk of type-2 diabetes
- rs12255372 linked to type-2 diabetes and breast cancer
- rs2395029 asymptomatic HIV viral load set point
- rs324650 influences intelligence and alcohol dependence
- rs1799971 makes alcohol cravings stronger
- rs17822931 determines earwax

Caution

How Will You Use the Information?
How Good Are The Correlations?
What To Do With The Information?
Privacy Issues?
Group Differences? Discrimination?

- rs53576 in the oxytocin receptor influences social behavior and personality
- rs1815739 muscle performance
- rs7412 and rs429358 can raise the risk of Alzheimer's disease by more than 10x
- rs6152 can influence baldness
- rs333 resistance to HIV
- rs1800497 in a dopamine receptor may influence the sense of pleasure
- rs1805007 determines red hair and sensitivity to anesthetics
- rs9939609 triggers obesity and type-2 diabetes
- rs662799 prevents weight gain from high fat diets
- rs7495174 green eye color and rs12913832 for blue eye color
- rs7903146 in 3% of the population greatly increases the risk of type-2 diabetes
- rs12255372 linked to type-2 diabetes and breast cancer
- rs1799971 makes alcohol cravings stronger
- rs17822931 determines earwax
- rs4680 varied cognitive effects
- rs1333049 coronary heart disease
- rs1801133 folate metabolism and several cancers
- rs1051730 and rs3750344 nicotine dependence
- rs4988235 lactose intolerance

# Examples of SNPs in SNPedia Database

# Examples of Whole Genomes in SNPedia Database

(http://www.snpedia.com/)

These are the 105 public genomes. They are from real people who've chosen to share their data to help all of us learn more about our genomes. but be sure to check the report header to understand how up to date each is.

| | Platform | Raw data available | Summary |
|---|---|---|---|
| Aaron Vollrath | 23andMe v2 | | Male on 23andMe |
| Almelina | 23andMe v2 | true | Female with cancer 2x |
| Bgreshake | 23andMe v3 | | German Male 23andMe v3 |
| Blainebettinger | 23andMe v2<br>FTDNA Family Finder | | 23andMe v. FTDNA Family Finder data |
| Corpas aunt | 23andMe v3 | | Visualized family snps; whole family 23andMe genotypes |
| Corpas dad | 23andMe v3 | | Visualized family snps; whole family 23andMe genotypes |
| Corpas mom | 23andMe v3 | | Visualized family snps; whole family 23andMe genotypes |
| Corpas sister | 23andMe v3 | | Visualized family snps; whole family 23andMe genotypes |
| Daniel MacArthur | 23andMe v3<br>23andMe v2<br>Lumigenix | | 1st Lumigenix, GNZ, 23andMe v2+3, blogger |
| David Ewing Duncan | Complete Genomics<br>23andMe v2<br>DeCODEme<br>Navigenics | | Full genome from Complete Genomics, but also microarrays from 23a |
| DeCODEme | DeCODEme | | Sample deCODEme male |
| Denisova | Full Sequencing | | a 41k year old member of the genus Homo |
| Dichro | 23andMe v2<br>23andMe v1 | true | 23andMe male (v1 + v2) |

# Using SNPs and Association Studies To Construct a Personal Genome Profile

## It's All About Me

Along with the flood of discoveries in human genetics, 2007 saw the birth of a new industry: personal genomics. Depending on your budget, you can either buy a rough scan of your genome or have the whole thing sequenced. The companies say the information will help customers learn about themselves and improve their health. But researchers worry that these services open up a Pandora's box of ethical issues.

At $300,000 to $1 million per genome, sequencing all 3 billion base pairs is still too costly for all but a few. Although dozens more personal genomes will probably be sequenced in the coming year, most will be done by public and private research organizations—including the institute run by genome maverick J. Craig Venter, whose personal genome was one of three completed in 2007 in the United States and China. In a lower-budget effort, Harvard's George Church this month will deliver initial DNA sequences for the protein-coding sections (1% of the genome) to the first 10 volunteers for his Personal Genome Project. Meanwhile, a new company called Knome is offering full-genome sequencing to 20 customers willing to pay $350,000.

A glimpse of one's genome is already within the reach of ordinary people, thanks to several companies. They include 23andMe, which has financing from Google and may let users link to others with shared traits; Navigenics, which will screen for about 20 medical conditions; and deCODE Genetics in Iceland, a pioneer in disease gene hunting. For $1000 to $2500, these companies will have consumers send in a saliva sample or cheek swab, then use "SNP chips" to scan their DNA for as many as 1 million markers. The companies will then match the results with the latest publications on traits, common diseases, and ancestry.

