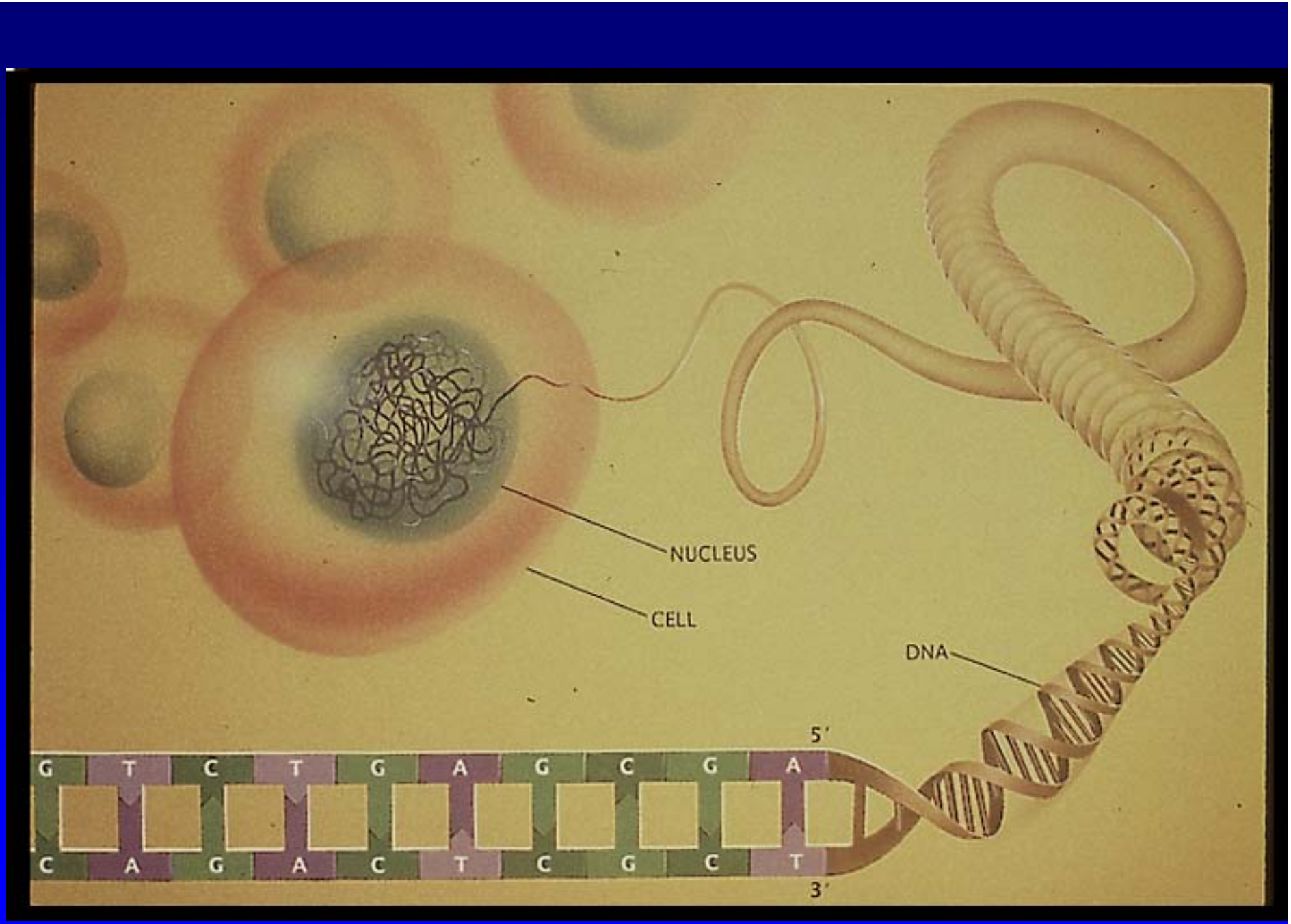
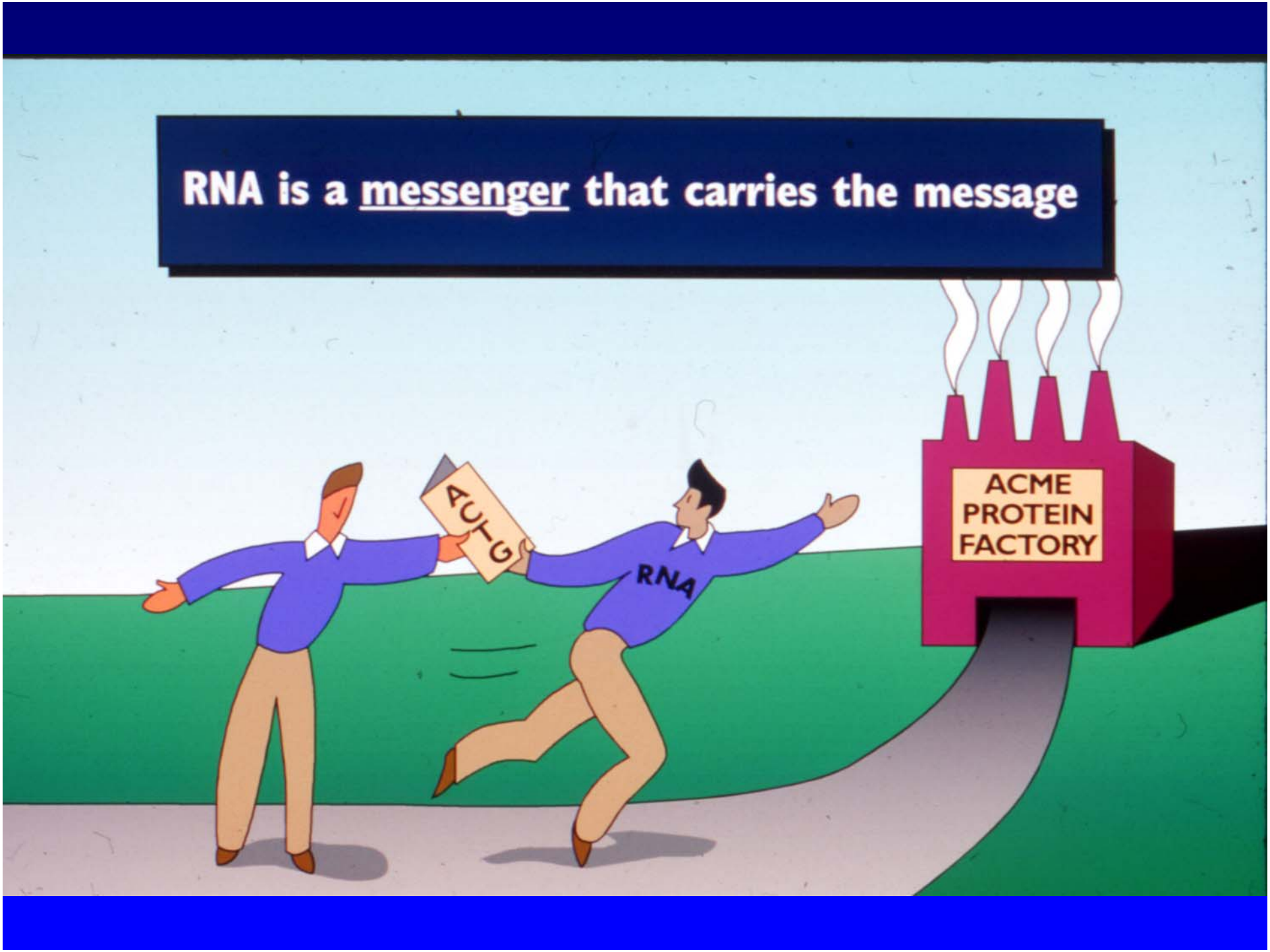
Genomics, Medicine, and Society

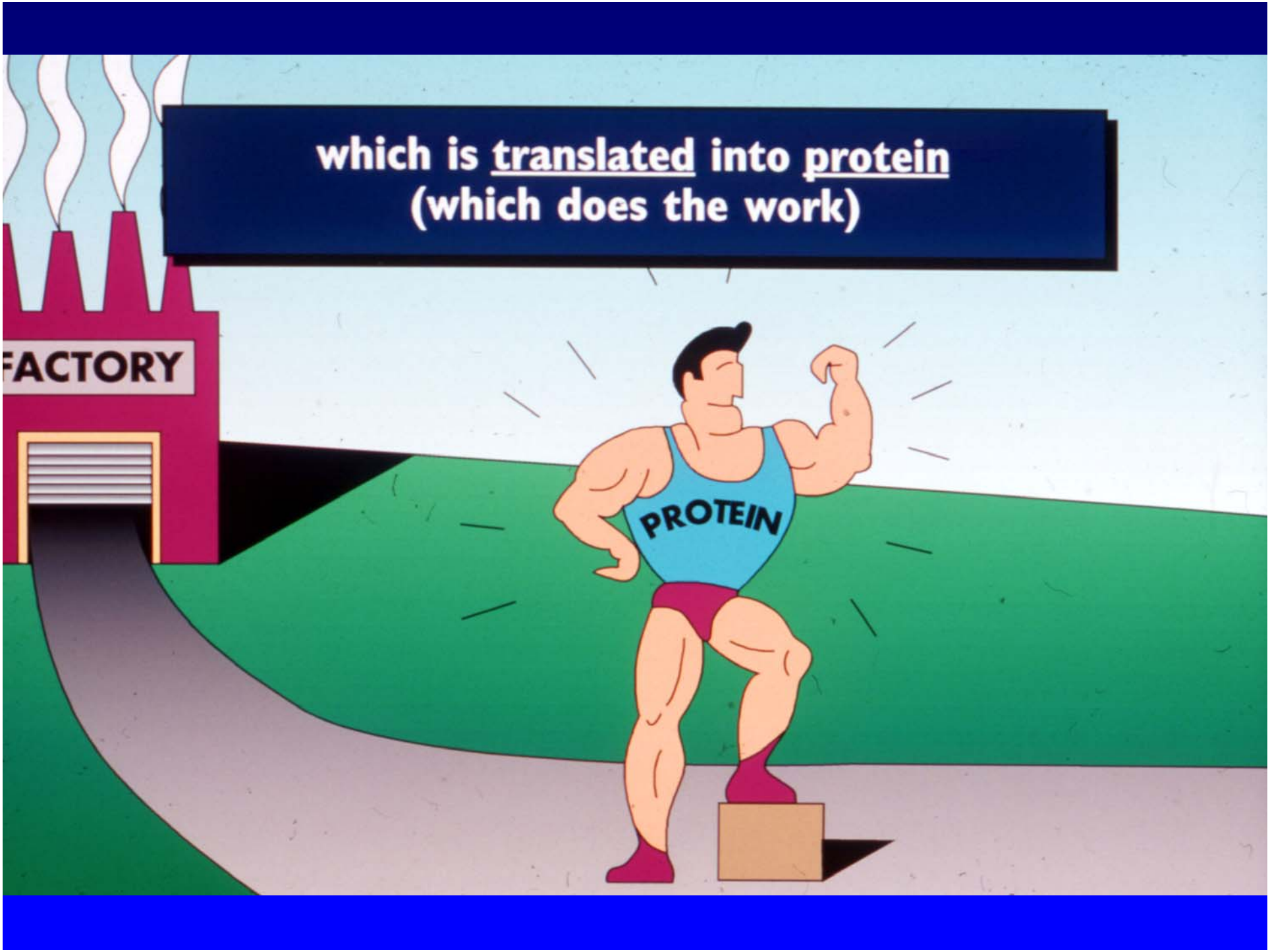
Francis S. Collins, M.D., Ph.D.
National Human Genome Research Institute

