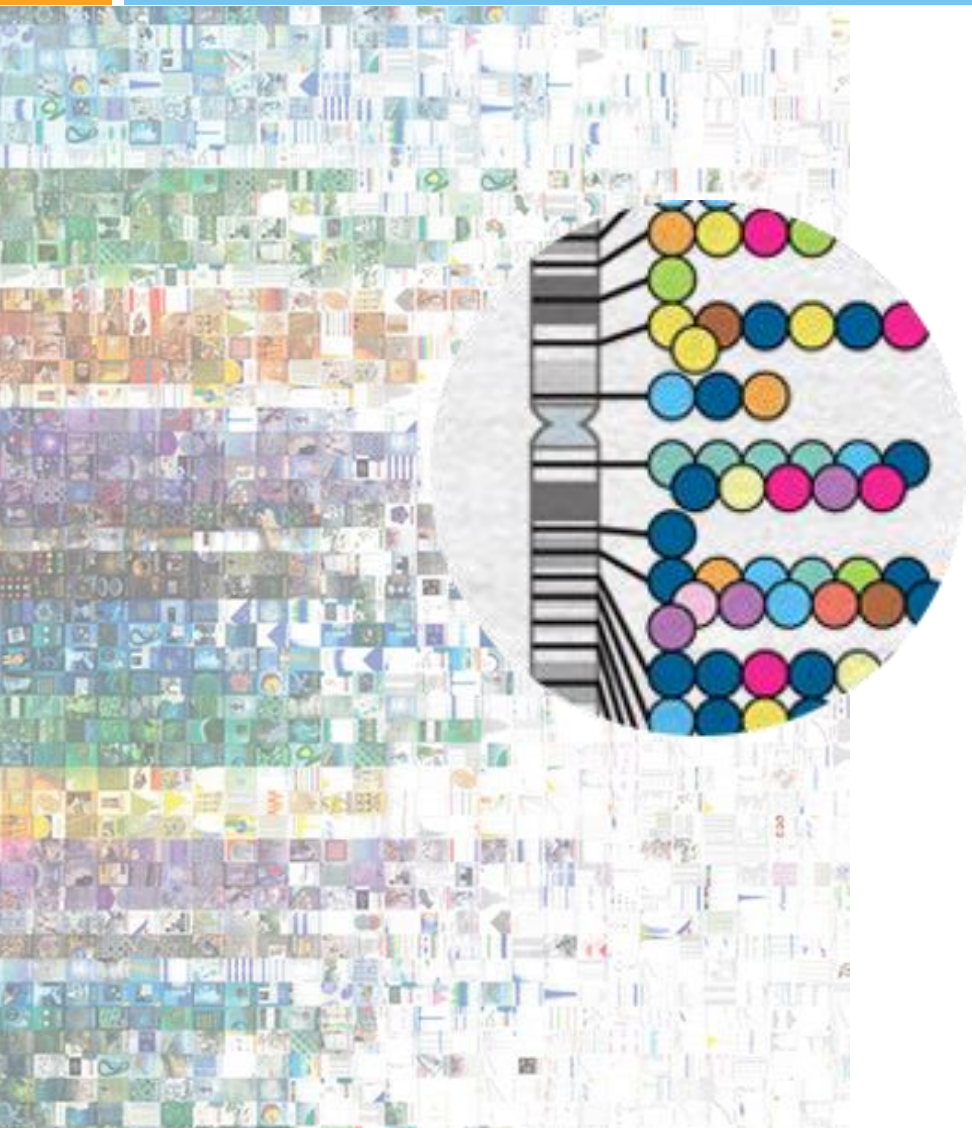
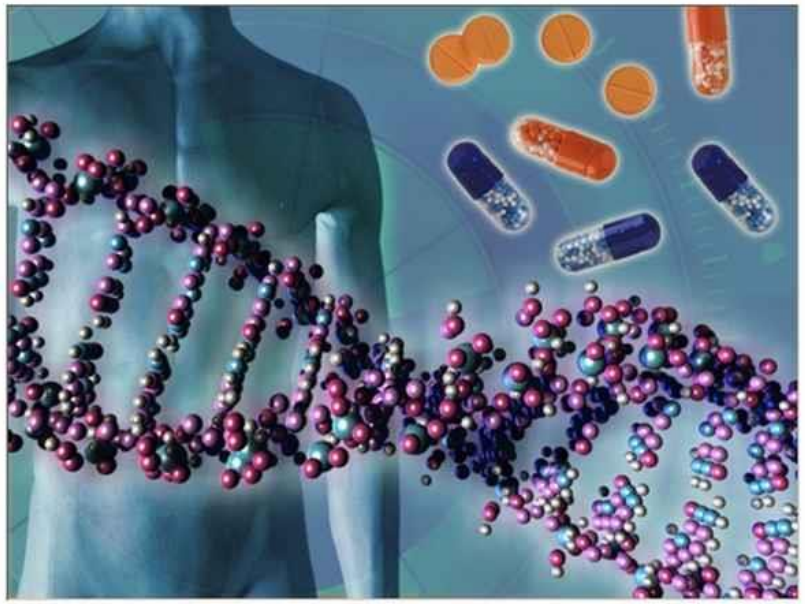
Post Genomic Medicine

Genome-wide association study

Bioinformatics at SNPs level





UMAPORN YORDPRATUM, PH.D.
Department of Microbiology
Faculty of Medicine
Khon Kaen University

Contents

Human genome project

Allele

Single Nucleotide Polymorphism (SNP)

