



The Human Genome Project

From genome to health

From human genome to other genomes and to gene function

Structural Genomics initiative

What is the Human Genome Project?

- U.S. govt. project coordinated by DOE and NIH
- Initiated in 1990
- ***goals (1998-2003)***
 - identify the approximate 100,000 genes in human DNA
 - determine the sequences of the >3 billion bases that make up human DNA
 - store this information in databases
 - develop tools for data analysis
 - address the ethical, legal, and social issues that arise from genome research

The Human Genome Project

*Project began in 1990 as a \$3 billion, 15-year effort
18 countries participate in the worldwide effort*

*About 3 yrs ago -A private
company - **Celera**
announced that the human
genome will be completed
within 3 years.*



The screenshot shows the top portion of the Celera Science website. At the top left is the Celera logo, which includes a stylized 'X' and the text 'CELERA A PHILIPPS-BROOKS COMPANY'. To the right of the logo are two navigation tabs: 'Discovery System' and 'News'. Below the navigation is the heading 'Welcome to CeleraScience'. The main text reads: 'CeleraScience accelerates scientific discovery by making genomic data, analysis tools, and resources available through a highly robust and integrated delivery platform.' Below this, it says: 'Now available to the Public in Beta - the Celera Discovery System™ provides integrated sequence data, including the Drosophila Genome, and analysis tools to the research community. Click on the Products & Services tab above for a detailed description of the expanded product offerings available through subscription.'

*The human genome was immediately associated with additional
sequencing projects - microbial, pathogen, plants, Fugu fish, mouse
etc*



The SOMP-BC Fugu Project is a community resource project funded by the British Columbia Science and Technology Centre, through the SOMP-BC Genome Project. The project is a collaboration between the University of British Columbia, the British Columbia Science and Technology Centre, and the SOMP-BC Genome Project. The project is a collaboration between the University of British Columbia, the British Columbia Science and Technology Centre, and the SOMP-BC Genome Project. The project is a collaboration between the University of British Columbia, the British Columbia Science and Technology Centre, and the SOMP-BC Genome Project.

The Human Genome Project

September 1999 public announcement:

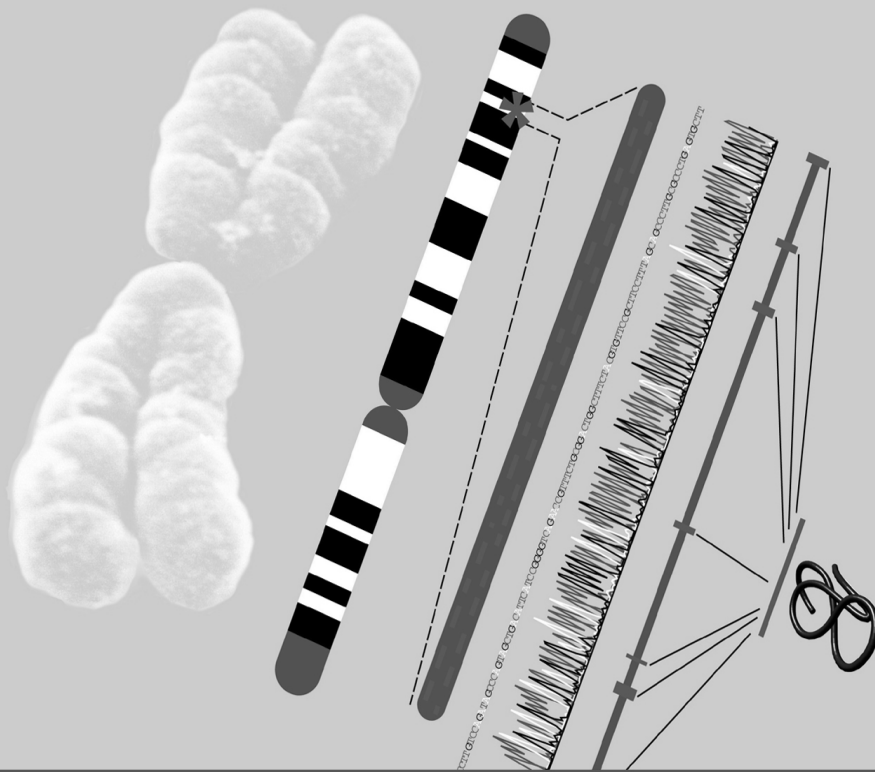
A rough draft by spring 2000

- ***The draft***

About 90% of the human genome

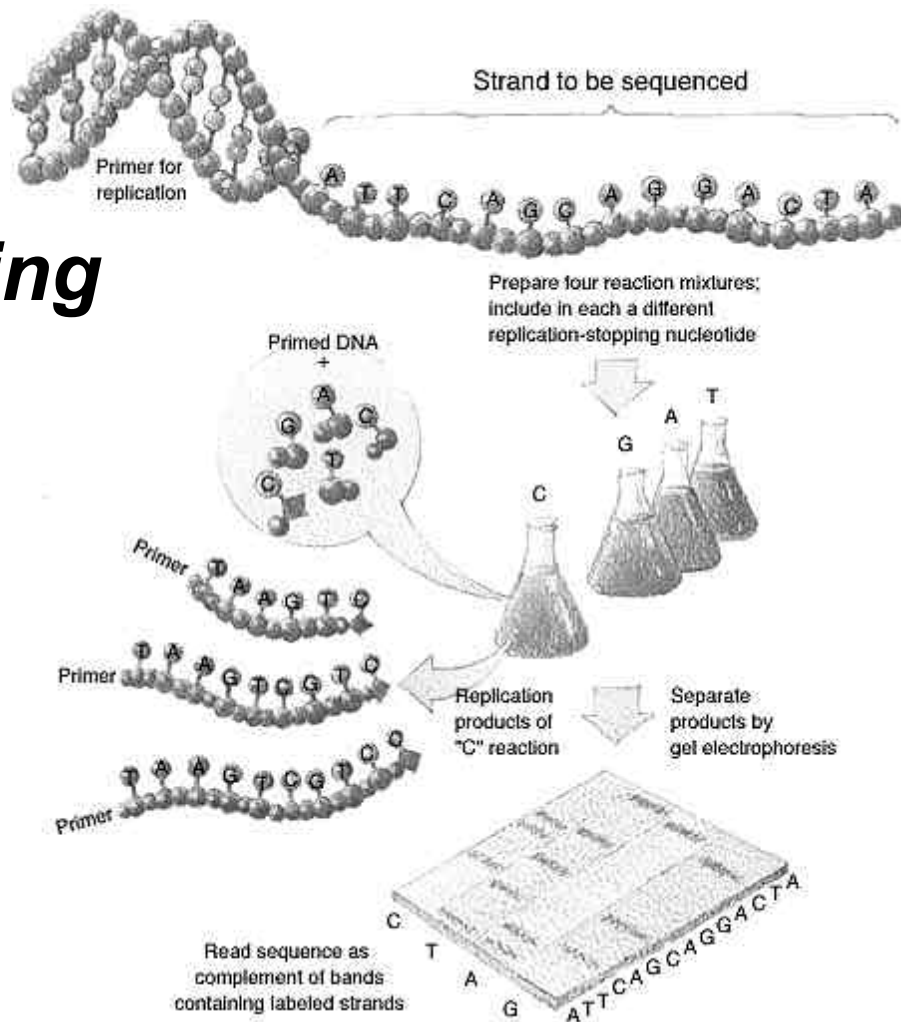
A complete, high-quality DNA sequence - by 2003

Generating the draft will require to sequence each piece of DNA only about 5 times, instead of the usual 10 times used for obtaining the highest quality sequence (99.99%).