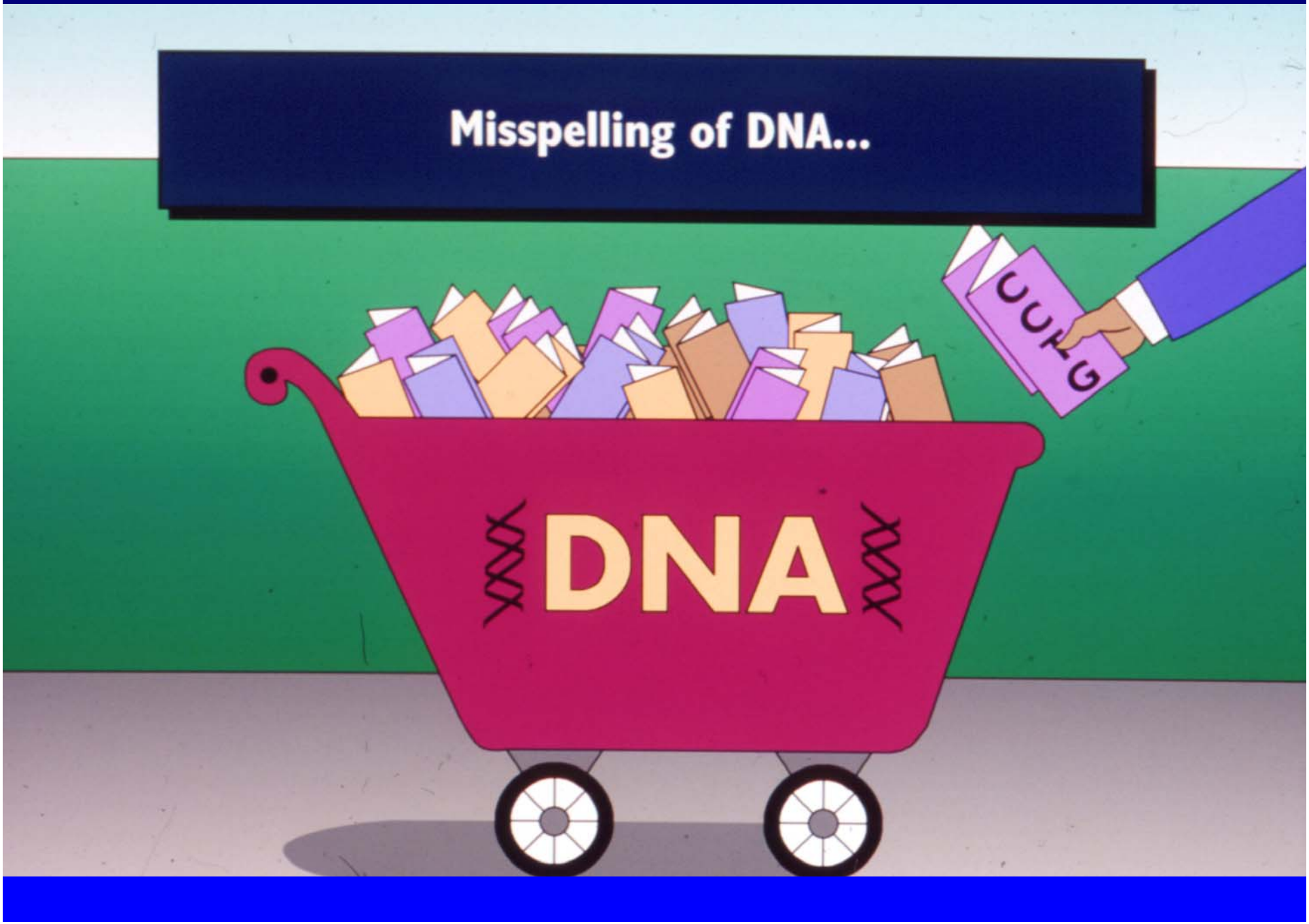
Presentation for House of Lords Visit
June 4, 2008


NUCLEUS
CELL
DNA
5′
G T C T G A G C G A
C A G A C T C G C T
3′

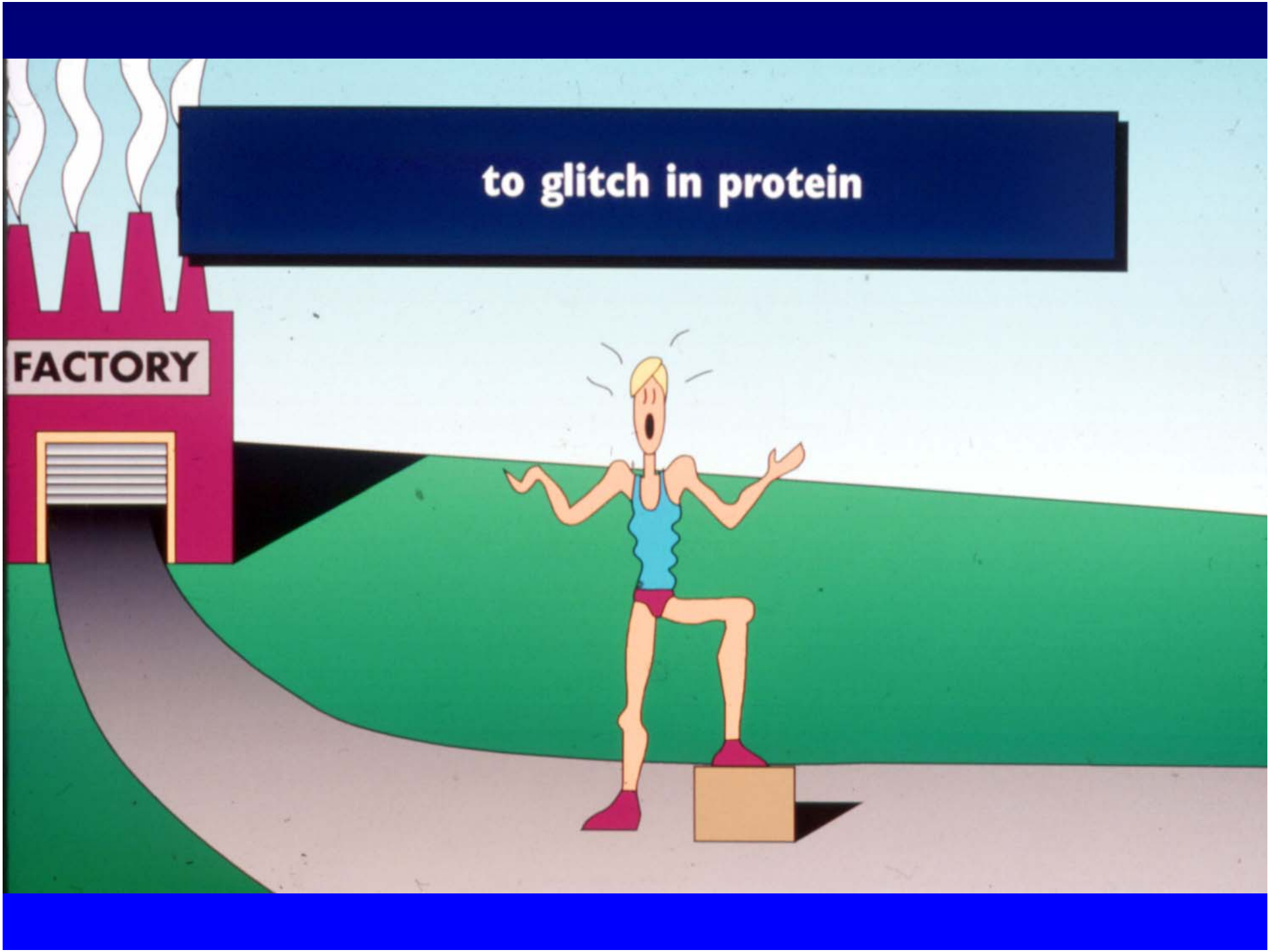

Instructions are transcribed into RNA
DNA


RNA is a messenger that carries the message
AUG
RNA
ACME
PROTEIN
FACTORY


which is translated into protein
(which does the work)
FACTORY
PROTEIN


Misspelling of DNA...
UUPG
DNA


Leads to glitch in RNA...
?
UUUG
RNA
ACME
PROTEIN
FACTORY


to glitch in protein
FACTORY

April, 1953 → April, 2003

No. 4356 April 25, 1953 NATURE

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

J. D. Watson
F. H. C. Crick

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge.