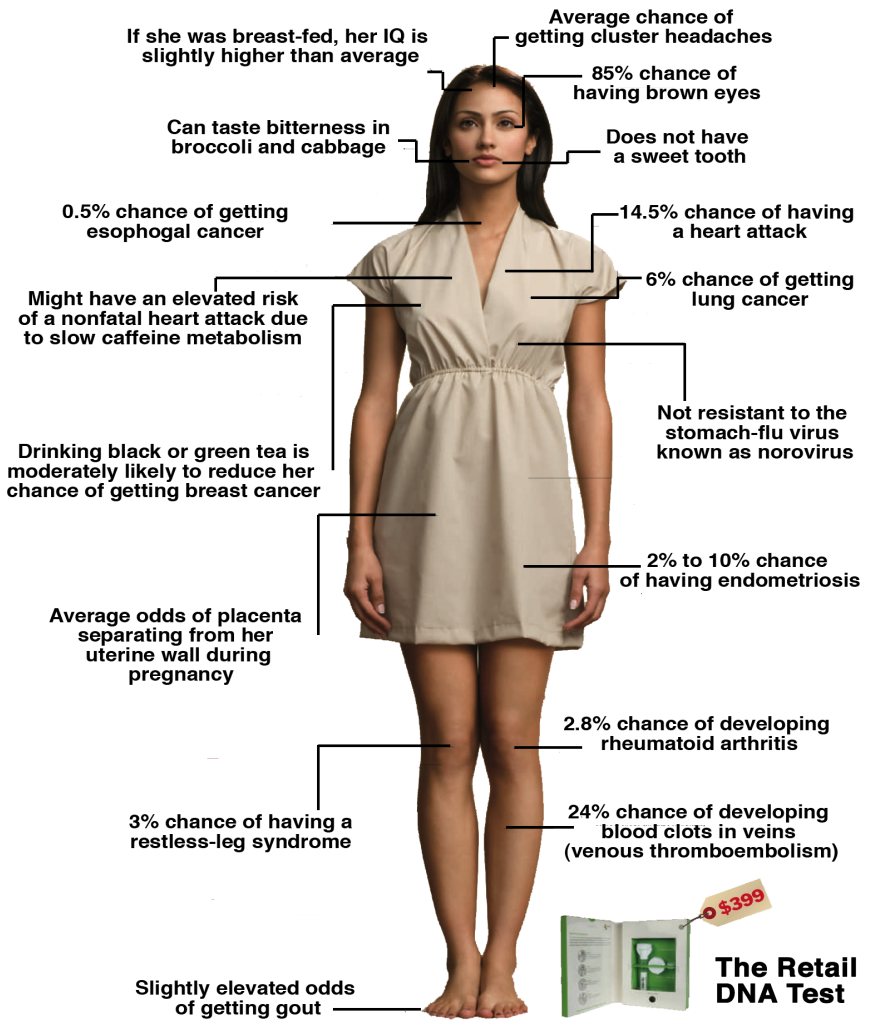
Although many customers may view this exercise as a way to learn fun facts about themselves—recreational genomics, some call it—bioethicists are wary. Most common disease markers identified so far raise risks only slightly, but they could cause needless worry. At the same time, some people may be terrified to learn they have a relatively high risk for an incurable disease such as Alzheimer's.

**Pandora's box?** This cheek-swab kit could reveal your intimate secrets.

The rush toward personal genome sequences also sharpens long-held worries about discrimination. A bill to prevent insurers and employers from misusing genetic data is stalled in Congress. Complicating matters, your genetic information exposes your relatives' DNA, too.

The most profound implications of having one's genome analyzed may not be what it reveals now—which isn't much—but what it may show later on. Perhaps to sidestep such questions, some companies will limit which markers to disclose. Others, however, will hand customers their entire genetic identity, along with all the secrets it may hold.

