

## The Human Genome Project



Cynthia P. Palmes-Saloma, Ph.D.  
Lab. of Molecular and Cellular Biology  
NIMBB, UP Diliman

## The Human Genome Project

The Human Genome Project is a global initiative to map and sequence the human genome

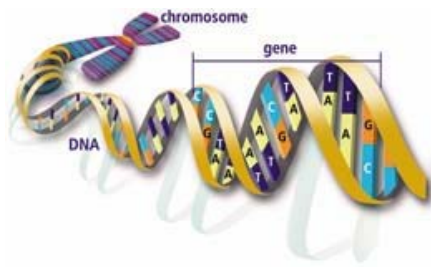


Image credit: U.S. Department of Energy Human Genome Program, <http://www.ornl.gov/hgms>

## The Human Genome Organization

- The Human Genome Organisation (HUGO) is the international organization of scientists involved in the Human Genome Project (HGP). HUGO was established in 1989 by a group of the world's leading genome scientists to promote international collaboration within the project.



## International Human Genome Sequencing Consortium

- The public consortium is composed of 16 genome centers all throughout the world with the five major sequencing centers being

- Sanger Centre in the UK;
- U.S. Department of Energy Joint Genome Institute in Walnut Creek, California;
- Washington Univ. School of Medicine in St. Louis;
- The Whitehead Institute for Biomedical Research in Cambridge, Massachusetts; and
- Baylor College of Medicine in Houston



- Privately-funded effort to sequence the human genome
- Celera Genomics of Rockville, Maryland led by J. Craig Venter



## The Human Genome Project

### Project goals

- determine the sequences of the 3 billion chemical base pairs that make up human DNA
- identify all the genes in human DNA
- store the information in databases
- improve tools for data analysis
- transfer related technologies to the private sector, and
- address the ethical, legal, and social issues (ELSI) that may arise from the project

[www.ncbi.nlm.nih.gov/genome/guide/human](http://www.ncbi.nlm.nih.gov/genome/guide/human)

## Whose genome is being sequenced?

- the first reference genome is a composite genome from several different people

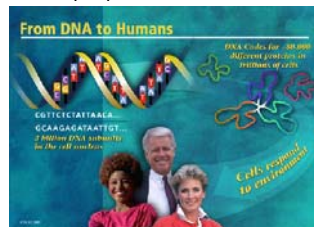
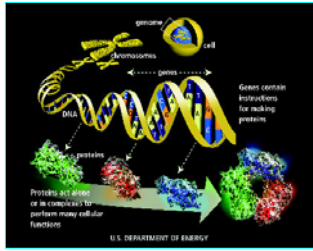


Image credit: U.S. Department of Energy Human Genome Program, <http://www.ornl.gov/hgms>

- generated from 10-20 primary samples taken from numerous anonymous donors across racial and ethnic groups

## What have been known so far from the HGS?

- ▶ size of human genome: 3.1647 billion base pairs (bp)
- ▶ number of human genes: ~30-35,000
- ▶ genes vary in length and can cover thousands of bases
  - avg. size: ~3,000 bp
- ▶ less than 2% of the human genome codes for proteins
- ▶ ("junk DNA") make up at least 50% of the human genome



- ▶ almost all (99.9%) nucleotide bases are exactly the same in all people

FACT: Humans share most of the same protein families with worms, flies, and plants, but the number of gene family members has expanded in humans, especially in proteins involved in development and immunity

- Mammalian models
  - MOUSE
  - RAT
- Non-mammalian models
  - *S. cerevisiae* (budding yeast)
  - *D. discoideum* (social amoebae)
  - *D. melanogaster* (fruit fly)
  - *D. rerio* (zebrafish)
  - *Xenopus* (frog)
- Plant
  - Arabidopsis

## Other genomes sequenced to date

Organism	Year	Genome size	Est. No. of Genes
<i>S. cerevisiae</i>	1996	12 MB	6,241
<i>C. elegans</i>	1998	97 MB	18,424
<i>Arabidopsis thaliana</i>	2000	100 MB	25,000
<i>D. melanogaster</i>	2000	122 MB	13,600
<i>Homo sapiens</i>	2001	3,000 MB	30-35,000
<i>Fugu rubripes</i>	2002	365 MB	38,000
<i>Anopheles gambiae</i>	2002	278 MB	14,000
<i>Plasmodium falciparum</i>	2002	23 MB	5,300
<i>Oryza sativa ssp. indica</i>	2002	466 MB	46-55,000
<i>Oryza sativa ssp. japonica</i>	2002	420 MB	32-50,000
<i>Mus musculus</i>	2002	3,000 MB	30,000

## Uses of the HGS information

- ▶ Improvements in medicine
- ▶ microbial genome research for fuel and environmental cleanup
- ▶ DNA forensics (identification)
- ▶ more accurate risk assessment
- ▶ improved agriculture and livestock
- ▶ better understanding of evolution and human migration