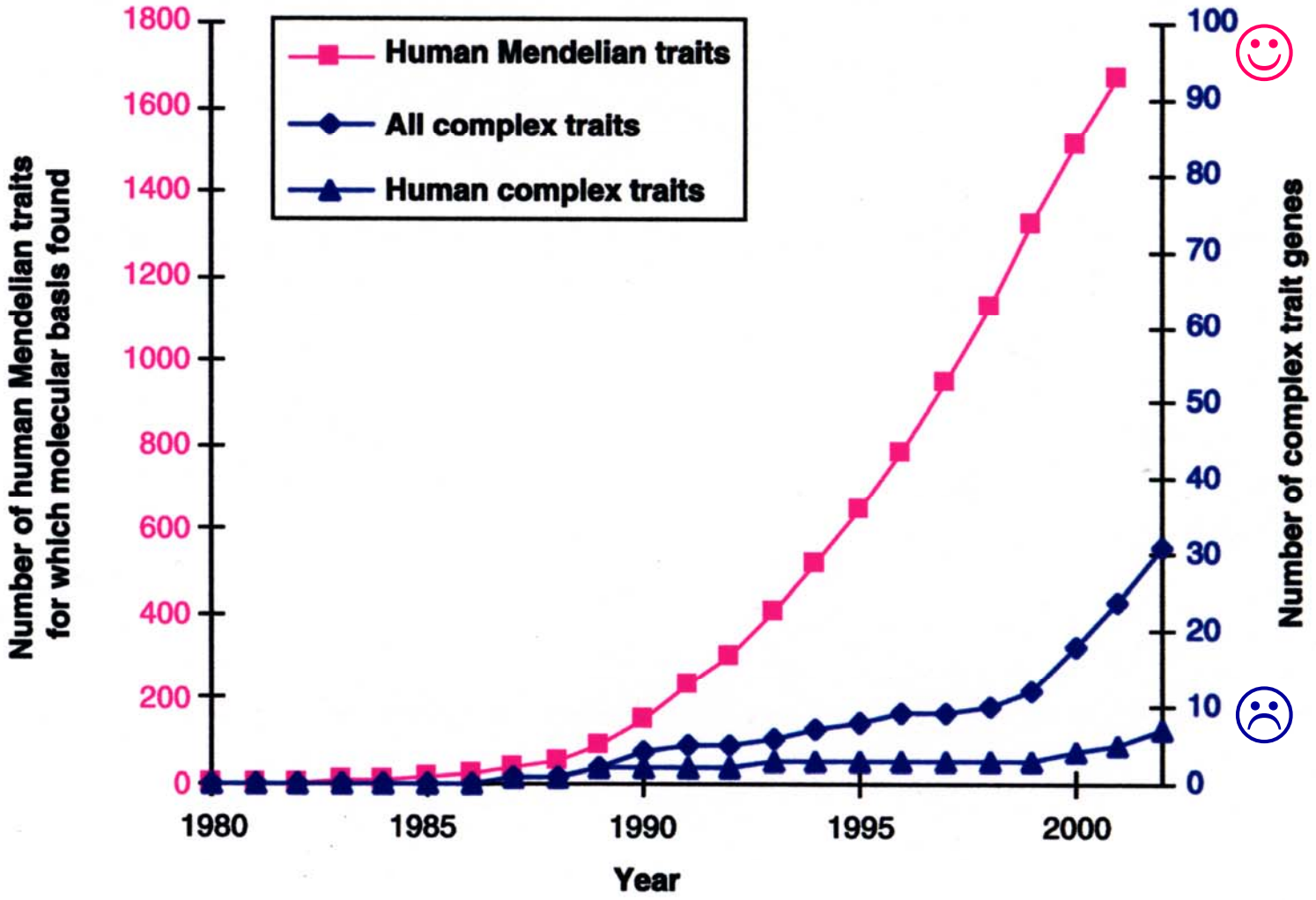
April 2.

Virtually All Diseases (Except Maybe Trauma) Have a Genetic Component

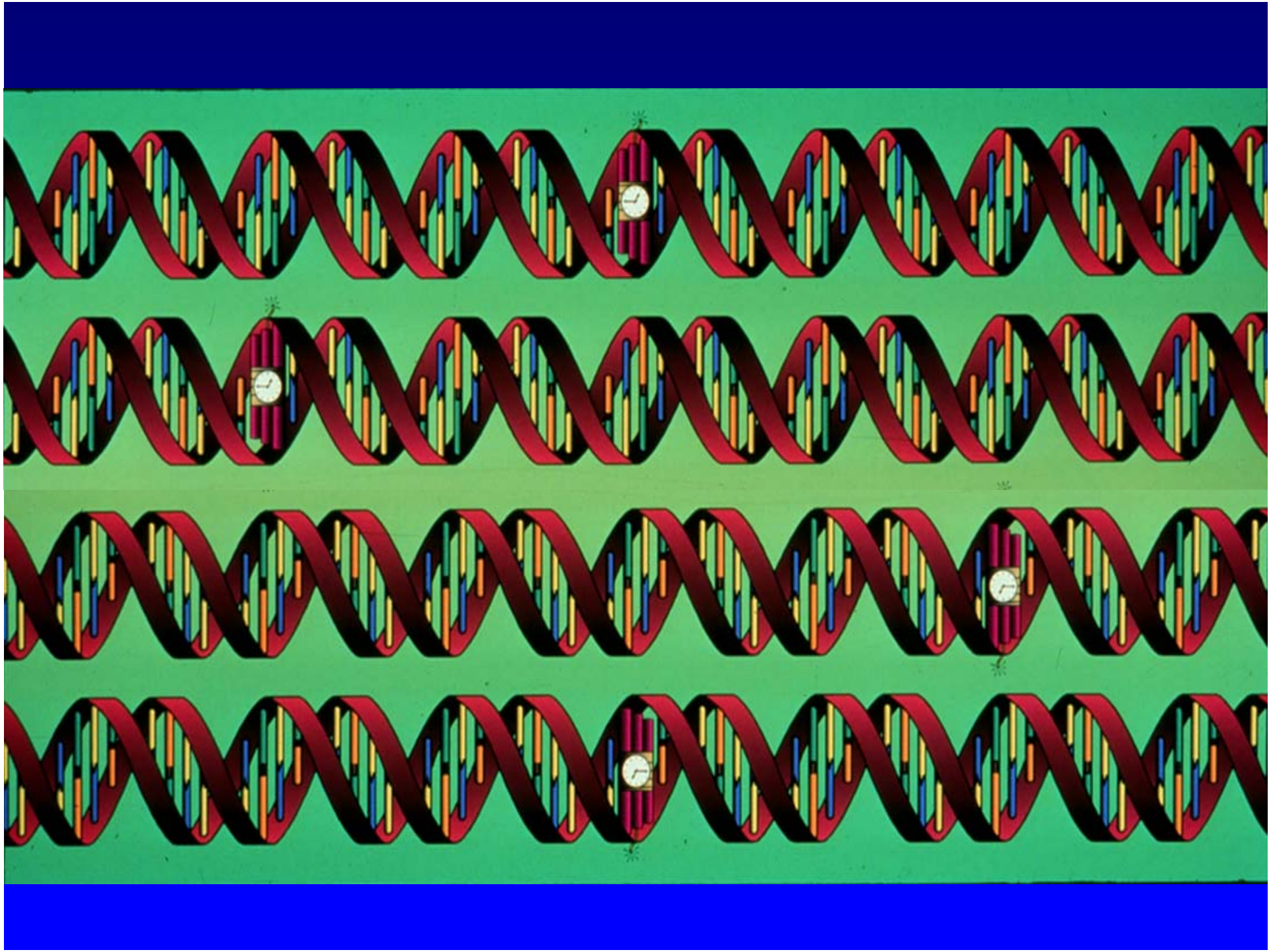
Cystic fibrosis	Adult onset diabetes	AIDS