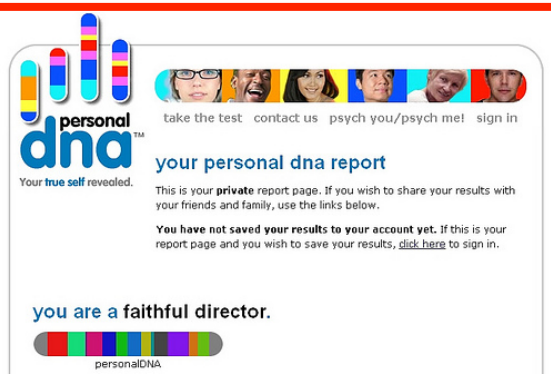
–JOCELYN KASER

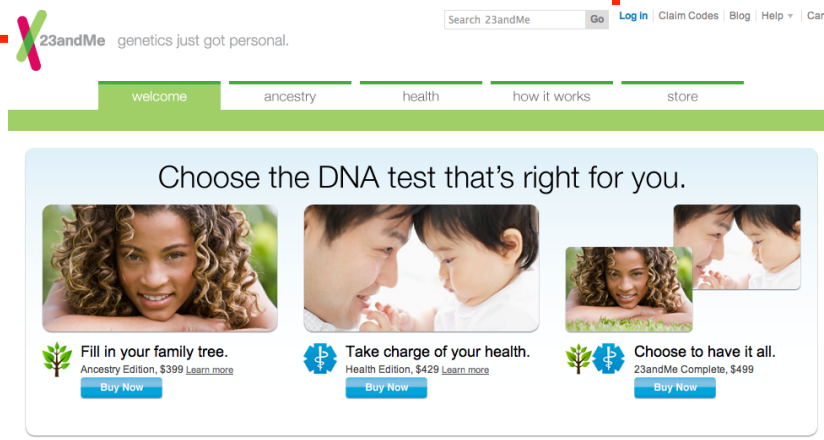


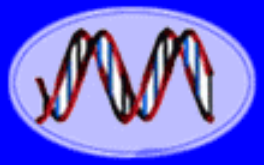


# The Problems With Human Genome Testing Companies Are?

a. Reliability of Results?
b. Validity of Association Studies
c. Privacy?
d. What To Do With Information Obtained?
e. FDA Regulatory Oversight?
f. Deceptive Practices
g. All of Above?

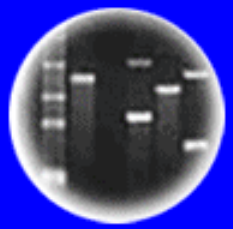


DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

# Problem: Different Companies-Different Predictions-No Oversight!

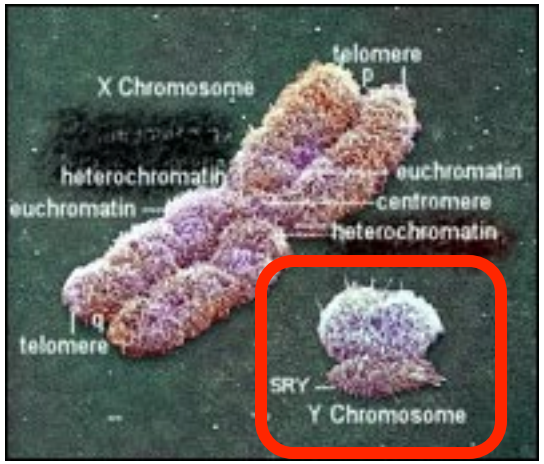
TABLE 1: PREDICTIONS FOR DISEASE RELATIVE RISKS FOR FIVE INDIVIDUALS