## Benefits of HGP Research

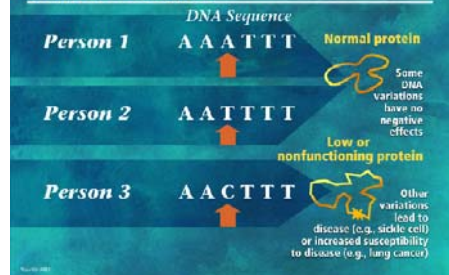
### Medical Benefits

- ▶ improved diagnosis of disease
- ▶ earlier detection of predispositions to disease
- ▶ rational drug design
- ▶ gene therapy and control systems for drugs
- ▶ pharmacogenomics "personal drugs"
- ▶ organ replacement



DNA in a bottle

## Health or Disease?



## RFLP- Restriction Fragment Length Polymorphism

- ➔ Restriction fragment length analyses uses restriction enzymes (RE) to cut DNA at specific 4-6 bp recognition sites
- ➔ Sample DNA is cut (digested) with one or more RE's and resulting fragments are separated according to molecular size using gel electrophoresis
- ➔ Ethidium bromide staining is used to reveal the fragments under UV (260 nm) light
- ➔ Molecular size standards are used to estimate fragment size
- ➔ Differences result from base substitutions, additions, deletions or sequence rearrangements within RE recognition sequences.
- ➔ Presence and absence of fragments resulting from changes in recognition sites are used identifying individuals, species or populations or in determining a healthy or a diseased gene.

[http://nematode.unl.edu/its\\_id/pcr\\_rflp.htm](http://nematode.unl.edu/its_id/pcr_rflp.htm)

## RFLP- Restriction Fragment Length Polymorphism

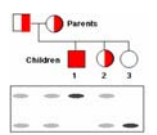
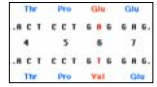
Restriction enzymes - are like molecular scissors that cut DNA at precise points (e.g. sequences)

RFLP products may be:

- a collection of DNA fragments of precisely defined length
- separated by electrophoresis with the smaller DNA fragments migrating faster than the larger ones

RFLPs are useful for:

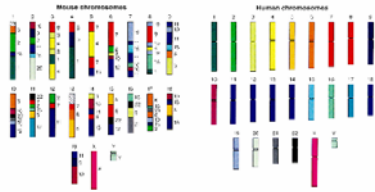
- screening human DNA for the presence of potentially deleterious genes
- providing evidence to establish the innocence of, or a probability of the guilt of, a crime suspect by DNA "fingerprinting"



## Understanding gene function

- To find genes found important in other species in humans

### Mouse and Human Genetic Similarities



## Benefits of HGP Research

### Microbial Genome Research

- new energy sources (biofuels)
- environmental monitoring to detect pollutants
- protection from biological and chemical warfare
- safe, efficient toxic waste cleanup

## Benefits of HGP Research

### DNA Forensics

- identify potential suspects at crime scenes
- exonerate wrongly accused persons
- identify crime and catastrophe victims
- establish paternity and other family relations
- identify endangered and protected species as an aid to wildlife officials (prosecution of poachers)

## Benefits of HGP Research

### DNA Forensics

- detect bacteria and other organisms that may pollute air, water, soil, and food
- match organ donors with recipients in transplant programs
- determine pedigree for seed or livestock breeds
- authenticate consumables such as caviar, meat and wine

## Benefits of HGP Research

### Evolution and Human Migration

- study migration of different population groups based on female genetic inheritance
- study mutations on the evolutionarily stable Y chrom. to trace lineage and migration



## Benefits of HGP Research

### Risk Assessment

- assess health damage and risks caused by radiation exposure, including low-dose exposures
- assess health damage and risks caused by exposure to mutagenic chemicals and cancer-causing toxins
- reduce the likelihood of heritable mutations



## Future challenges

### • Functional genomics

- transcriptomics: what, where and how are genes expressed
- proteomics: how are proteins expressed and what functions do they fulfill
- structural genomics: what are the functions of DNA sequences and the proteins they encode
- comparative genomics: how to identify human genes and predict their gene functions from studies of model organisms

## Future challenges

- Improve content and utility of databases
- Develop better tools for data generation, capture, and annotation
- Develop and improve tools and databases for comprehensive functional studies
- Develop and improve tools for representing and analyzing sequence similarity and variation
- Create mechanisms to support effective approaches for producing robust, exportable software that can be widely shared

## Information sources

To learn more about the Human Genome Project and its underlying science and applications

- [www.ornl.gov/hgmis](http://www.ornl.gov/hgmis)
- [www.ncbi.nlm.nih.gov/genome/guide/human](http://www.ncbi.nlm.nih.gov/genome/guide/human)
- *Nature* (Feb. 15, 2001)
- *Science* vol. 291 (16 February 2001) "The Human Genome" issue