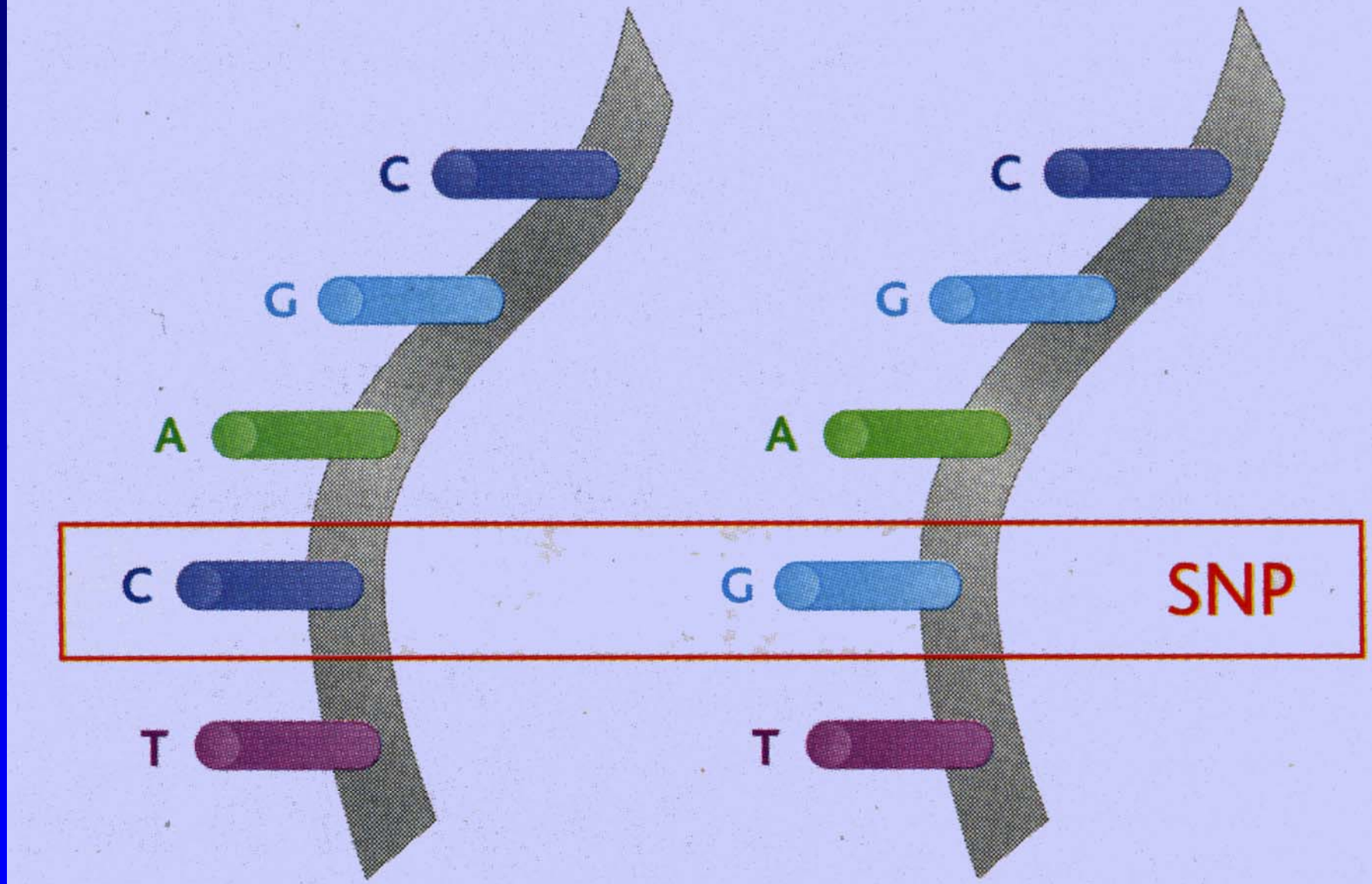
Genetic Component

Environmental Component

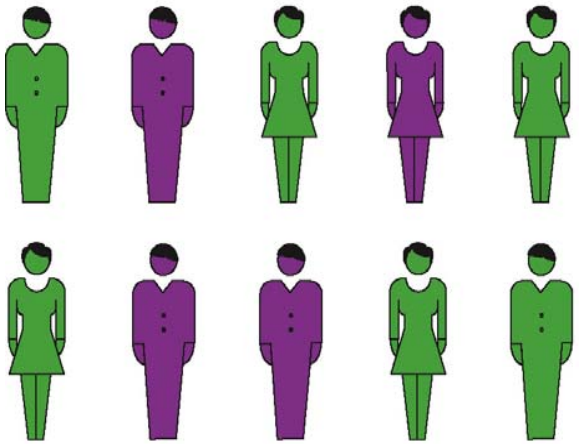


Glazier et al., Science 298:2345-9, 2002




C
G
A
C
T
C
G
A
G
T
SNP

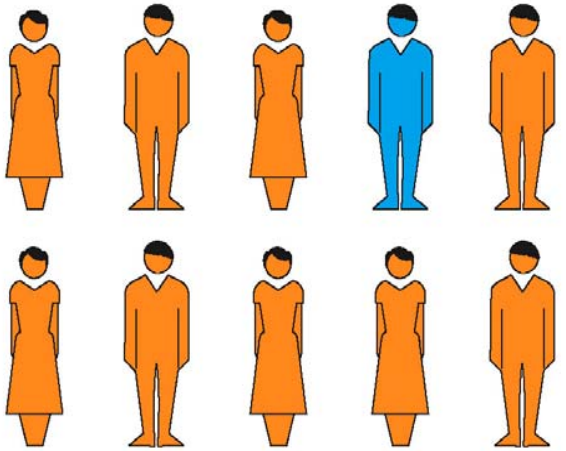
SNP A



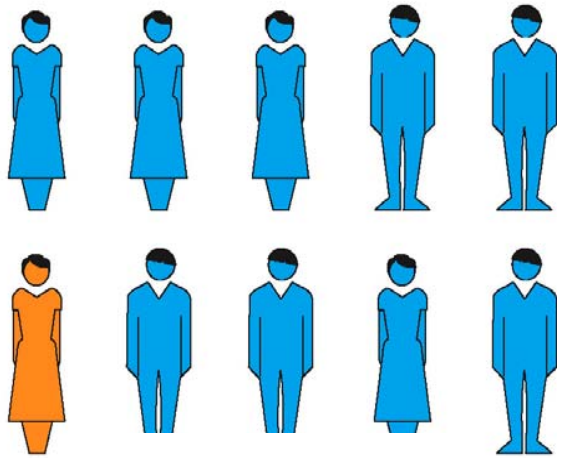
Diabetes

Unaffected

SNP B



Diabetes



Unaffected

"Genome Wide Association" Approach to Common Disease: The View from 2002

- Identify all 10 million common SNPs
- Collect 1000 cases and 1000 controls
- Genotype all DNAs for all SNPs
- That adds up to 20 billion genotypes
- At 50 cents a genotype, that's $10 billion for each disease – completely out of the question

Sequence from chromosome 7

GAAATAATTAATGTTTTCCTTCCTTCTCCTATTTTGTCCTTTACTTCAATTTATTTATTTATTATTAATATTATTATTTTTTG
AGACGGAGTTTCACTCTTGTTGCCAACCTGGAGTGCAGTGGCGTGATCTCAGCTCACTGCACACTCCGCTTTCC/TGG
TTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGTCACACACCACCACGCCCGGCTAATTTTTG
TATTTTTAGTAGAGTTGGGGTTTCACCATGTTGGCCAGACTGGTCTCGAACTCCTGACCTTGTGATCCGCCAGCCTCT
GCCTCCCAAAGAGCTGGGATTACAGGCGTGAGCCACCGCGCTCGGCCCTTTGCATCAATTTCTACAGCTTGTTTTCTT
TGCCTGGACTTTACAAGTCTTACCTTGTTCTGCCTTCAGATATTTGTGTGGTCTCATTCTGGTGTGCCAGTAGCTAAAA
ATCCATGATTTGCTCTCATCCCACTCCTGTTGTTCATCTCCTCTTATCTGGGGTCAC**A/C**TATCTCTTCGTGATTGCATTC
TGATCCCCAGTACTTAGCATGTGCGTAACAACTCTGCCTCTGCTTTCCCAGGCTGTTGATGGGGTGCTGTTCATGCCT
CAGAAAAATGCATTGTAAGTTAAATTATTAAAGATTTTAAATATAGGAAAAAAGTAAGCAAACATAAGGAACAAAAAG
GAAAGAACATGTATTCTAATCCATTATTTATTATACAATTAAGAAATTTGGAAACTTTAGATTACACTGCTTTTAGAGAT
GGAGATGTAGTAAGTCTTTTACTCTTTACAAAATACATGTGTTAGCAATTTTGGGAAGAATAGTAACTCACCCGAACA
GTGTAATGTGAATATGTCACTTACTAGAGGAAAGAAGGCACTTGAAAAACATCTCTAAACCGTATAAAAACAATTACA
TCATAATGATGAAAACCCAAGGAATTTTTTTAGAAAACATTACCAGGGCTAATAACAAAGTAGAGCCACATGTCATTT
ATCTTCCCTTTGTGTCTGTGTGAGAATTCTAGAGTTATATTTGTACATAGCATGGAAAAATGAGAGGCTAGTTTATCAA
CTAGTTCATTTTTAAAAGTCTAACACATCCTAGGTATAGGTGAACTGTCCTCCTGCCAATGTATTGCACATTTGTGCCC
AGATCCAGCATAGGGTATGTTTGCCATTTACAAACGTTTATGTCTTAAGAGAGGAAATATGAAGAGCAAAACAGTGCA
TGCTGGAGAGAGAAAGCTGATACAAATATAAATGAAACAATAATTGGAAAAATTGAGAAACTACTCATTTTCTAAATT
ACTCATGTATTTTCCTAGAATTTAAGTCTTTTAATTTTTGATAAATCCCAATGTGAGACAAGATAAGTATTAGTGATGGT
ATGAGTAATTAATATCTGTTATATAATATTCATTTTCATAGTGGAAGAAATAAAATAAAGGTTGTGATGATTGTTGATTA
TTTTTTCTAGAGGGGTTGTCAGGGAAAGAAATTGCTTTTTTTCATTCTCTCTTTCCACTAAGAAAGTTCAACTATTAATT
TAGGCACATACAATAATTACTCCATTCTAAAATGCCAAAAAGGTAATTTAAGAGACTTAAAACTGAAAAGTTTAAGATA
GTCACACTGAACTATATTAAAAAATCCACAGGGTGGTTGGAACTAGGCCTTATATTAAAGAGGCTAAAAATTGCAATA
AGACCACAGGCTTTAAATATGGCTTTAAACTGTGAAAGGTGAAACTAGAATGAATAAAATCCTATAAATTTAAATCAA
AAGAAAGAAACAAACT**A/G**AAATTAAAGTTAATATACAAGAATATGGTGGCCTGGATCTAGTGAACATATAGTAAAGA
TAAAACAGAATATTTCTGAAAAATCCTGGAAAATCTTTTGGGCTAACCTGAAAACAGTATATTTGAAACTATTTTTAAA