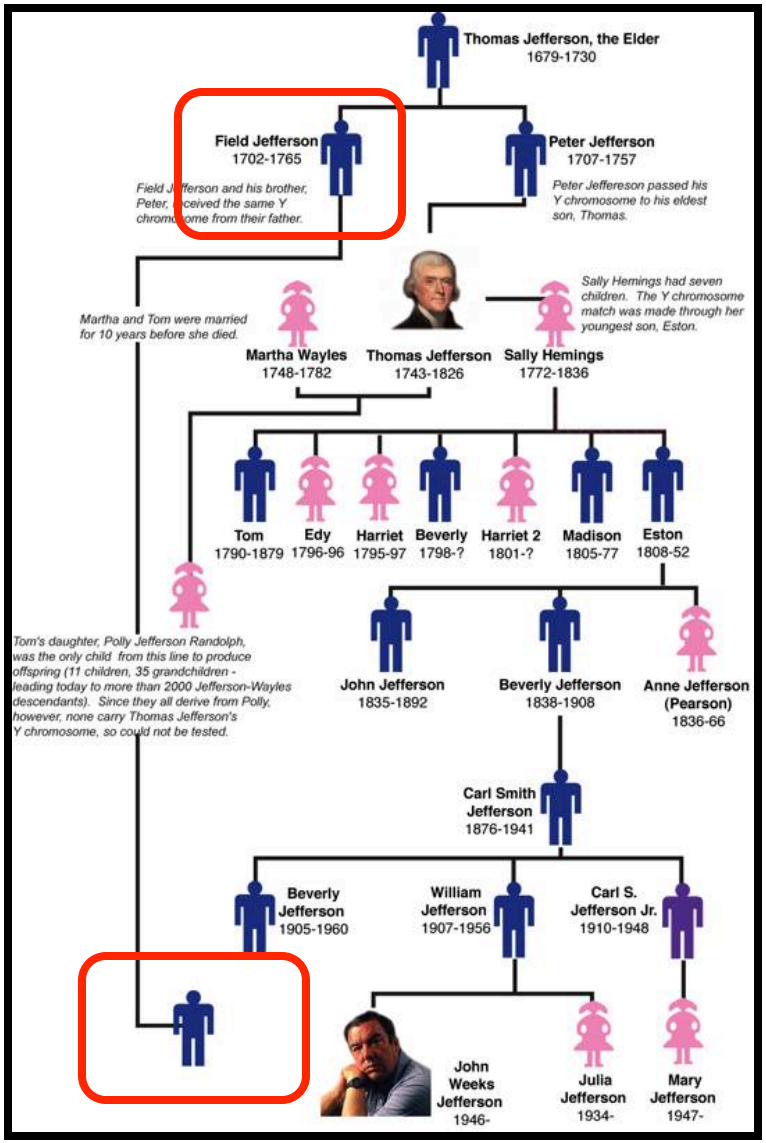
| Disease | Female A | Female B | Female C | Male D | Male E |
|---|---|---|---|---|---|
| Breast cancer | ↑↑ | ↑↑ | ↓↓ | | |
| Coeliac disease | ↓↓ | ↓↓ | ↓↓ | ↓↓ | ↓↓ |
| Colon cancer | == | == | =↓ | ↑↑ | =↓ |
| Crohn's disease | ↓↑ | ↓↑ | ↓↓ | ↓↓ | ↓= |
| Heart attack | ↓↓ | =↓ | =↓ | =↓ | ↑↑ |
| Lupus | ↑↓ | ↓↓ | ↓↓ | ↑= | ↑= |
| Macular degeneration | ↓↓ | ↓↓ | ↑= | ↓↓ | ↓↓ |
| Multiple sclerosis | ↑↑ | | ↓↓ | ↓↓ | ↓↓ |
| Prostate cancer | | | | ↑↑ | ↓↑ |
| Psoriasis | ↓↑ | | ↑↓ | ↑↑ | ↓↓ |
| Restless legs syndrome | =↓ | ↑↑ | ↓= | ↓↑ | ↑↑ |
| Rheumatoid arthritis | ↑↑ | ↑↑ | ↓↓ | ↓↓ | ↑↑ |
| Type 2 diabetes | ↓↓ | =↓ | ↓↓ | ↑↓ | =↓ |