Genetic linkage map

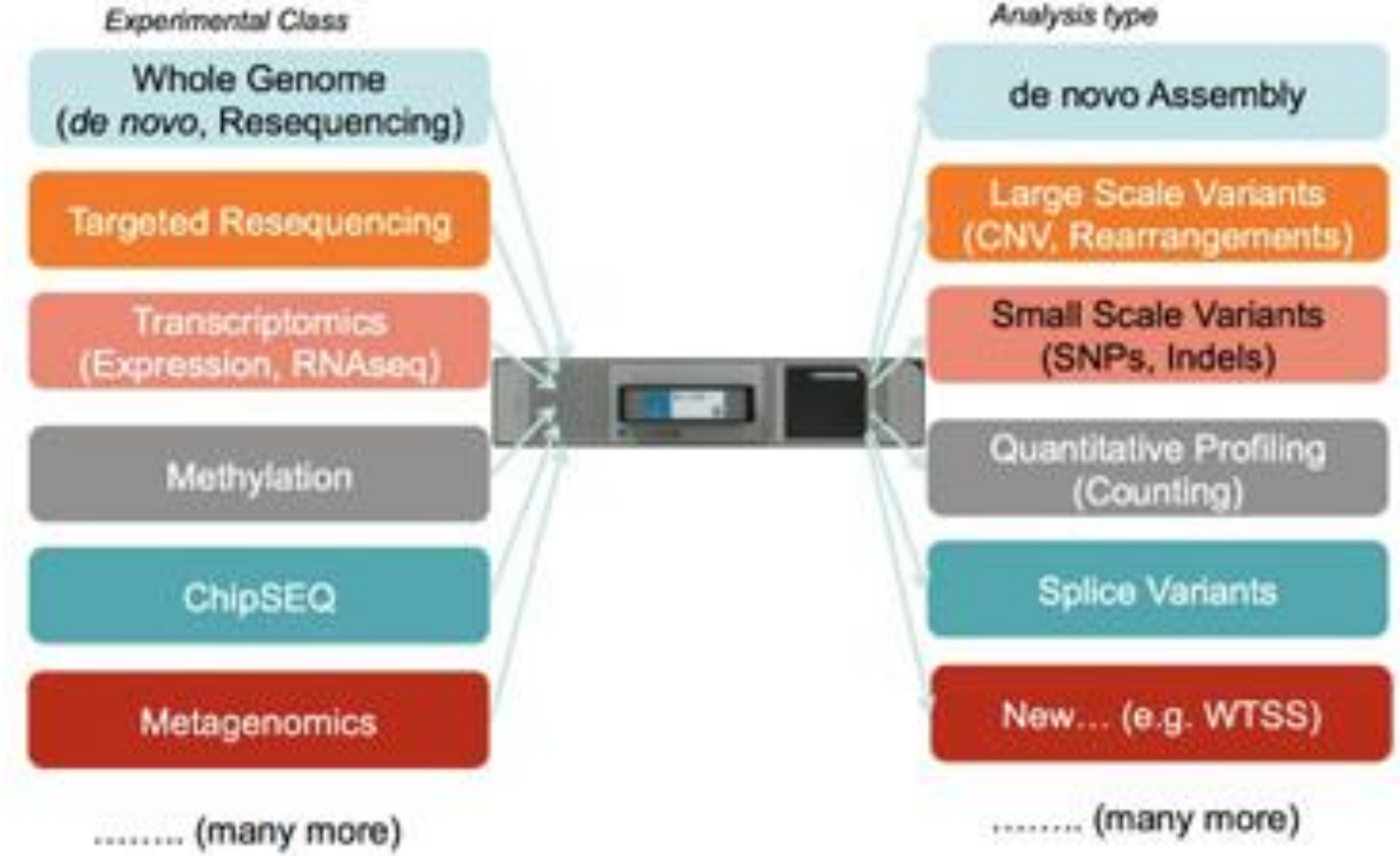
Genome-Wide Association Study **

Personal whole genome sequencing

DNA sequencing technologies

- **Shotgun**
- **Illumina**
- **Pyrosequencing**
- **Ion Torrent**

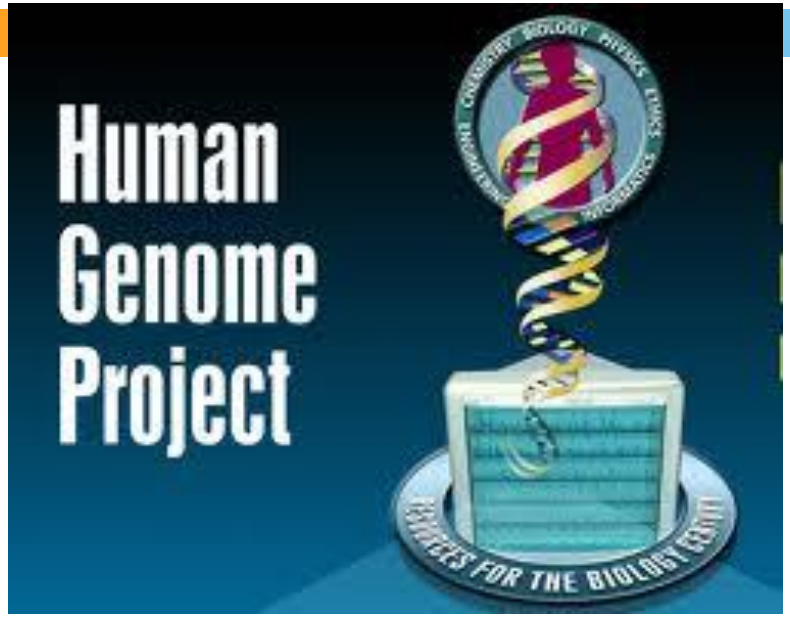
Sequencing Applications



The human genome project

The Human Genome Project (HGP) is an international scientific research project goals:
-determining the sequence of chemical base pairs which make up DNA,
-identifying and mapping the **approximately 20,000–25,000 genes** of the human genome from both a physical and functional standpoint.