From Chromosomes to Proteins
Courtesy U.S. Department of Energy Human Genome Program

DNA sequencing



The Human Genome Project

From the news:

The Human Genome Project international consortium announced (April 2000) that 2 billion of the 3 billion “letters” that constitute the genetic instruction book of humans have been deciphered and deposited into GenBank.

GenBank, the public database of DNA sequence operated by the National Institutes of Health, is accessible freely and without restrictions to all scientists in industry and academia.

The Human Genome Project

Some numbers:

Human Genome Project assembles 12,000 bases every minute.

15 billion raw base pairs were sequenced to reach the two billion.

Each area of a chromosome at least four to five times to insure that the data deposited is accurate.

The “depth of coverage,”

Cancer genome projects and the search for disease genes

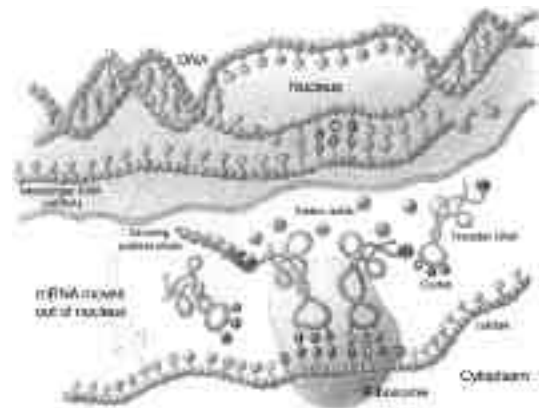
- *One of the challenges of the post-genomic era is to establish the role of genes in common **multifactorial disease.***
 - ***Cancer in the western world***
 - *About one in three people will get cancer over their lifetime, and about one in five will die of it.*
- The goal - discovery of the disease genes***

Cancer genome projects

During life, human cells acquire changes to their DNA as a result of :

Exposure to the environment - chemicals, radiation or viruses

Errors in the enzymatic machinery that copies DNA during cell division



Cancer genome projects

Most of these changes are put right by the cell's repair mechanisms.

If mistakes occur in genes known as oncogenes or cancer genes, the cell in which those changes occur can start to behave abnormally.

The cell may continue dividing when it should be stopping



Cancer genome projects

Following isolating the genes...

Which genes are abnormal in which cancers

Which cancer have many abnormalities and which have few abnormalities

In Sanger center (UK)

screen 2000 cancers of all types - childhood, adult, epithelial and soft tissue

DNA Sequence Variation in a Gene Can Change the Protein Produced by the Genetic Code