Three single nucleotide polymorphisms (SNPs) are present

Sequence from chromosome 7

GAAATAATTAATGTTTTCCTTCCTTCTCCTATTTTGTCCTTTACTTCAATTTATTTATTTATTATTAATATTATTATTTTTTG
AGACGGAGTTTCACTCTTGTTGCCAACCTGGAGTGCAGTGGCGTGATCTCAGCTCACTGCACACTCCGCTTTCC/TGG
TTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGTCACACACCACCACGCCCGGCTAATTTTTG
TATTTTTAGTAGAGTTGGGGTTTCACCATGTTGGCCAGACTGGTCTCGAACTCCTGACCTTGTGATCCGCCAGCCTCT
GCCTCCCAAAGAGCTGGGATTACAGGCGTGAGCCACCGCGCTCGGCCCTTTGCATCAATTTCTACAGCTTGTTTTCTT
TGCCTGGACTTTACAAGTCTTACCTTGTTCTGCCTTCAGATATTTGTGTGGTCTCATTCTGGTGTGCCAGTAGCTAAAA
ATCCATGATTTGCTCTCATCCCACTCCTGTTGTTCATCTCCTCTTATCTGGGGTCAC**A/C**TATCTCTTCGTGATTGCATTC
TGATCCCCAGTACTTAGCATGTGCGTAACAACTCTGCCTCTGCTTTCCCAGGCTGTTGATGGGGTGCTGTTCATGCCT
CAGAAAAATGCATTGTAAGTTAAATTATTAAAGATTTTAAATATAGGAAAAAAGTAAGCAAACATAAGGAACAAAAAG
GAAAGAACATGTATTCTAATCCATTATTTATTATACAATTAAGAAATTTGGAAACTTTAGATTACACTGCTTTTAGAGAT
GGAGATGTAGTAAGTCTTTTACTCTTTACAAAATACATGTGTTAGCAATTTTGGGAAGAATAGTAACTCACCCGAACA
GTGTAATGTGAATATGTCACTTACTAGAGGAAAGAAGGCACTTGAAAAACATCTCTAAACCGTATAAAAACAATTACA
TCATAATGATGAAAACCCAAGGAATTTTTTTAGAAAACATTACCAGGGCTAATAACAAAGTAGAGCCACATGTCATTT
ATCTTCCCTTTGTGTCTGTGTGAGAATTCTAGAGTTATATTTGTACATAGCATGGAAAAATGAGAGGCTAGTTTATCAA
CTAGTTCATTTTTAAAAGTCTAACACATCCTAGGTATAGGTGAACTGTCCTCCTGCCAATGTATTGCACATTTGTGCCC
AGATCCAGCATAGGGTATGTTTGCCATTTACAAACGTTTATGTCTTAAGAGAGGAAATATGAAGAGCAAAACAGTGCA
TGCTGGAGAGAGAAAGCTGATACAAATATAAATGAAACAATAATTGGAAAAATTGAGAAACTACTCATTTTCTAAATT
ACTCATGTATTTTCCTAGAATTTAAGTCTTTTAATTTTTGATAAATCCCAATGTGAGACAAGATAAGTATTAGTGATGGT
ATGAGTAATTAATATCTGTTATATAATATTCATTTTCATAGTGGAAGAAATAAAATAAAGGTTGTGATGATTGTTGATTA
TTTTTTCTAGAGGGGTTGTCAGGGAAAGAAATTGCTTTTTTTCATTCTCTCTTTCCACTAAGAAAGTTCAACTATTAATT
TAGGCACATACAATAATTACTCCATTCTAAAATGCCAAAAAGGTAATTTAAGAGACTTAAAACTGAAAAGTTTAAGATA
GTCACACTGAACTATATTAAAAAATCCACAGGGTGGTTGGAACTAGGCCTTATATTAAAGAGGCTAAAAATTGCAATA
AGACCACAGGCTTTAAATATGGCTTTAAACTGTGAAAGGTGAAACTAGAATGAATAAAATCCTATAAATTTAAATCAA
AAGAAAGAAACAAACT**A/G**AAATTAAAGTTAATATACAAGAATATGGTGGCCTGGATCTAGTGAACATATAGTAAAGA
TAAAACAGAATATTTCTGAAAAATCCTGGAAAATCTTTTGGGCTAACCTGAAAACAGTATATTTGAAACTATTTTTAAA

Are the SNPs correlated with their neighbors?

These three SNPs could theoretically occur in 8 different haplotypes

…C…A…A…

…C…A…G…

…C…C…A…

…C…C…G…

…T…A…A…

…T…A…G…

…T…C…A…

…T…C…G…

But in practice, only two are observed

…C…A…A…

…C…A…G…

…C…C…A…

…C…C…G…

…T…A…A…

…T…A…G…

…T…C…A…

…T…C…G…

27 October 2005 | www.nature.com/nature

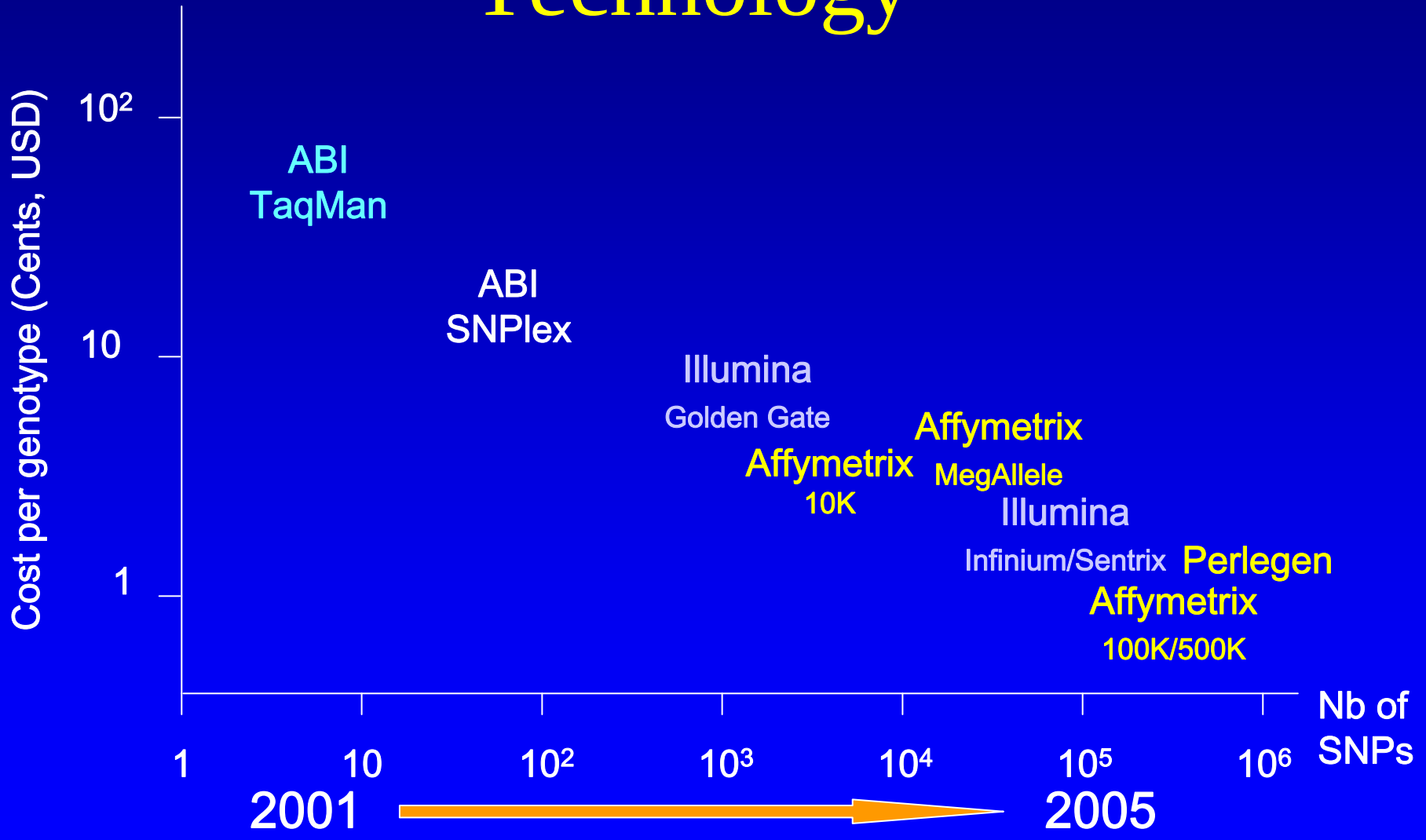
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