The project began in **October 1990** and was initially headed by Aristides Patrinos, head of the Office of Biological and Environmental Research in the U.S. Department of Energy's Office of Science. Francis Collins directed the US National Institutes of Health (NIH) National Human Genome Research Institute efforts. **A working draft of the genome was announced in 2000** and a **complete one in 2003**, with further, more detailed analysis still being published.

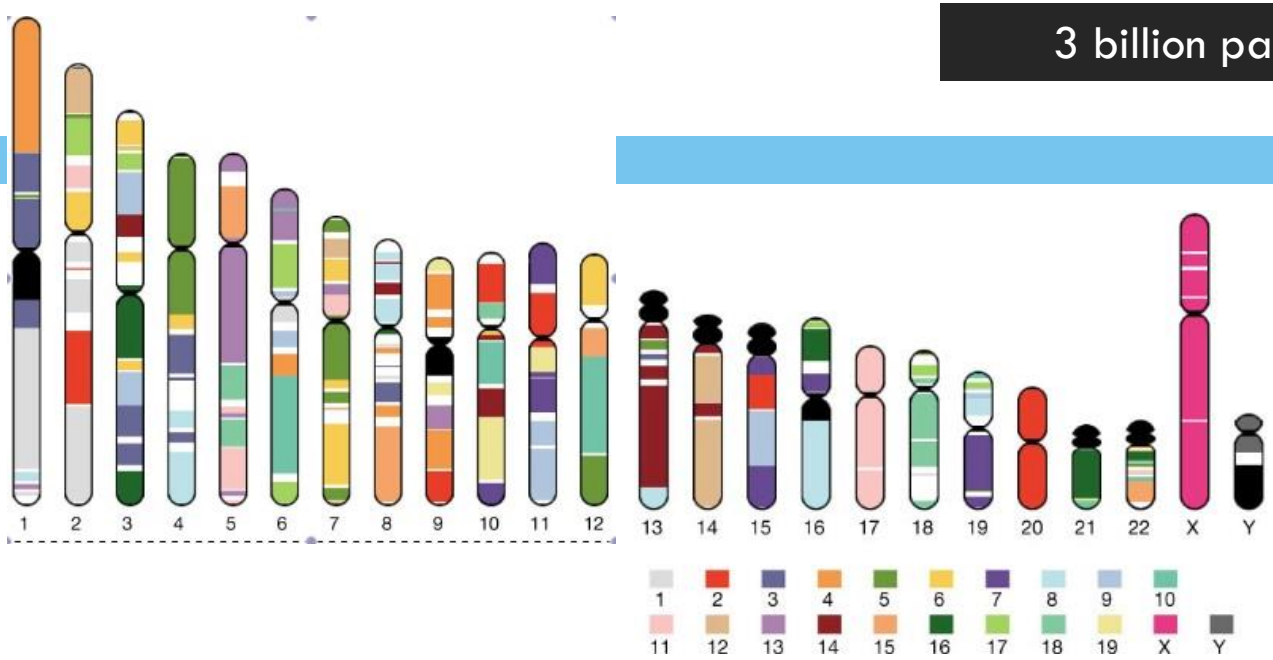

Human
Genome
Project

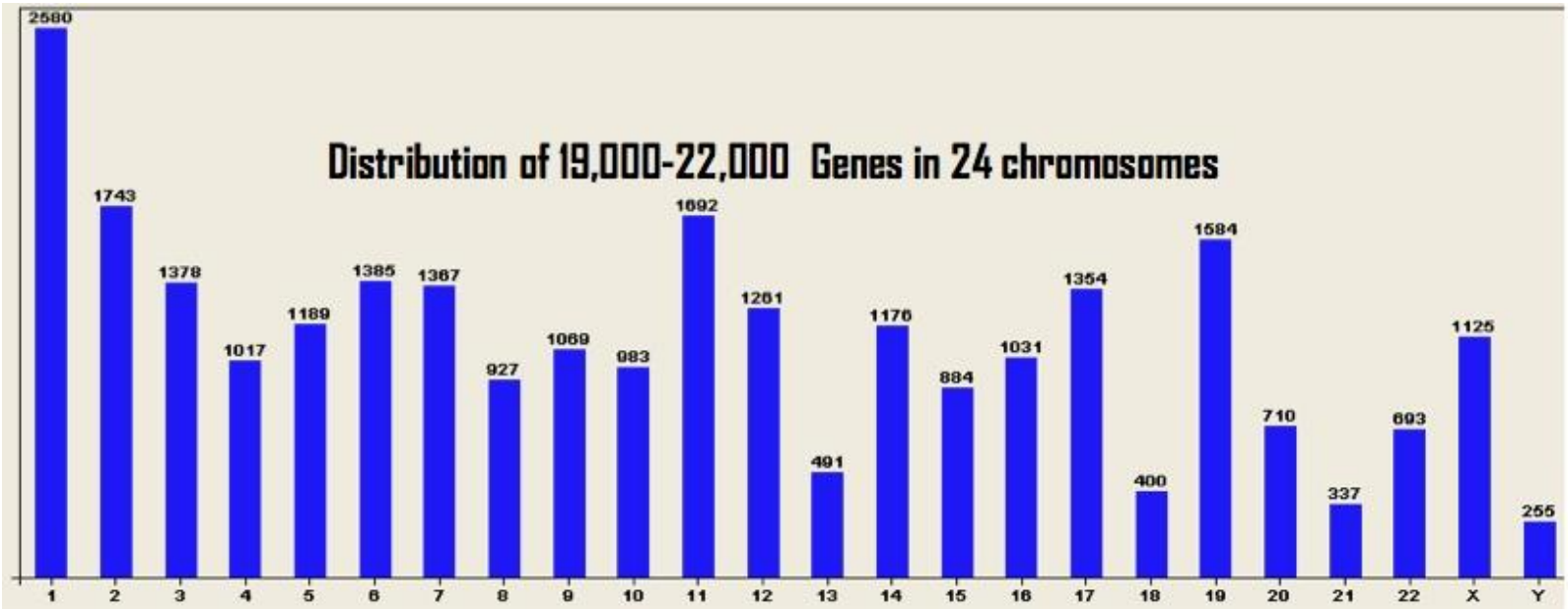

Science
THE
HUMAN
GENOME