Gene A from Person 1

GCA AGA GAT AAT TGT...
↓ ↓ ↓ ↓ ↓
Ala Arg Asp Asn Cys ...
1 2 3 4 5

Protein Products



Gene A from Person 2

Codon change made no difference in amino acid sequence

GCG AGA GAT AAT TGT...
↓ ↓ ↓ ↓ ↓
Ala Arg Asp Asn Cys ...
1 2 3 4 5

Gene A from Person 3

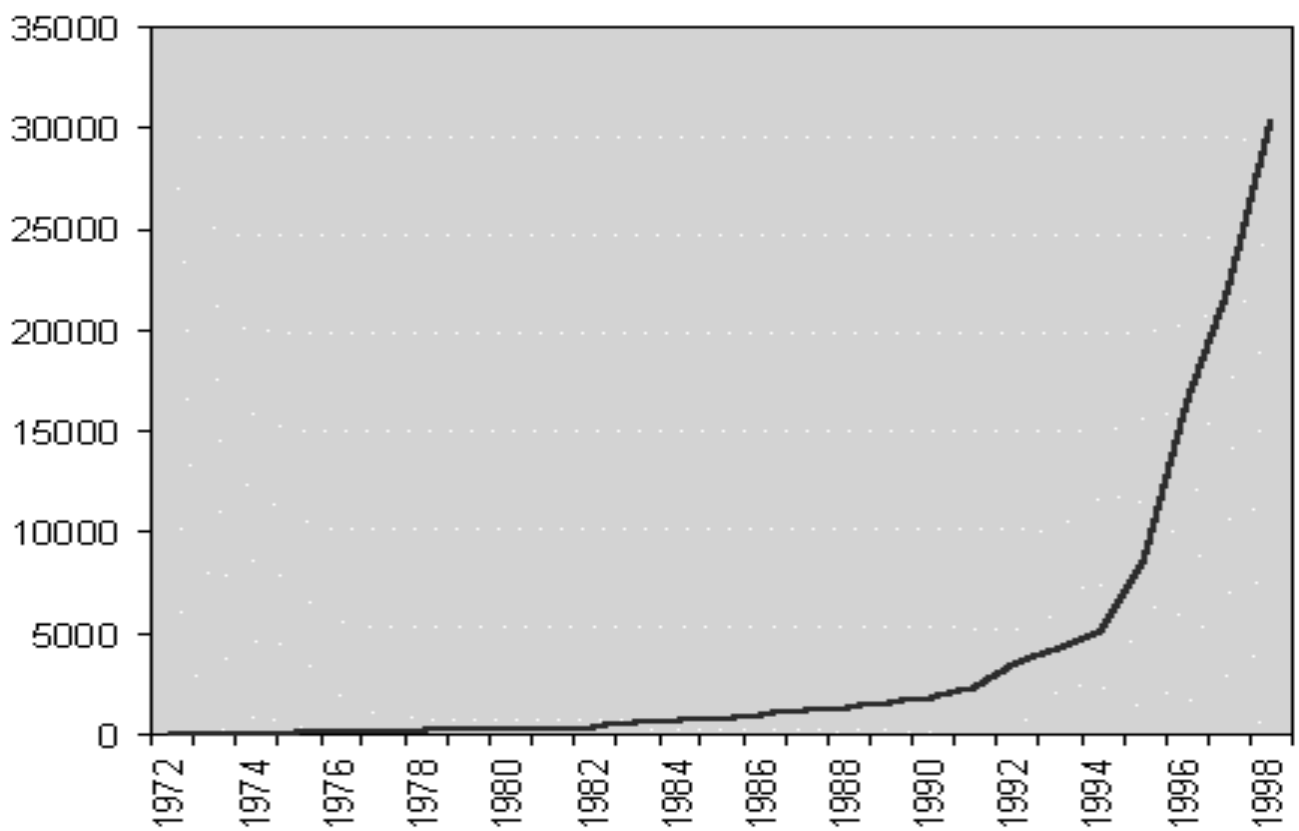
Codon change resulted in a different amino acid at position 2