nature

THE HAPMAP PROJECT

Chapter and verse on human genetic variation

Progress in Genotyping Technology

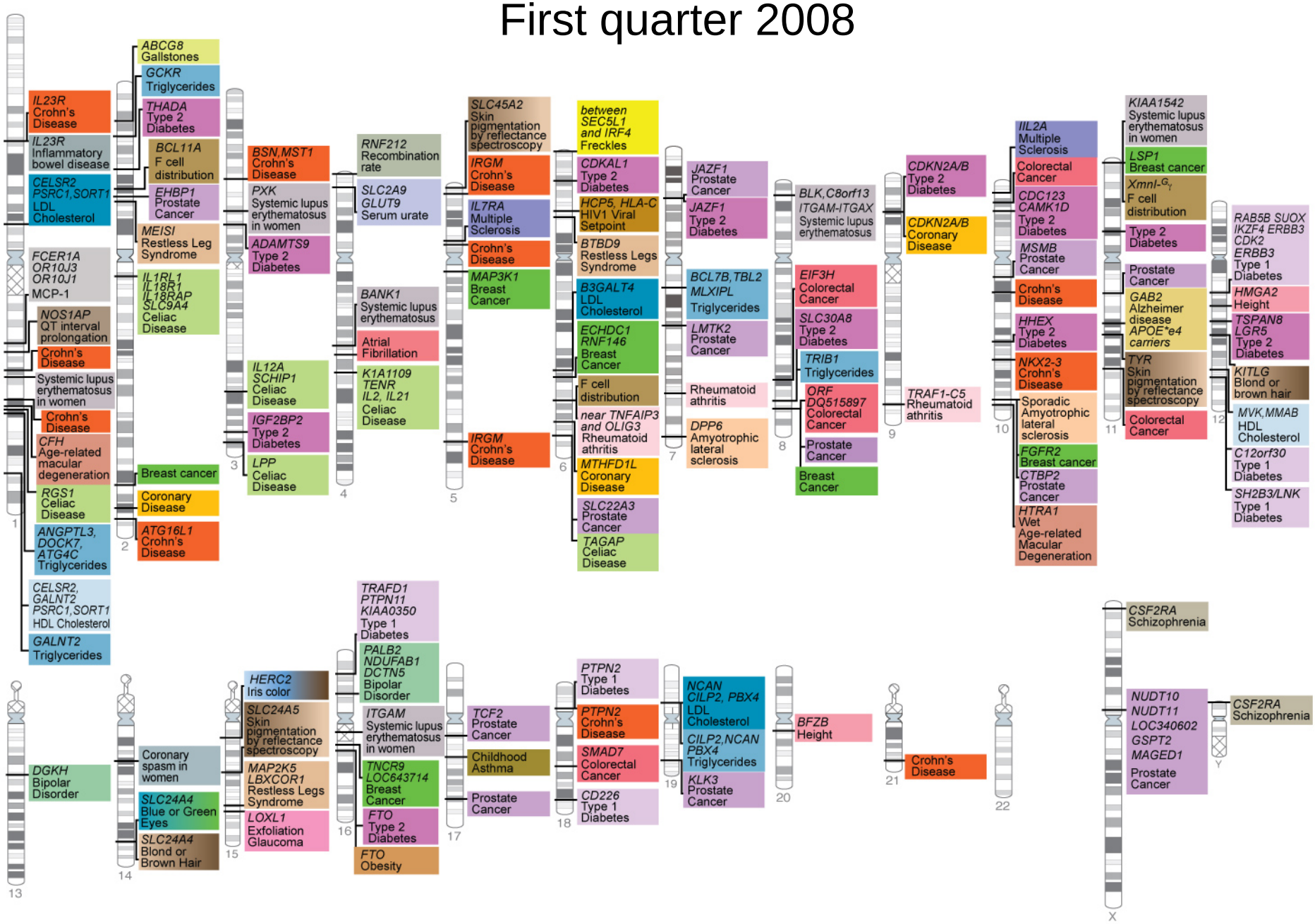


Courtesy S. Chanock, NCI

Genome Wide Association Approach to Common Disease: The View from 2008

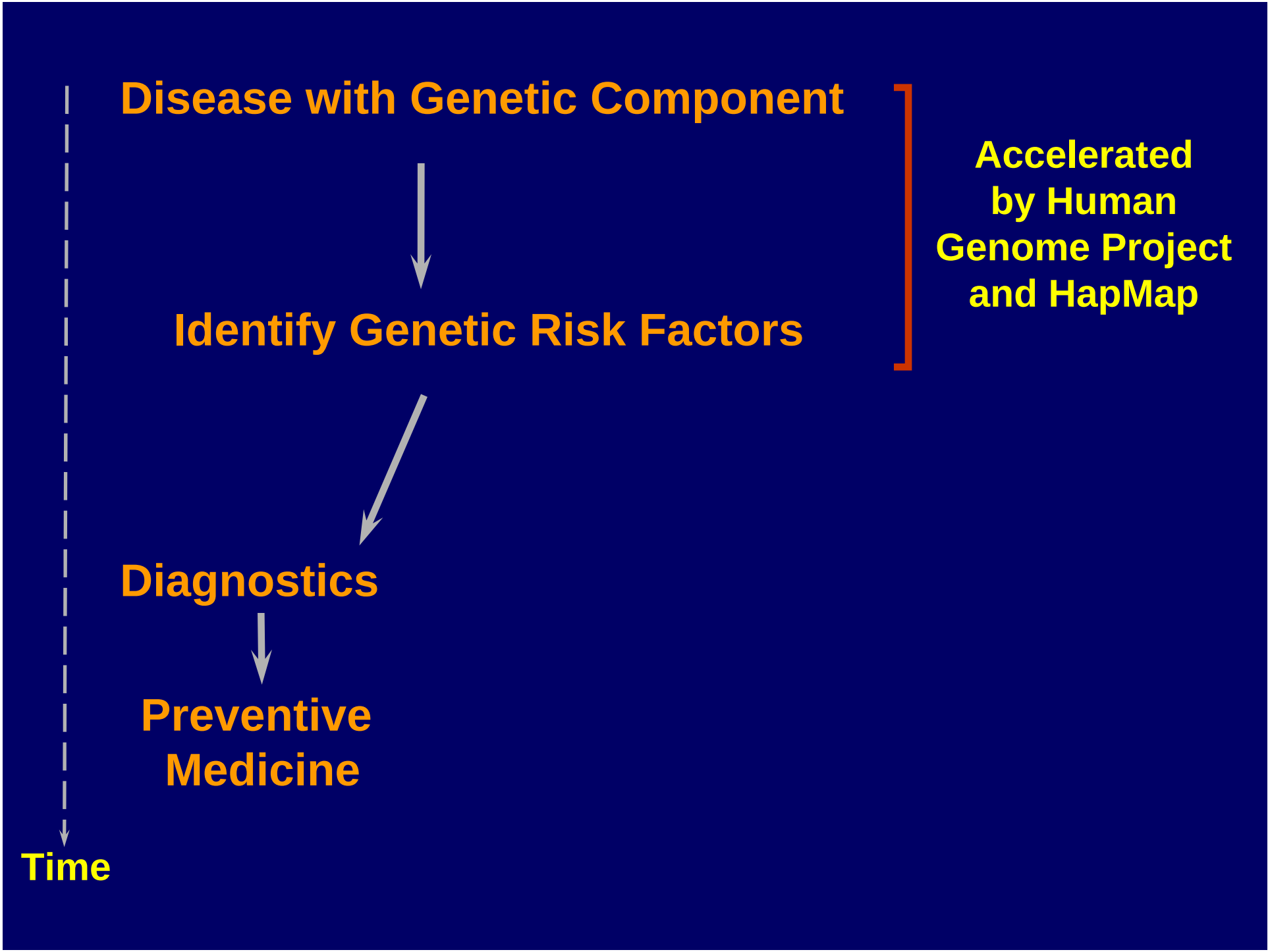
- Identify an optimum set of 300,000 tag SNPs
- Collect 1000 cases and 1000 controls
- Genotype all DNAs for all SNPs
- That adds up to 600 million genotypes
- Genotyping just dropped to $0.0010, so that's $600,000 for each disease