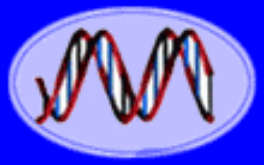
↑ increased risk (RR > 1.05), ↓ decreased risk (relative risk (RR) < 0.95), = average risk (0.95 ≤ RR ≤ 1.05). First prediction is from 23andMe; second prediction is from Navigenics. Different predictions are highlighted in beige.

# SNPs Can Be Used To Determine Paternity

## Using Y Chromosome SNPs and RFLPs To Determine That Thomas Jefferson and Sally Hemmings Had Children

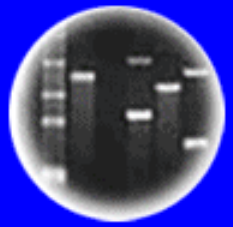




DNA
Genetic Code of Life

Entire Genetic Code of a Bacteria

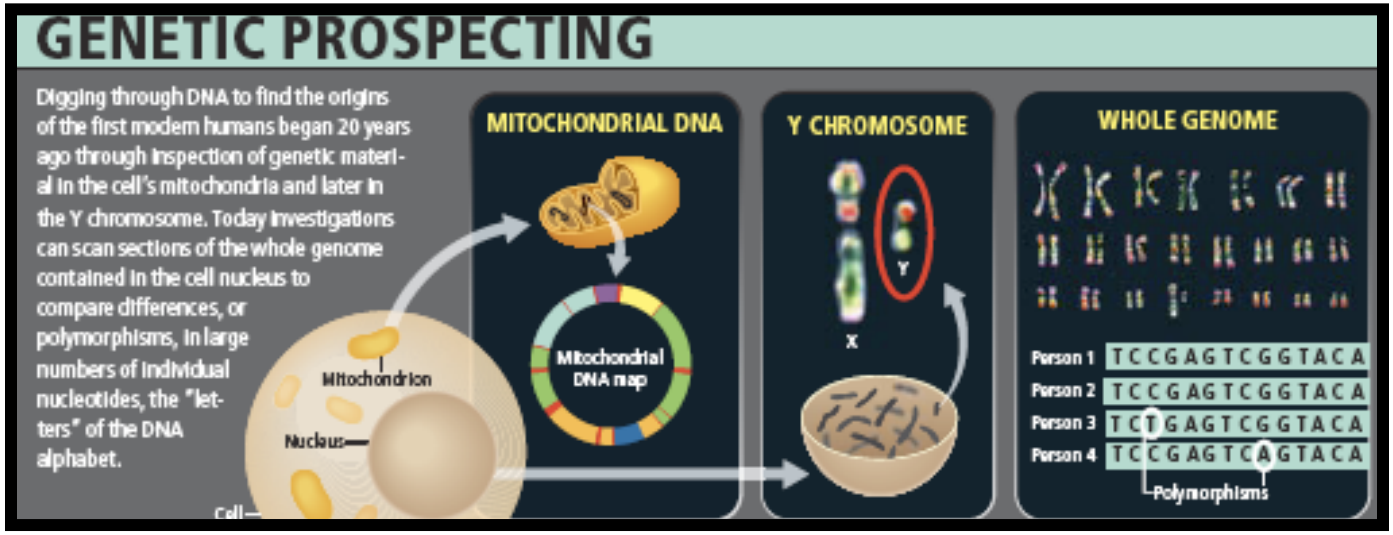


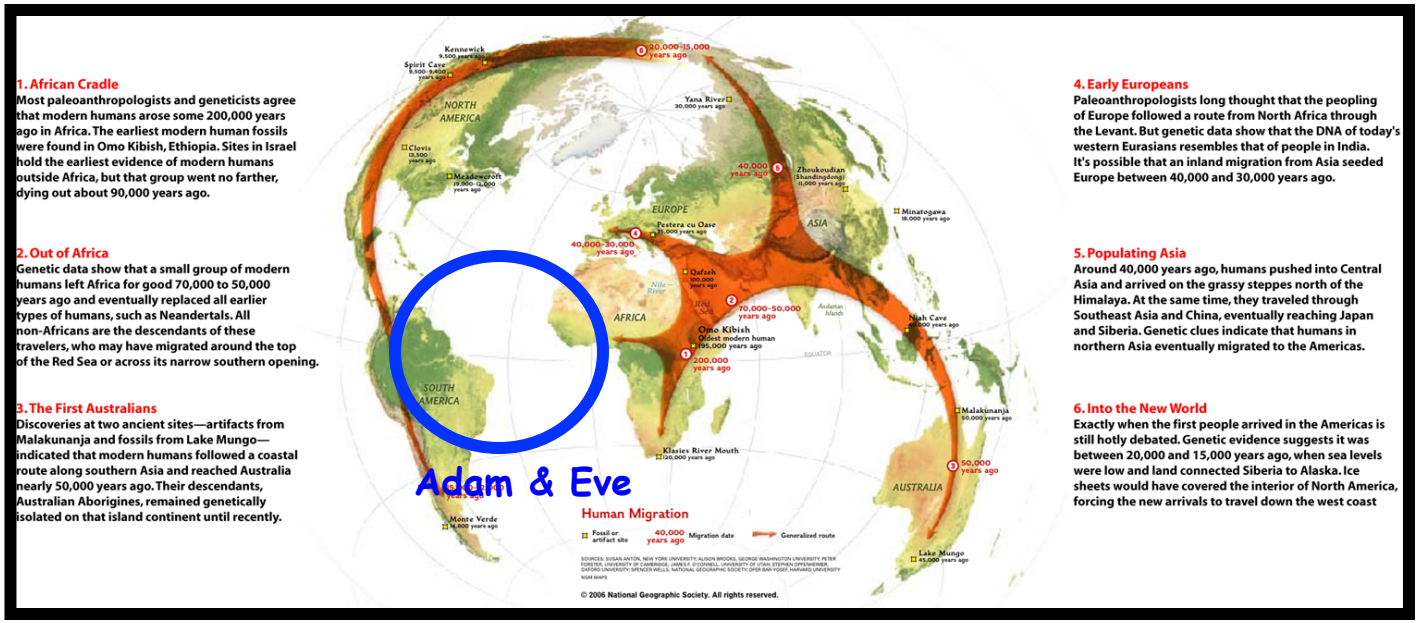
DNA Fingerprinting

Cloning: Ethical Issues and Future Consequences

Plants of Tomorrow

# SNPS Can Be Used To Trace Human Populations & Origins



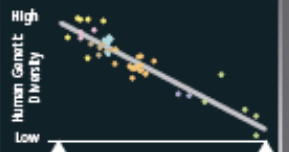
# Human Population Relationships and Origins Using Whole-Genome Comparisons

[WHOLE-GENOME RESULTS]