GCA AAA GAT AAT TGT...
↓ ↓ ↓ ↓ ↓
Ala Lys Asp Asn Cys ...
1 2 3 4 5

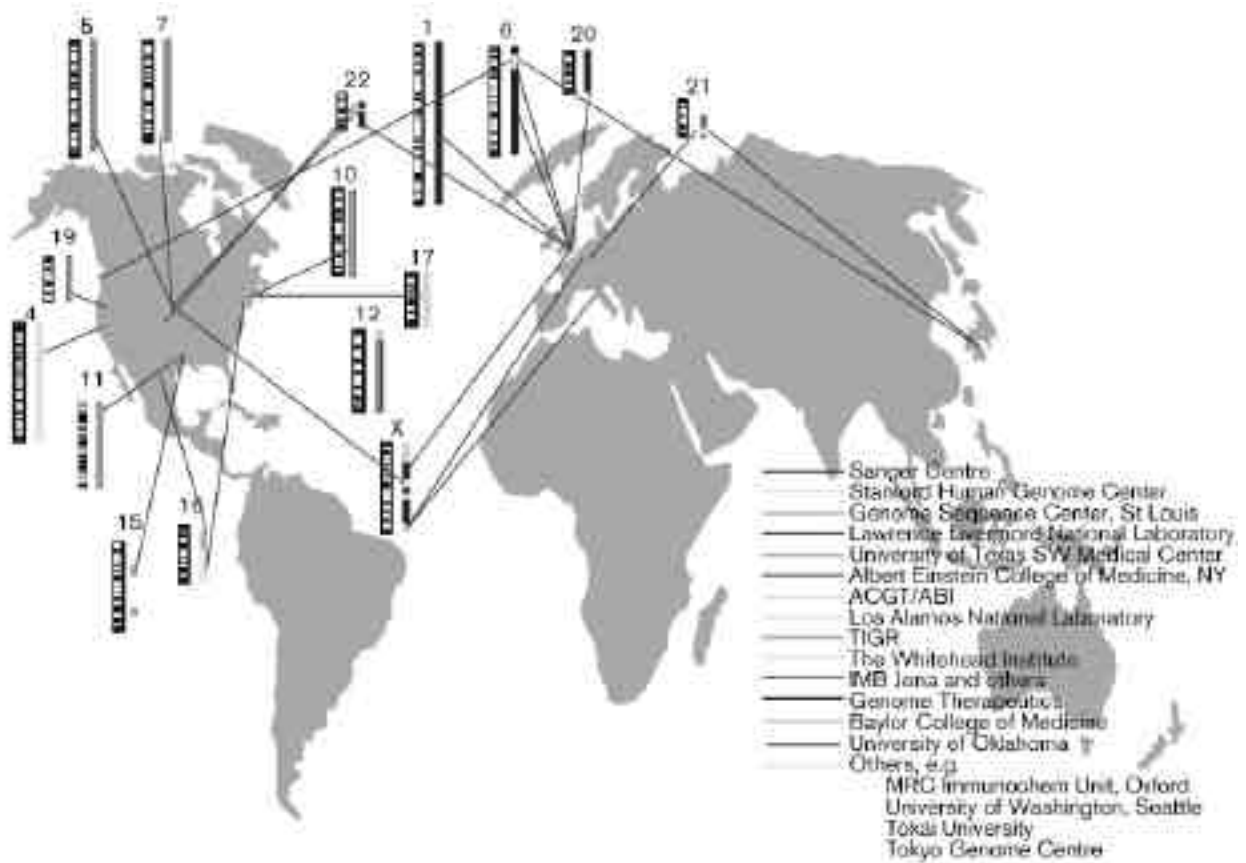
OR



The growth in sequencing - last decade



Chromosomes -over the globe



Clifford Cooper in Biotechnology

Chromosomes content

***Chromosome 5 -
194 million bases,
6 % of the genome.***

***Disease-linked genes:
colorectal cancer,
basal cell carcinoma,
acute myelogenous leukemia,
salt-resistant hypertension
type of dwarfism.***

***Chromosome 16 98 million 3 % of the
genome.***

***breast and prostate cancer,
Crohn's disease
adult polycystic kidney disease (5 million
people worldwide. 50% require dialysis or kidney
transplant).***

***Chromosome 19 60 million, 2% of the
genome.***

***DNA damage
atherosclerosis
diabetes mellitus***

Chromosomes 22 landscape

Chromosome 22 - the longest sequenced piece (up to Chr.21)

545 genes, 134 pseudogenes and 200 to 300 additional ones.

Size of a gene - 1,000 to 583,000 bases of DNA.

39% of the chromosome is copied into RNA (exons and introns).