Finished Human Genome
Human Genome Project
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
Wellcome Trust Sanger Institute

Human genome (15 years) Project (1988-2003)

3 billion pairs of bases





Allele

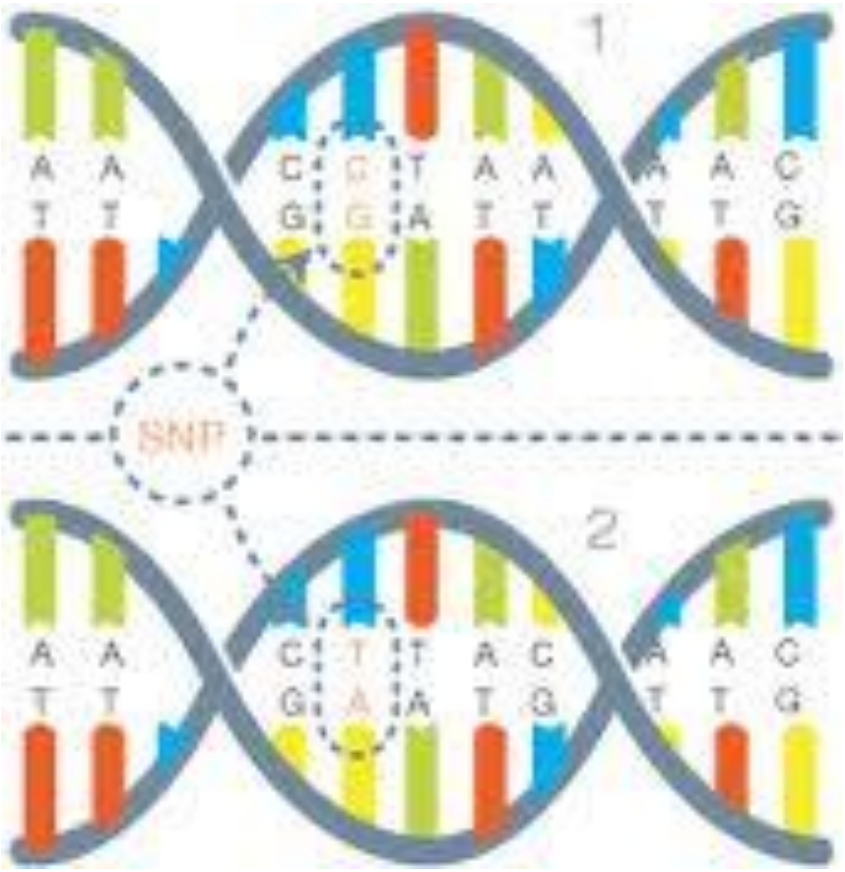
Hair gene

One member of a pair (or any of the series) of genes occupying a specific spot on a chromosome (called locus) that controls the same trait.

Example: a pair of alleles controlling the same trait, i.e. eye color: one allele codes for blue eyes, another allele for brown eyes.

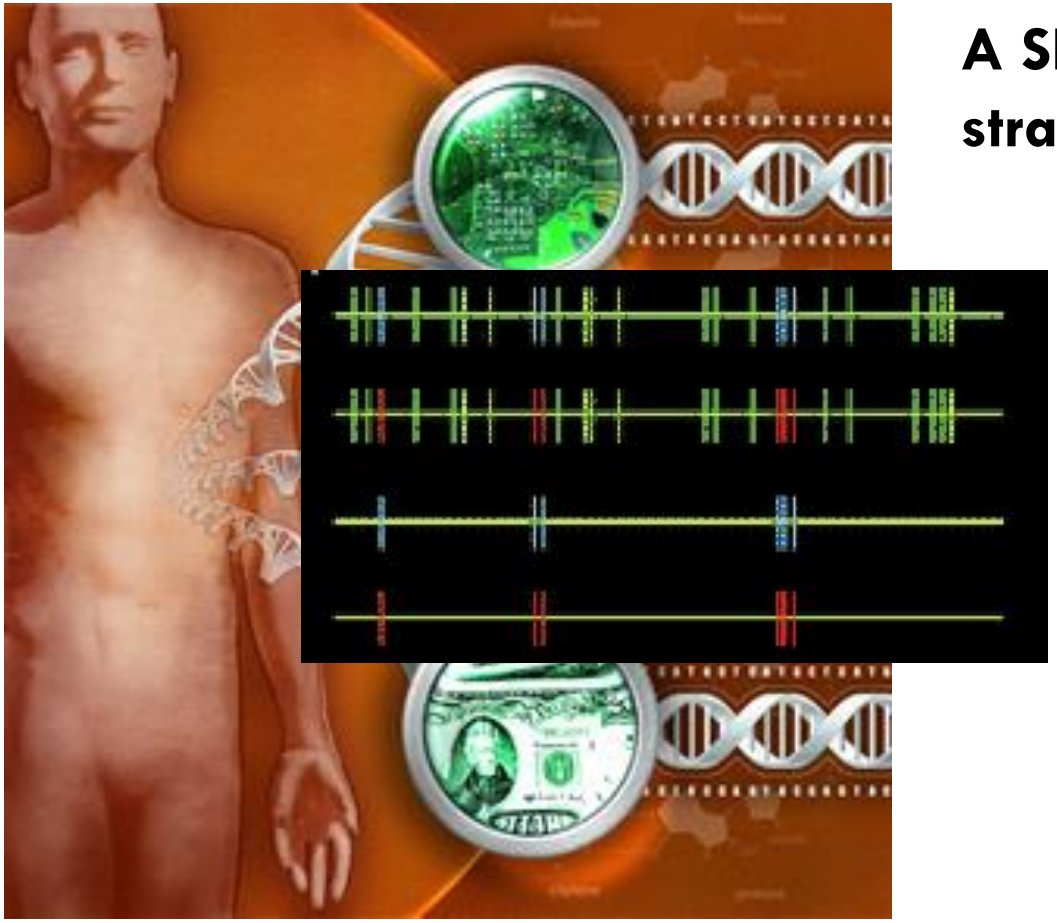
Single Nucleotide Polymorphism (SNP)