## LOOKING FAR AND WIDE

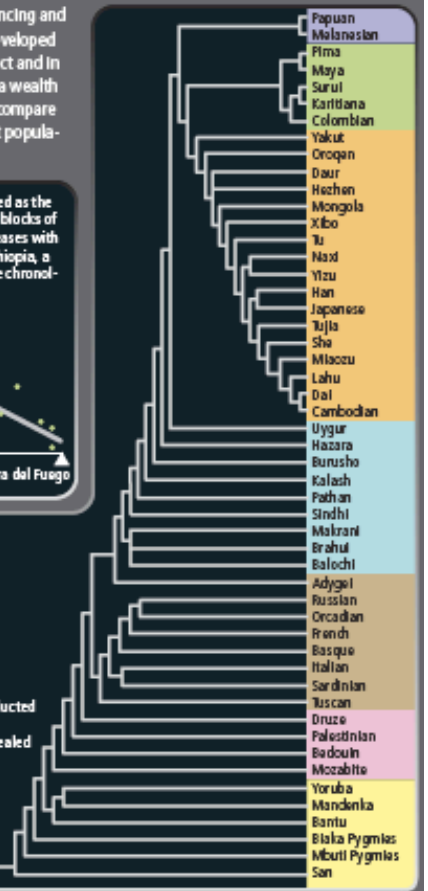
High-powered genetic sequencing and computational techniques developed for the Human Genome Project and in its aftermath have furnished a wealth of data that lets researchers compare genomes drawn from distinct populations around the globe.

The diversity of DNA—measured as the variation of nucleotides within blocks of DNA called haplotypes—decreases with distance from Addis Ababa, Ethiopia, a pattern that corresponds to the chronology of human migrations.

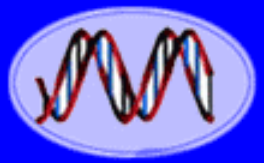



**Geographic Region**
- Oceania
- Americas
- East Asia
- Central and South Asia
- Europe
- Middle East
- Africa

Whole-genome analysis conducted by researchers at Stanford University and elsewhere revealed many of the populations that form the branches of a genetic tree beginning in Africa and expanding out to the rest of the world.

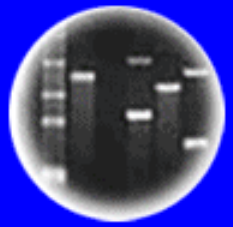



A genomic map of the world, crafted by researchers at the University of Michigan at Ann Arbor, shows that genetic diversity decreases outside of Africa. Each colored tile represents a common haplotype. Africa has more tiles than found on other continents and ones that correspond to haplotypes found nowhere else.

DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria



DNA Fingerprinting

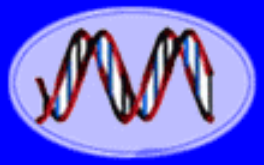
Cloning: Ethical Issues
and Future Consequences

Plants of Tomorrow

# HUMAN DIVERSITY

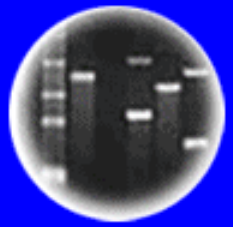
RICHARD LEWONTIN

*Scientific American Library*
*1982 ISBN 07167-14698*



DNA
Genetic Code of Life

Entire Genetic Code
of a Bacteria



DNA Fingerprinting

Cloning: Ethical Issues
and Future Consequences

Plants of Tomorrow

## Human Races Have a Genetic Basis:

a. Yes
b. No

# There is More Genetic Diversity Within Populations than Between Populations!! So Much for the Concept of Racial "Purity"!!!!

Proportion of genetic diversity accounted for within and between populations and races

| Gene | Total $H_{species}$ | Proportion | | |
|---|---|---|---|---|
| | | Within Populations | Within Races between Populations | Between Races |
| *Hp* | .994 | .893 | .051 | .056 |
| *Ag* | .994 | .834 | — | — |
| *Lp* | .639 | .939 | — | — |
| *Xm* | .869 | .997 | — | — |
| *Ap* | .989 | .927 | .062 | .011 |
| *6PGD* | .327 | .875 | .058 | .067 |
| *PGM* | .758 | .942 | .033 | .025 |
| *Ak* | .184 | .848 | .021 | .131 |
| *Kidd* | .977 | .741 | .211 | .048 |
| *Duffy* | .938 | .636 | .105 | .259 |
| *Lewis* | .994 | .966 | .032 | .002 |
| *Kell* | .189 | .901 | .073 | .026 |
| *Lutheran* | .153 | .694 | .214 | .092 |
| *P* | 1.000 | .949 | .029 | .022 |
| *MNS* | 1.746 | .911 | .041 | .048 |
| *Rh* | 1.900 | .674 | .073 | .253 |
| *ABO* | 1.241 | .907 | .063 | .030 |
| Mean | | .854 | .083 | .063 |

Source: R. C. Lewontin, *Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).

**More Genetic Diversity Within Any Population Than Between Polulations**

1. 85% of Human Genetic Variations Occurs within Populations & Between Individuals in that Populations!
2. Remaining 15% of Human Genetic Variation Split Between Different Populations of Same "race" (8%) & Between Different "races" (6%)
3. Only 6% of Human Genetic Variation are to Differences between races!!! Mostly Geographic. Note: THERE ARE GROUP DIFFERENCES!