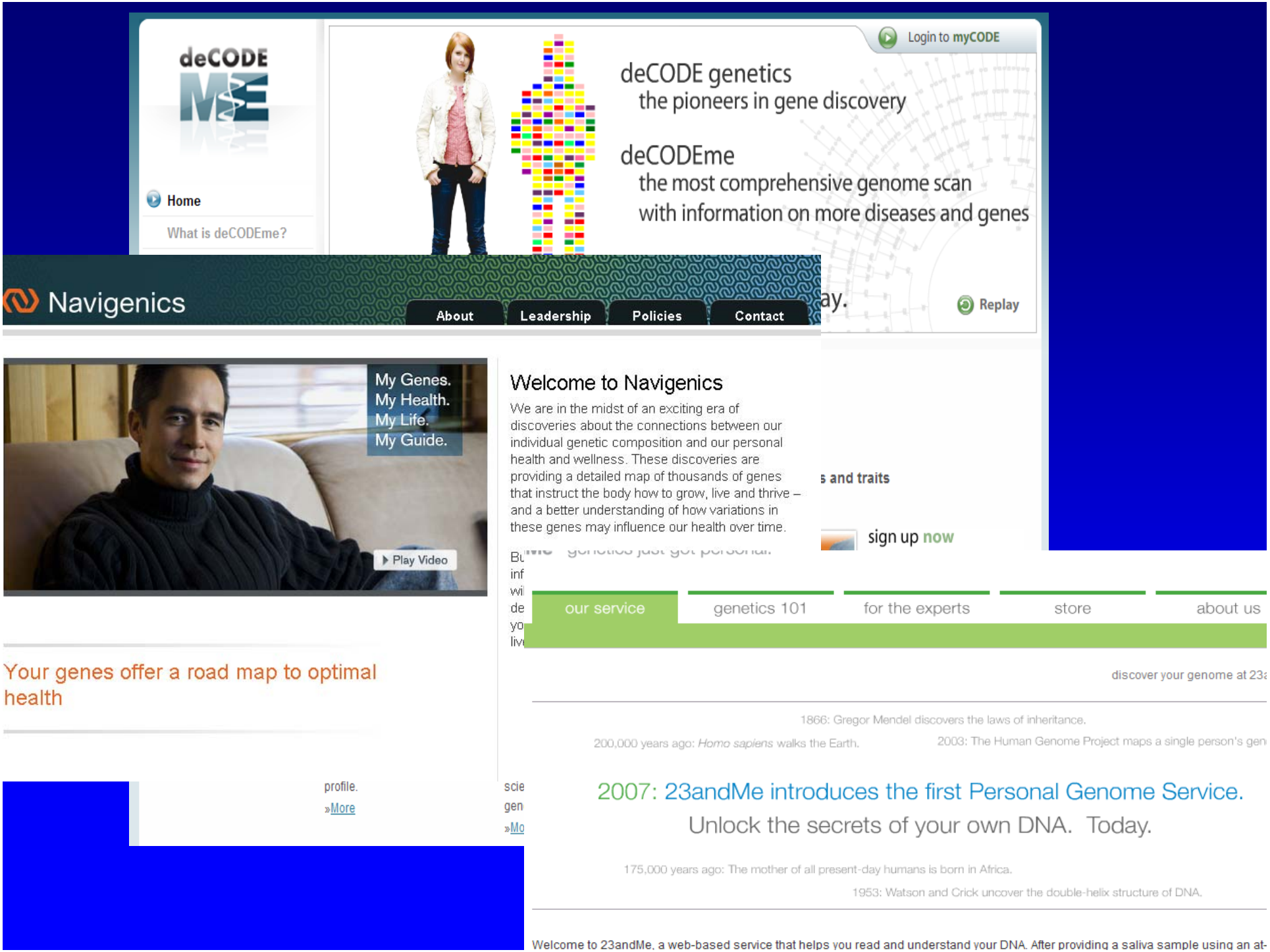
First quarter 2008

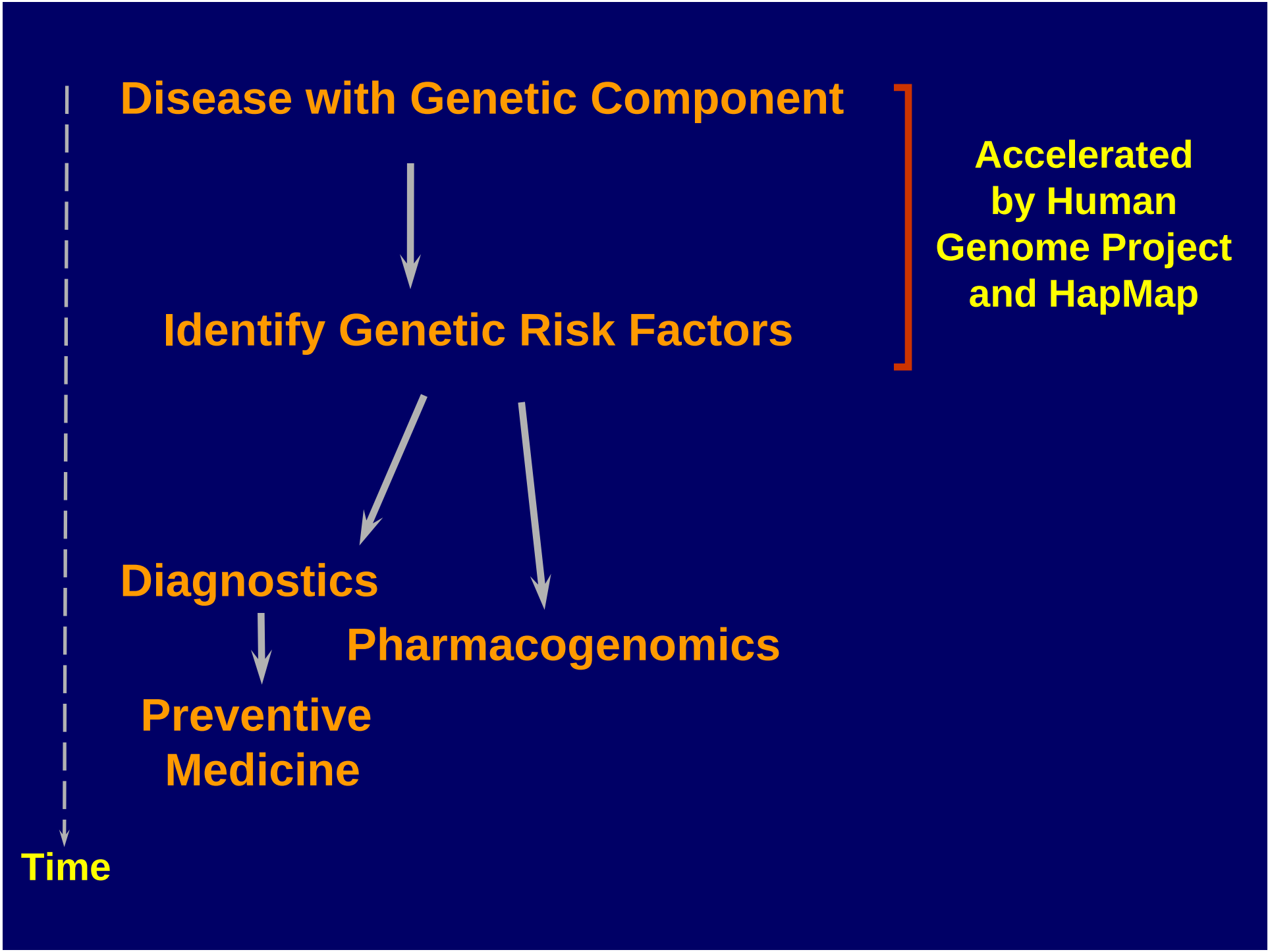


Manolio, Brooks, Collins, J. Clin. Invest., May 2008

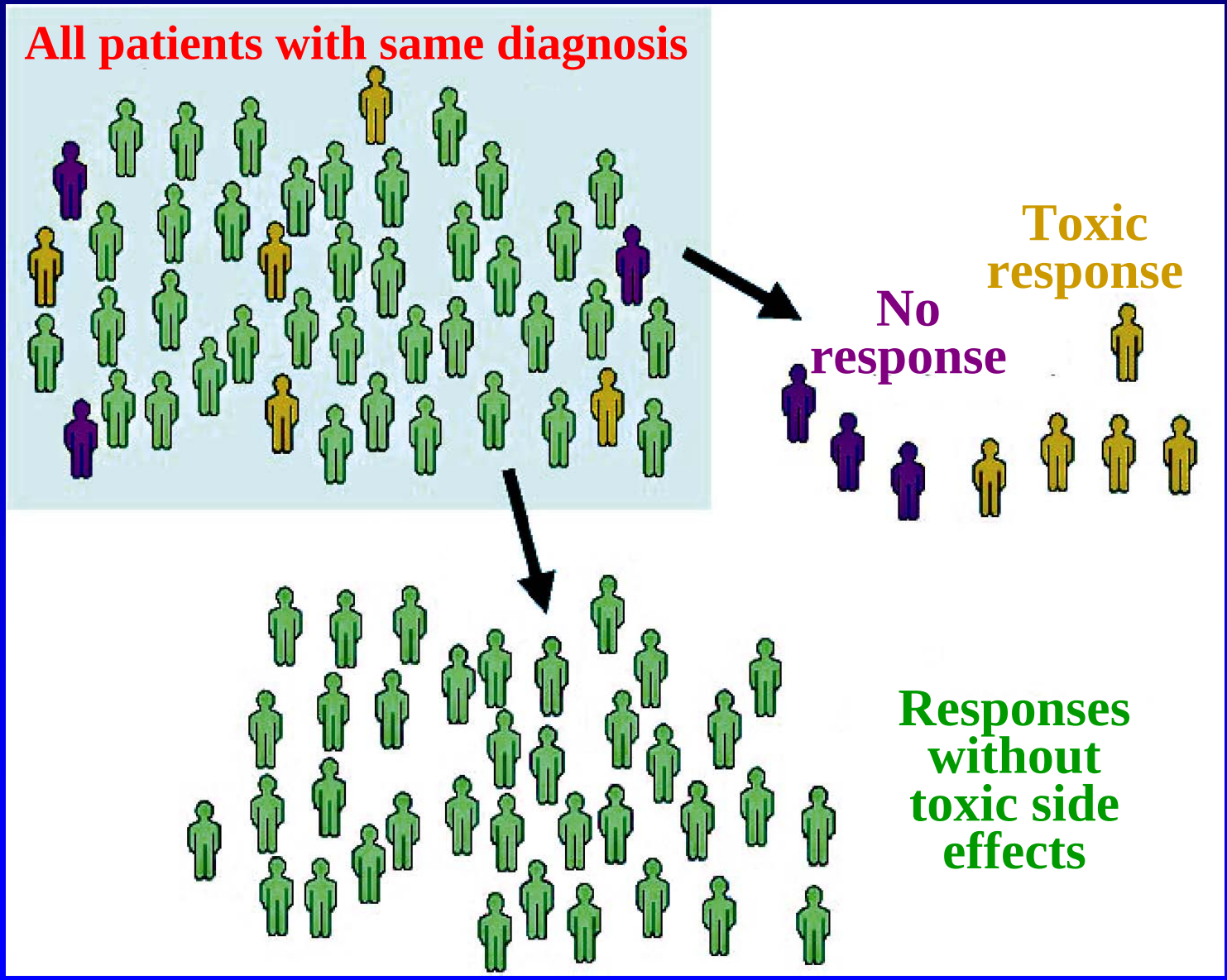
What will be the impact of genomics on the practice of medicine?


Disease with Genetic Component
Identify Genetic Risk Factors
Diagnostics
Preventive Medicine
Time
Accelerated by Human Genome Project and HapMap