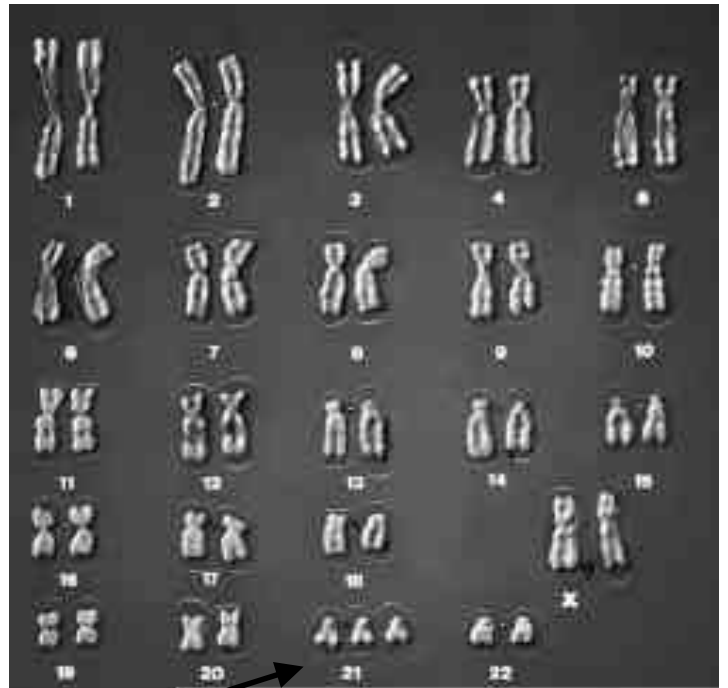
247 genes were revealed to be identical to previously identified genes.

There are families of genes that are distributed over large chromosomal regions.

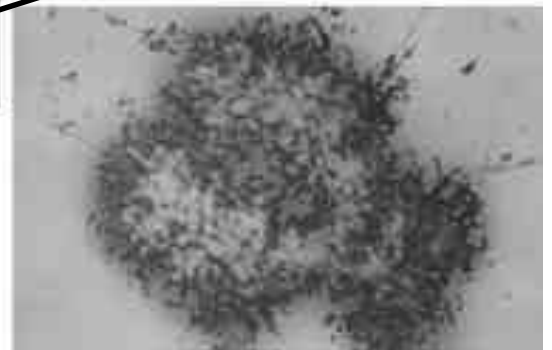
Unexpected long-range complexity of the chromosome with an elaborate array of repeat sequences near the centromere.

several regions where recombination is increased, and others where it is suppressed.

Human chromosomes



***Trisomy of chromosome 21
Down syndrome***



Nature, 18th May 2000

The latest -18 May 2000

articles

The DNA sequence of human chromosome 21

The chromosome 21 mapping and sequencing consortium

M. Hattori¹, A. Fujiyama², T. D. Taylor³, H. Watanabe⁴, T. Yada⁵, H.-S. Park⁶, A. Toyoda⁷, K. Ishii⁸, Y. Tofuku⁹, D.-K. Choi¹⁰, E. Soeda¹¹, M. Ohki¹², T. Takagi¹³, Y. Sakaki¹⁴; S. Taudien¹⁵, K. Blochschmidt¹⁶, A. Polley¹⁷, U. Menzies¹⁸, J. Delabar¹⁹, K. Kumpff²⁰, R. Lehmann²¹, D. Patterson²², K. Reichwald²³, A. Rump²⁴, M. Schilfhabel²⁵, A. Schudy²⁶, W. Zimmermann²⁷, A. Rosenthal²⁸; J. Kudoh²⁹, K. Shibuya³⁰, K. Kawasaki³¹, S. Asakawa³², A. Shintani³³, T. Sasaki³⁴, K. Nagamine³⁵, S. Nishiyama³⁶, S. E. Antonarakis³⁷, S. Minoshima³⁸, N. Shimizu³⁹, G. Nordstiek⁴⁰, K. Hornischer⁴¹, P. Brandt⁴², M. Scharfe⁴³, G. Schön⁴⁴, A. Desario⁴⁵, J. Reichel⁴⁶, G. Kauer⁴⁷, H. Blöcker⁴⁸; J. Ramser⁴⁹, A. Beck⁵⁰, S. Hagos⁵¹, S. Hennig⁵², L. Hesselmann⁵³, E. Dagand⁵⁴, T. Haaf⁵⁵, S. Wehrmeyer⁵⁶, K. Boczyn⁵⁷, K. Gardiner⁵⁸, D. Nizetic⁵⁹, F. Franco⁶⁰, H. Lehrach⁶¹, R. Reinhardt⁶² & M.-L. Yaspo⁶³