# Within Population Differences Account For 95% of Human Genetic Variation

## Genetic Structure of Human Populations

Noah A. Rosenberg,[1]* Jonathan K. Pritchard,[2] James L. Weber,[3] Howard M. Cann,[4] Kenneth K. Kidd,[5] Lev A. Zhivotovsky,[6] Marcus W. Feldman[7]

We studied human population structure using genotypes at 377 autosomal microsatellite loci in 1056 individuals from 52 populations. Within-population differences among individuals account for 93 to 95% of genetic variation; differences among major groups constitute only 3 to 5%. Nevertheless, without using prior information about the origins of individuals, we identified six main genetic clusters, five of which correspond to major geographic regions, and subclusters that often correspond to individual populations. General agreement of genetic and predefined populations suggests that self-reported ancestry can facilitate assessments of epidemiological risks but does not obviate the need to use genetic information in genetic association studies.

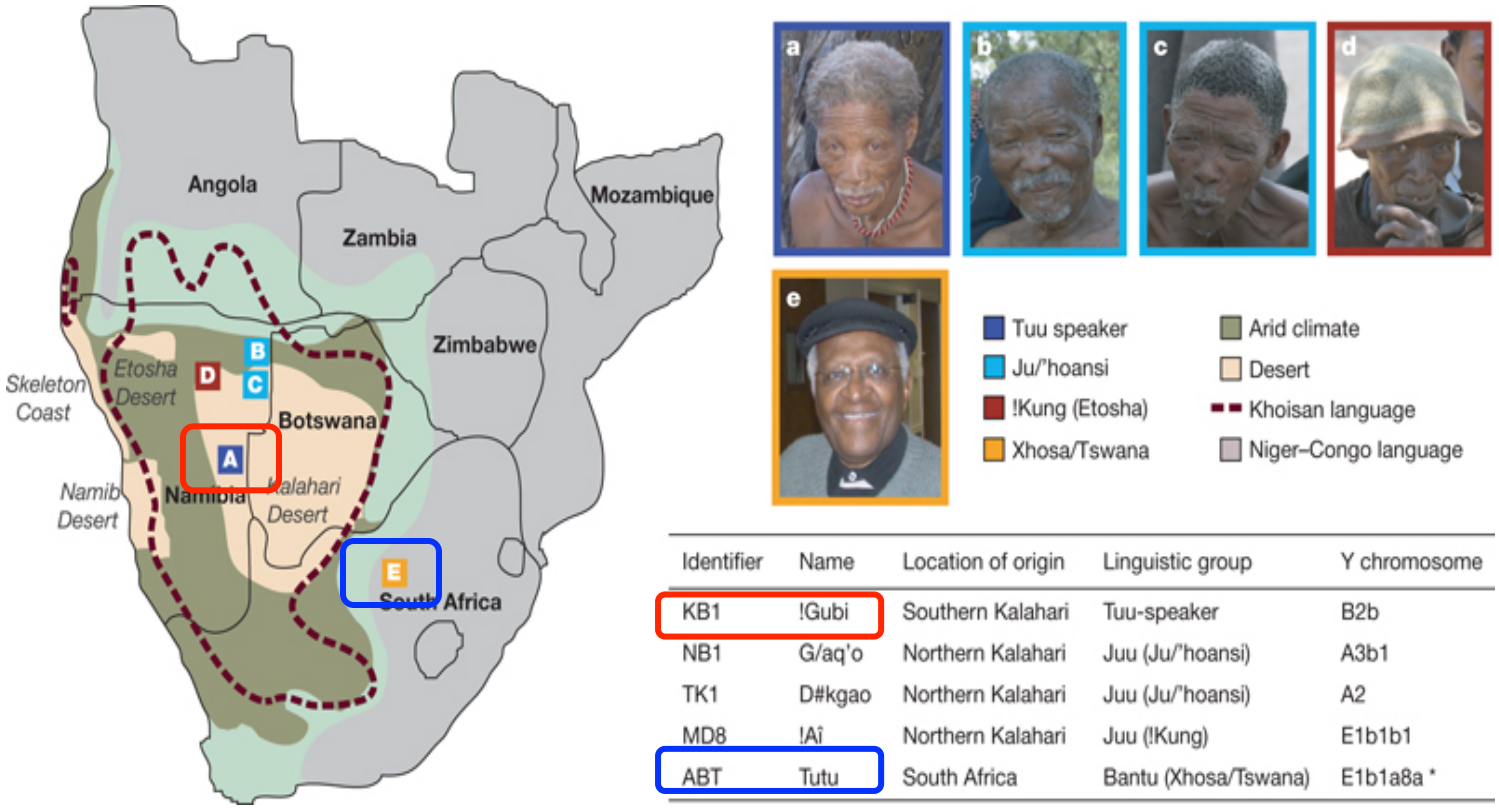
**Table 1.** Analysis of molecular variance (AMOVA). Eurasia, which encompasses Europe, the Middle East, and Central/South Asia, is treated as one region in the five-region AMOVA but is subdivided in the seven-region design. The World-B97 sample mimics a previous study (*6*).