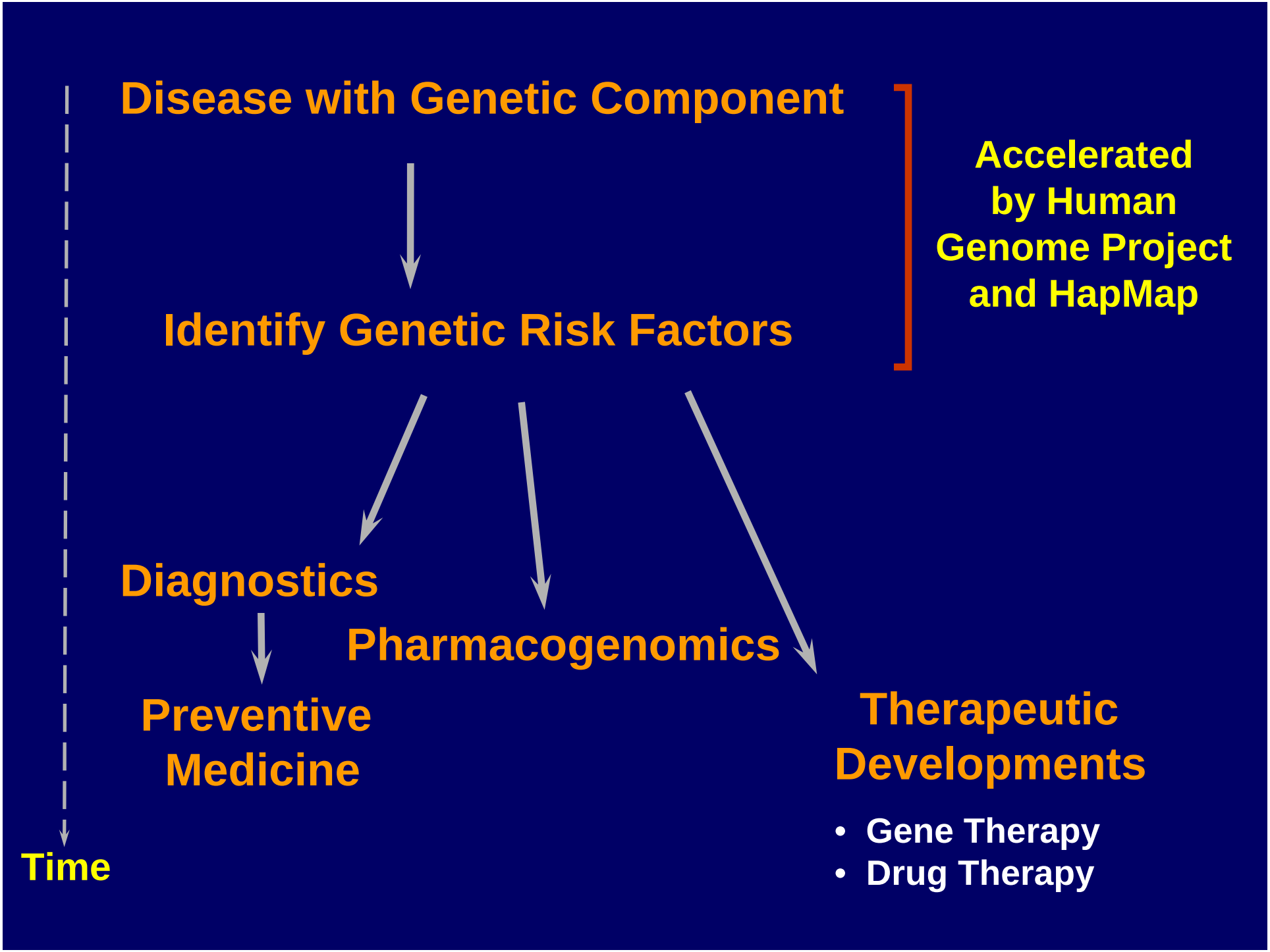

deCODE
Home
What is deCODEme?
Login to myCODE
deCODE genetics
the pioneers in gene discovery
deCODEme
the most comprehensive genome scan
with information on more diseases and genes
Replay
Navigenics
About
Leadership
Policies
Contact
My Genes.
My Health.
My Life.
My Guide.
Play Video
Welcome to Navigenics
We are in the midst of an exciting era of discoveries about the connections between our individual genetic composition and our personal health and wellness. These discoveries are providing a detailed map of thousands of genes that instruct the body how to grow, live and thrive – and a better understanding of how variations in these genes may influence our health over time.
s and traits
sign up now
our service
genetics 101
for the experts
store
about us
Your genes offer a road map to optimal health
discover your genome at 23
1866: Gregor Mendel discovers the laws of inheritance.
200,000 years ago: Homo sapiens walks the Earth.
2003: The Human Genome Project maps a single person's gen
profile.
»More
2007: 23andMe introduces the first Personal Genome Service.
Unlock the secrets of your own DNA. Today.
175,000 years ago: The mother of all present-day humans is born in Africa.
1953: Watson and Crick uncover the double-helix structure of DNA.
Welcome to 23andMe, a web-based service that helps you read and understand your DNA. After providing a saliva sample using an at-

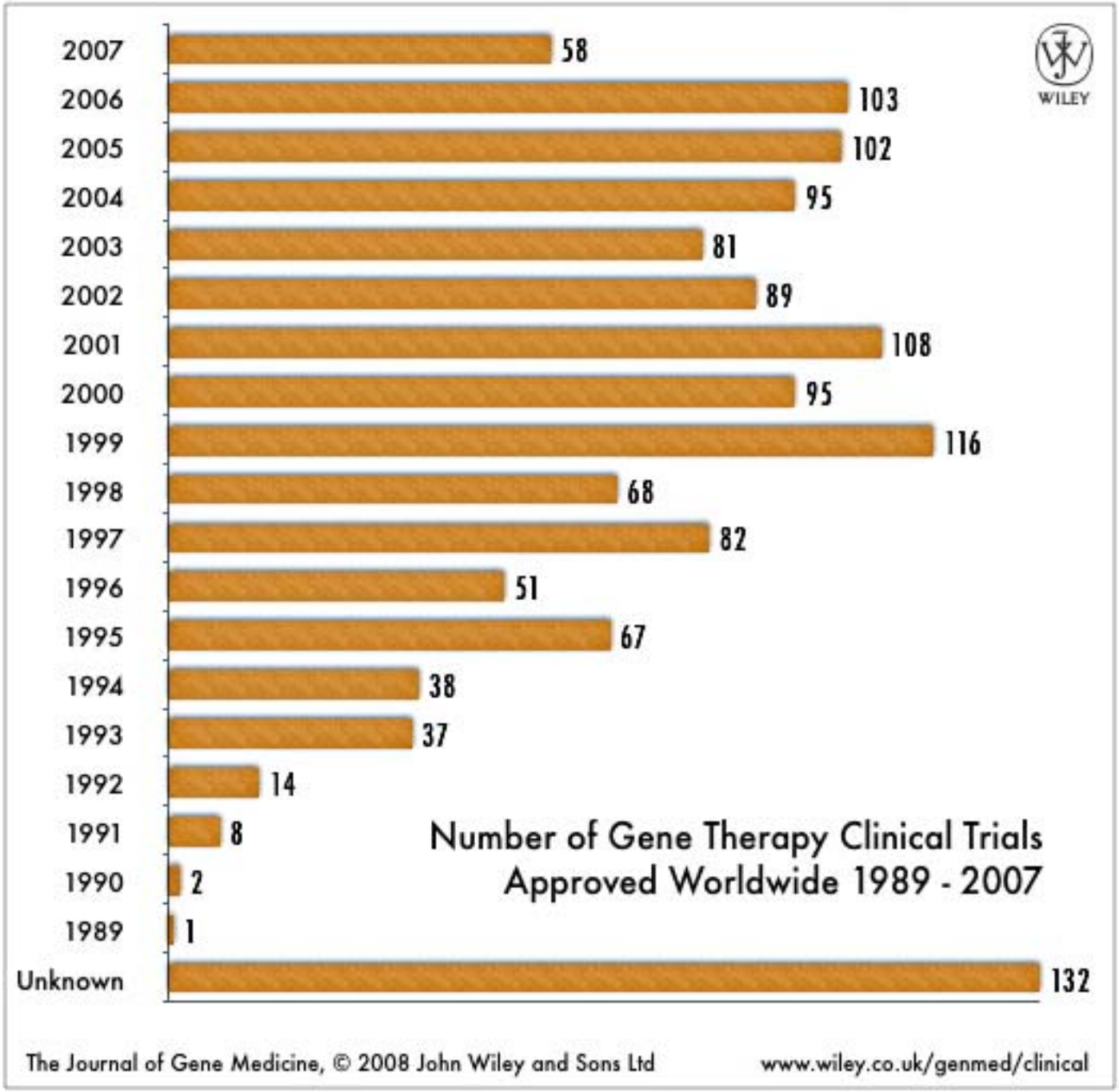

Disease with Genetic Component
Identify Genetic Risk Factors
Diagnostics
Pharmacogenomics
Preventive Medicine
Time
Accelerated by Human Genome Project and HapMap

Pharmacogenomics: Unlocking the Human Genome for Better Drug Therapy



McLeod & Evans, *Ann Rev Pharmacol Toxicol* 41:101-21, 2001

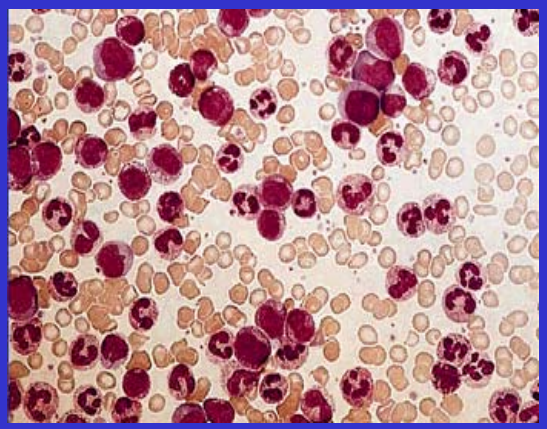

Disease with Genetic Component
Identify Genetic Risk Factors
Accelerated by Human Genome Project and HapMap
Diagnostics
Preventive Medicine
Pharmacogenomics
Therapeutic Developments
• Gene Therapy
• Drug Therapy
Time


2007
58
2006
103
2005
102
2004
95
2003
81
2002
89
2001
108
2000
95
1999
116
1998
68
1997
82
1996
51
1995
67
1994
38
1993
37
1992
14
1991
8
1990
2
1989
1
Unknown
132
WILEY
Number of Gene Therapy Clinical Trials
Approved Worldwide 1989 - 2007
The Journal of Gene Medicine, © 2008 John Wiley and Sons Ltd
www.wiley.co.uk/genmed/clinical

Gleevec™ – Specifically Targets an Abnormal Protein, Blocking Its Ability to Cause Chronic Myeloid Leukemia

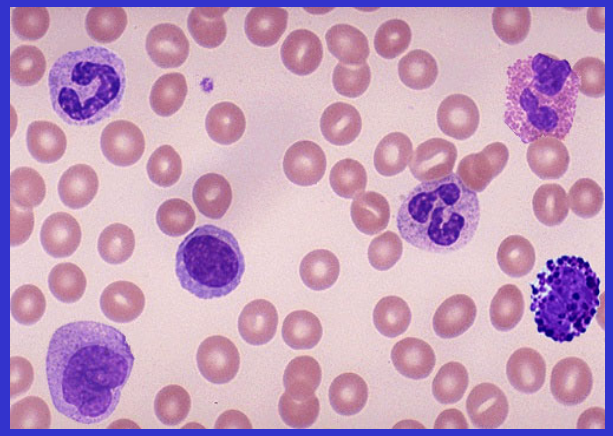
Leukemia arises in mutated blood cell

Leukemia Protein



Leukemia

Leukemia Protein



Remission


GTACT
CAAGT
CTA
TAC
CCGA
TAGG
ATCC
TGG
ATCC
S. Harris