Correspondence: Reinhardt

¹RIKEN, Genomic Science Center, Saitama 351-8571, Japan

²Institut für Molekulare Biotechnologie, Genomanalyse, D-02748 Jena, Germany

³Department of Molecular Biology, Keio University School of Medicine, Tokyo 160-8582, Japan

⁴GDF (German Research Centre for Biotechnology), Genome Analysis, D-38124 Braunschweig, Germany

⁵Bio-PARC/Armita für Molekulare Genetik, D-14195 Berlin-Adlershof, Germany

Collaborating Institutions

⁶RIKEN, Life Science Tokyo/Ramsey Center, Tsukuba 305-0874, Japan

⁷Cancer Genomics Division, National Cancer Center Research Institute, Tokyo 104-0965, Japan

⁸Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo 108-8501, Japan

⁹UMR 6602 CNRS, UFR Nancy-Environnement, Paris 75731, France

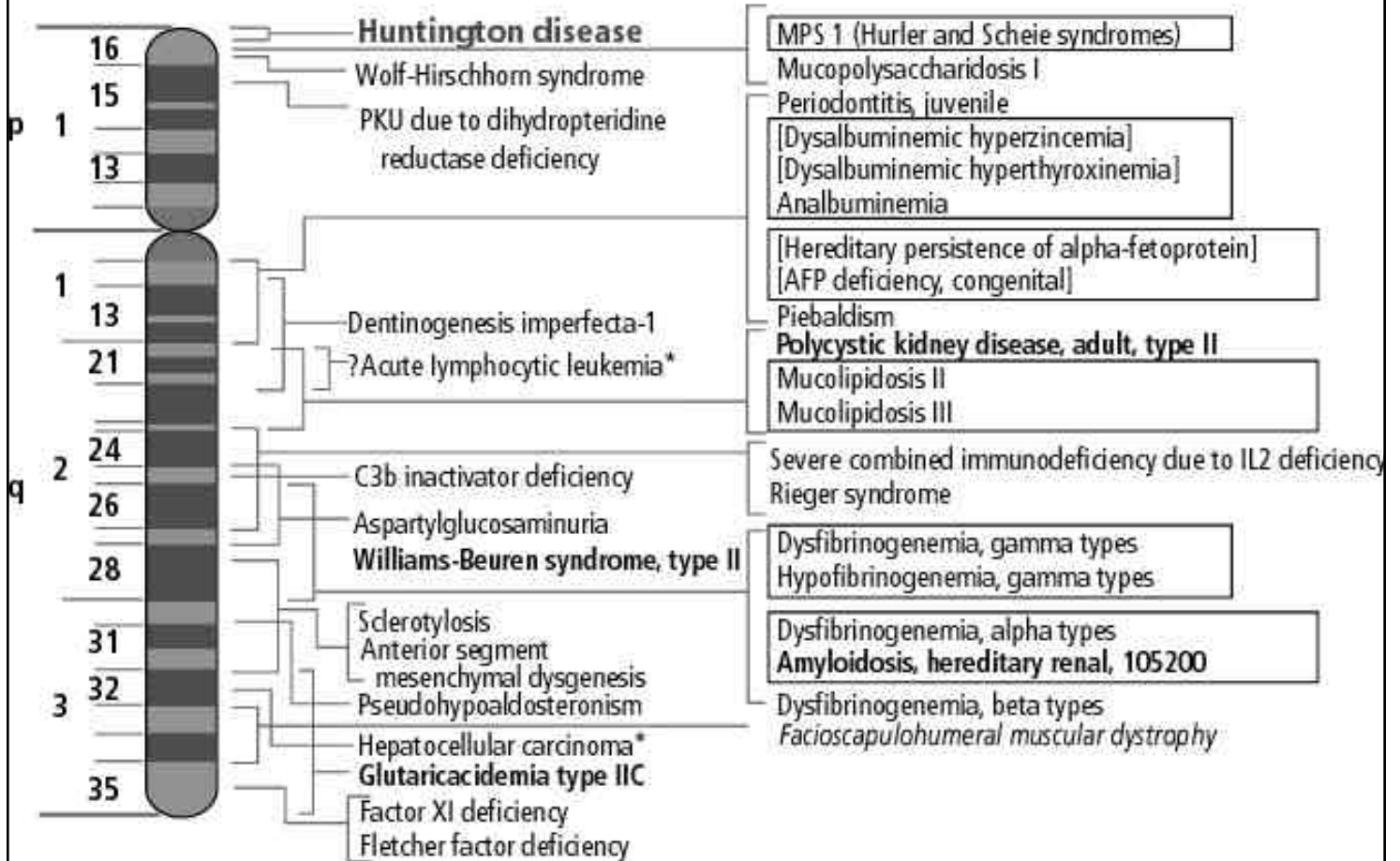
¹⁰Evans Research Institute, Denver, Colorado 80206, USA