A single-nucleotide variation in a genetic sequence; a common form of variation in the human genome



C -> T

Human genome has ~3 million SNPs distributed randomly



A SNP profile can be used for stratify patient

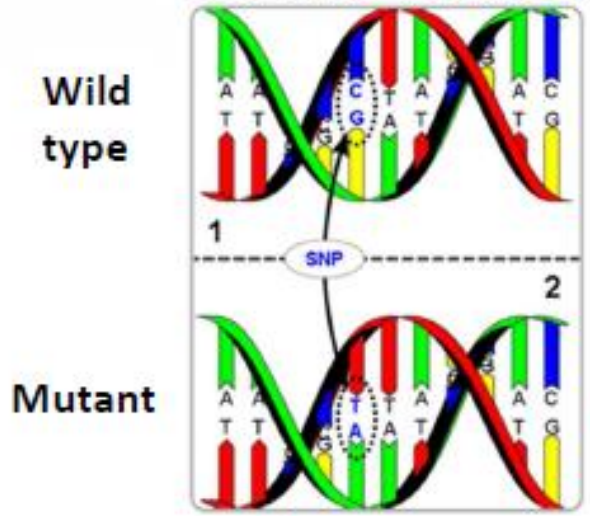
Drug treatment worked

Drug treatment didn't work

SNPs predictive of efficacy

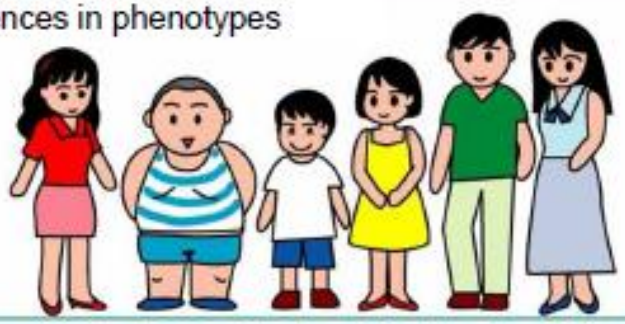
SNPs predictive of NO efficacy

Genetic variations



Making differences in quality and/or quality of gene products (proteins, functional RNAs)