| Sample | Number of regions | Number of populations | Variance components and 95% confidence intervals (%) | | |
|---|---|---|---|---|---|
| | | | Within populations | Among populations within regions | Among regions |
| World | 1 | 52 | 94.6 (94.3, 94.8) | 5.4 (5.2, 5.7) | |
| World | 5 | 52 | 93.2 (92.9, 93.5) | 2.5 (2.4, 2.6) | 4.3 (4.0, 4.7) |
| World | 7 | 52 | 94.1 (93.8, 94.3) | 2.4 (2.3, 2.5) | 3.6 (3.3, 3.9) |
| World-B97 | 5 | 14 | 89.8 (89.3, 90.2) | 5.0 (4.8, 5.3) | 5.2 (4.7, 5.7) |
| Africa | 1 | 6 | 96.9 (96.7, 97.1) | 3.1 (2.9, 3.3) | |
| Eurasia | 1 | 21 | 98.5 (98.4, 98.6) | 1.5 (1.4, 1.6) | |
| Eurasia | 3 | 21 | 98.3 (98.2, 98.4) | 1.2 (1.1, 1.3) | 0.5 (0.4, 0.6) |
| Europe | 1 | 8 | 99.3 (99.1, 99.4) | 0.7 (0.6, 0.9) | |
| Middle East | 1 | 4 | 98.7 (98.6, 98.8) | 1.3 (1.2, 1.4) | |
| Central/South Asia | 1 | 9 | 98.6 (98.5, 98.8) | 1.4 (1.2, 1.5) | |
| East Asia | 1 | 18 | 98.7 (98.6, 98.9) | 1.3 (1.1, 1.4) | |
| Oceania | 1 | 2 | 93.6 (92.8, 94.3) | 6.4 (5.7, 7.2) | |
| America | 1 | 5 | 88.4 (87.7, 89.0) | 11.6 (11.0, 12.3) | |



But - There Are Differences! But…They Fall Into Geographical Groups -- Groups Divided Originally by Geographic Barriers (Ocean, Desert, Mountains). The 5% Difference Allows Us to Mark and Trace Ancestry!

# Recent Sequencing of Two African Genomes Reveals Remarkable Genetic Diversity




| Identifier | Name | Location of origin | Linguistic group | Y chromosome |
|---|---|---|---|---|
| KB1 | !Gubi | Southern Kalahari | Tuu-speaker | B2b |
| NB1 | G/aq'o | Northern Kalahari | Juu (Ju/'hoansi) | A3b1 |
| TK1 | D#kgao | Northern Kalahari | Juu (Ju/'hoansi) | A2 |
| MD8 | !Aî | Northern Kalahari | Juu (!Kung) | E1b1b1 |
| ABT | Tutu | South Africa | Bantu (Xhosa/Tswana) | E1b1a8a * |

*SC Schuster et al. Nature 463, 943-947 (2010)*

**Each Genome Contains One Million SNPs Not Found in Any Other Genome**

# Conclusions

1. If 85% of Human Genetic Variation Occurs Between Different People Within Any Given Population (localized)
2. If only 7% of Human Genetic Variation Occurs Between "Races" (novel alleles specific to "races") e.g. $F_{yB}$ $^{ES}$
3. Then Losing all "Races" Except One Retains 94% of all Human Genetic Variation!

[85% + (15%-7%) ]= 94%

Variation That Occurs in Ancestral Population

85% Within Population genetic variability

8% Between Populations of Same "Race"

7% Between "Race" Genetic Variability

4. ∴ Human Highly Heterozygous or Hybrids- & If Above Not True- Most of Us Would Not Be Here- Need Genetic Variation to Survive!

# So What is a "Race"?

1. Primarily a sociological concept- but could be a localized or inbred population that has a higher frequency of alleles at a very small number of loci. Affects few physical features.
2. High frequency alleles in one "race" are present at lower frequencies in other "races". All humans have same genes- differ in form mostly within populations!
3. Heterozygosity (variation) high in human populations- all populations. None homozygous at all loci!
4. No such thing as a "pure" race - would have little variation
5. Genes affecting physical features not representation of genes across genome-" selected" traits

Geographical Ancestry is relevant-many "racial" groups now have multiple ancestries because of admixture and migration

A Better Term is POPULATIONS!

# *Knowledge or Certainty: The Ascent of Man Series*

*Jacob Bronowski, 1973*