¹¹Medical Genetics Division, University of Geneva Medical School, Geneva 1221, Switzerland

¹²CNRS UPR 1742, Institut de Biologie, Montpellier, 34093, France

¹³School of Pharmacy, University of London, London WC2N 1AL, UK

Chromosome 4



Latest news:

Genetic information on human chromosomes 5, 16 and 19.

The chromosomes contain an estimated 10-15,000 genes.

Genetically linked diseases :

kidney disease, prostate and colorectal cancer, leukemia, hypertension, diabetes and atherosclerosis.

300 million base pairs, or an estimated 11 percent of the total human genome

Cancer genome projects

How to isolate those genes?

About 100 cancer genes have been found

Most of the 100 genes have been found in the rare leukaemias and lymphomas, which account for less than 10 % of all human cancer

For the common adult cancers of the breast, colon, prostate, lung and ovary, which account for 80 % of the cancer burden, only about 30 genes are known.

Cancer genome projects

Future medicine...

The genes are likely to be the key targets for new types of therapy.

For example:

- ***If a gene is mutated so it is permanently active, a therapy could inactivate the protein.***
- ***If a gene has been turned off by the mutation, perhaps it could be turned back on or a second, normal copy introduced.***

Screening genome projects

Future medicine...

- ***similar kinds of screen can be performed for other human diseases that we know have a genetic component -***
- ***for example diabetes and heart disease.***
- ***Long term goal: Asthma, Dyslexia, Autism..***

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

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 - cont.

- 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 and wine

Benefits of HGP Research

Agriculture and Livestock

- disease-, insect-, and drought-resistant crops
- healthier, more productive, disease-resistant farm animals
- more nutritious produce
- biopesticides
- edible vaccines incorporated into food products
- new environmental cleanup uses for plants like tobacco

Benefits of HGP Research

Evolution and Human Migration

- use germline mutations in lineages to study evolution
- study migration of different population groups based on female genetic inheritance
- study mutations on the evolutionarily stable Y chromosome to trace lineage and migration
- compare breakpoints in the evolution of mutations with ages of populations and historical events

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 medicine

99.9 per cent of our DNA sequence is identical to that of all other humans

One nucleotide base in every 1000 of the 3 billion total, differs from one individual to another.

A simple blood sample could provide a complete genetic profile that predicts our susceptibility to disease, our response to drugs, our propensity to obesity or anxiety.

Differences among individuals

Each of us has a unique combination of polymorphisms (SNPs).

This polymorphism underlies the fact that some have blue eyes and some brown, some are tall and some short - and...

some suffer heart attacks in middle age while others are 'very healthy' at >90 years of age.

The role of Bioinformatics in the Human Genome Project