- Differences in phenotypes



- Differences in risk of various diseases

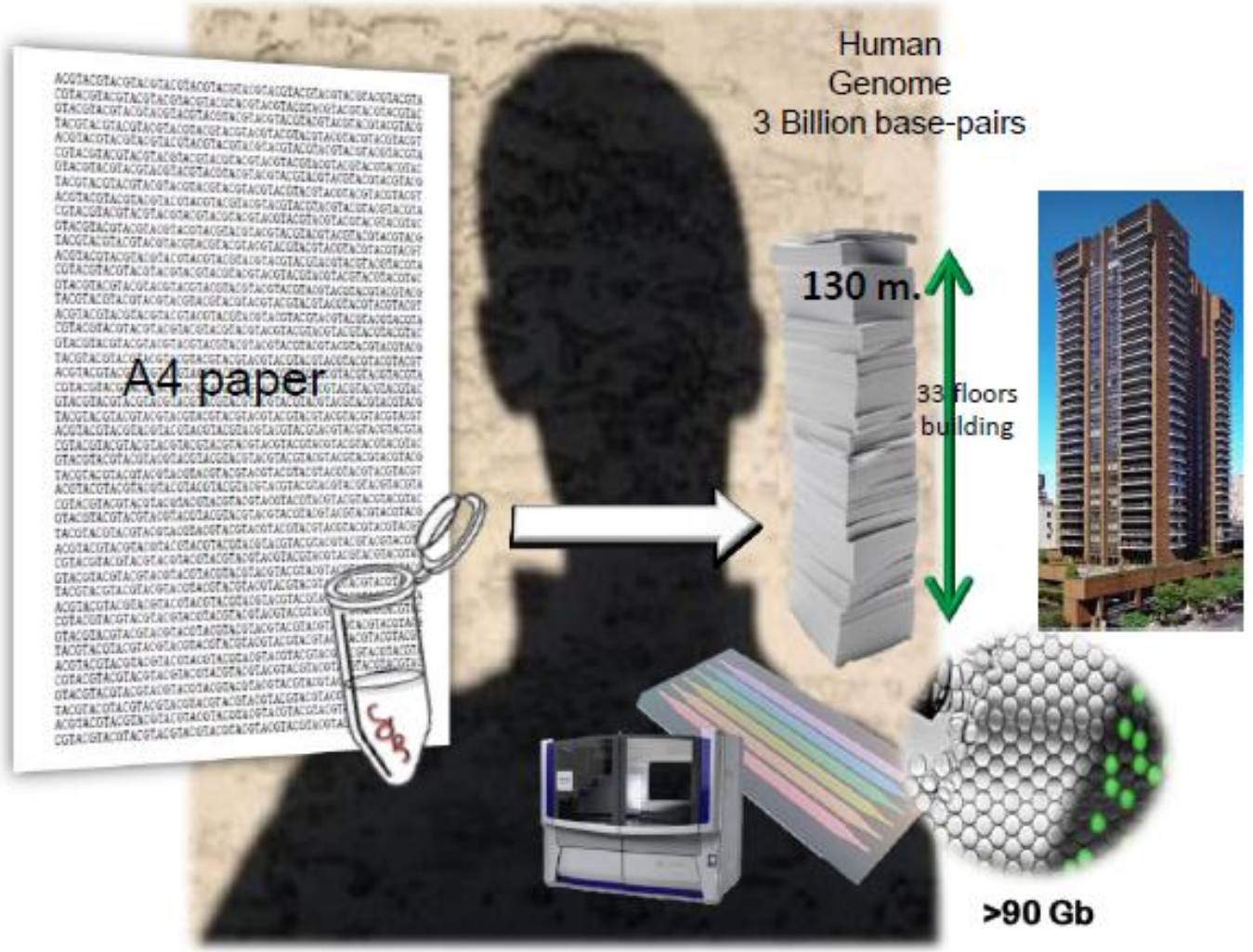
- Differences in responses to treatment

What are we looking for ?

Search for predictive marker associated with disease/ADR

What do we gain ?

Three billion possibilities, how can we do it?



Bioinformatics is the answer. Leading to 3 technologies from past, present and future

Past:

“ Genetic Linkage Map ”

Present:

“ Genome Wide Association Study ”

Future:

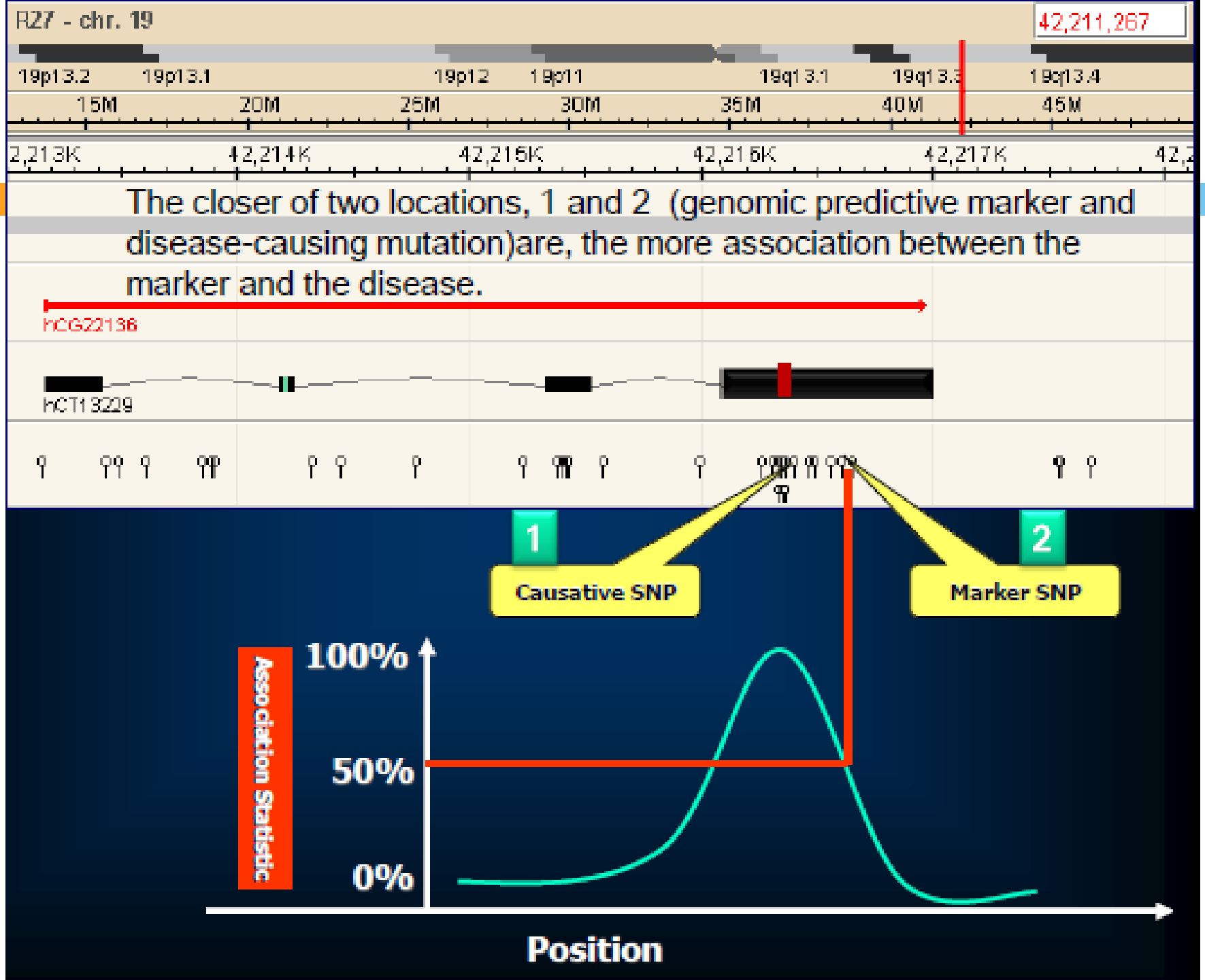
“ Personal (whole) genome sequencing ”

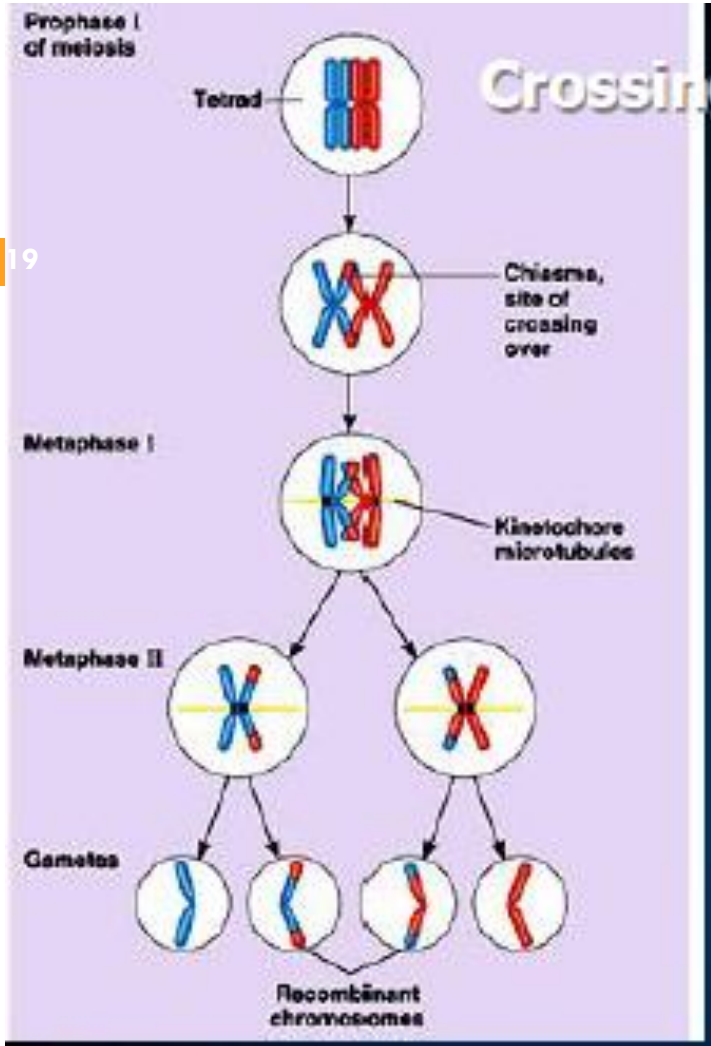
Genetic Linkage Map

Genetic linkage is the tendency of genes that are located proximal to each other on the chromosome to be inherited together during meiosis.

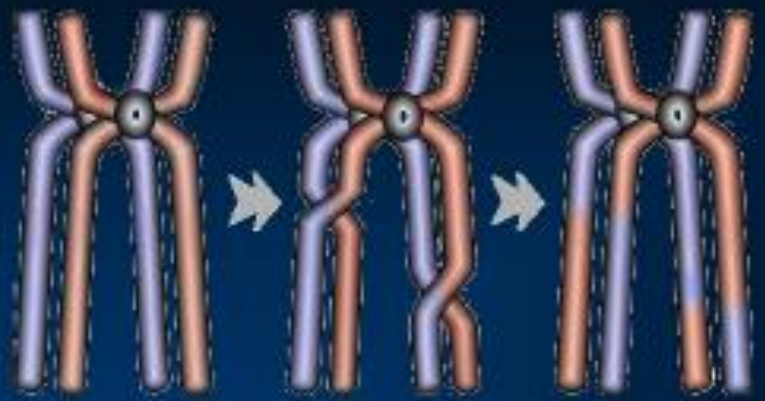
Genes whose loci are nearer to each other are less likely to be separated onto different chromatids during chromosomal crossover, and therefor said to be genetically linked.

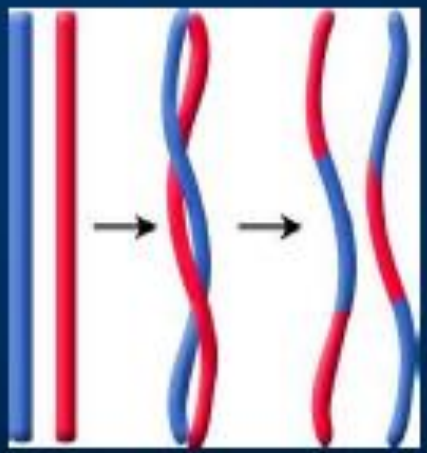
The closer of two locations, 1 and 2 (genomic predictive marker and disease-causing mutation)are, the more association between the marker and the disease.





Crossing Over of Chromosomes in Meiosis





A crossover occurs between two non-sister chromatids

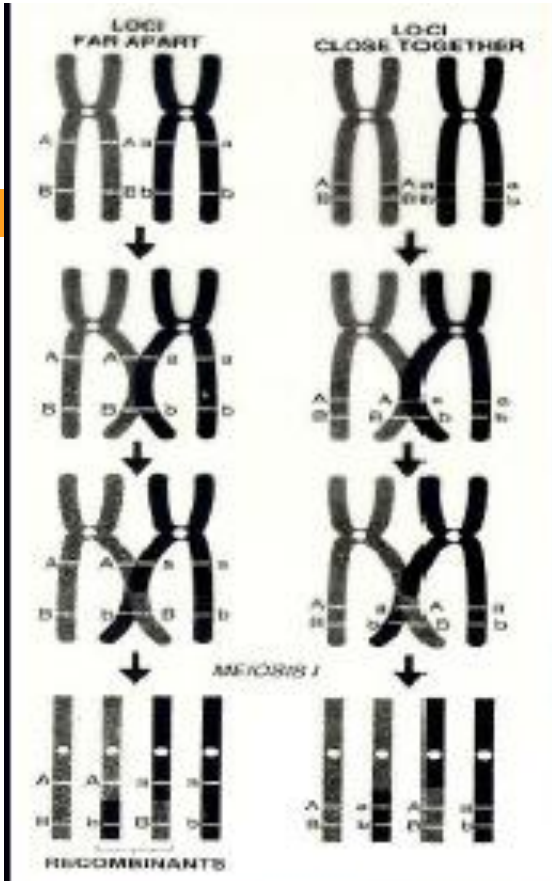
Haplotype:

A set of closely linked genetic markers present on one chromosome which tend to be inherited together (not easily separable by recombination). Some haplotypes may be in linkage disequilibrium.

Hapotype

A set of DNA variations, or polymorphisms, that tend to be inherited together.

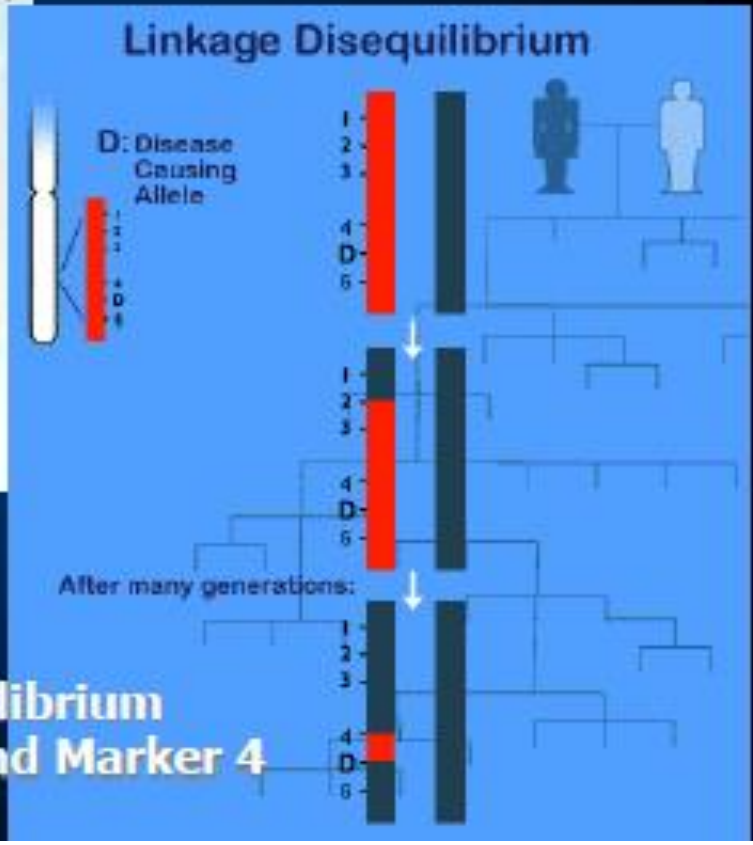
A haplotype can refer to a combination of alleles or to a set of single-nucleotide polymorphisms found on the same chromosome.



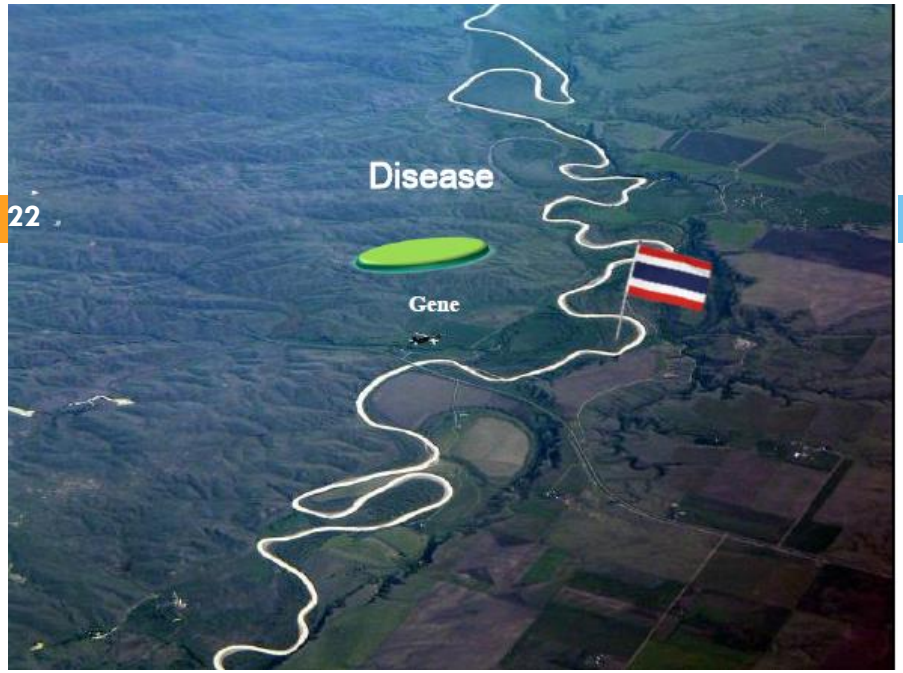
The closer together two gene locations are, the more frequently they are inherited together

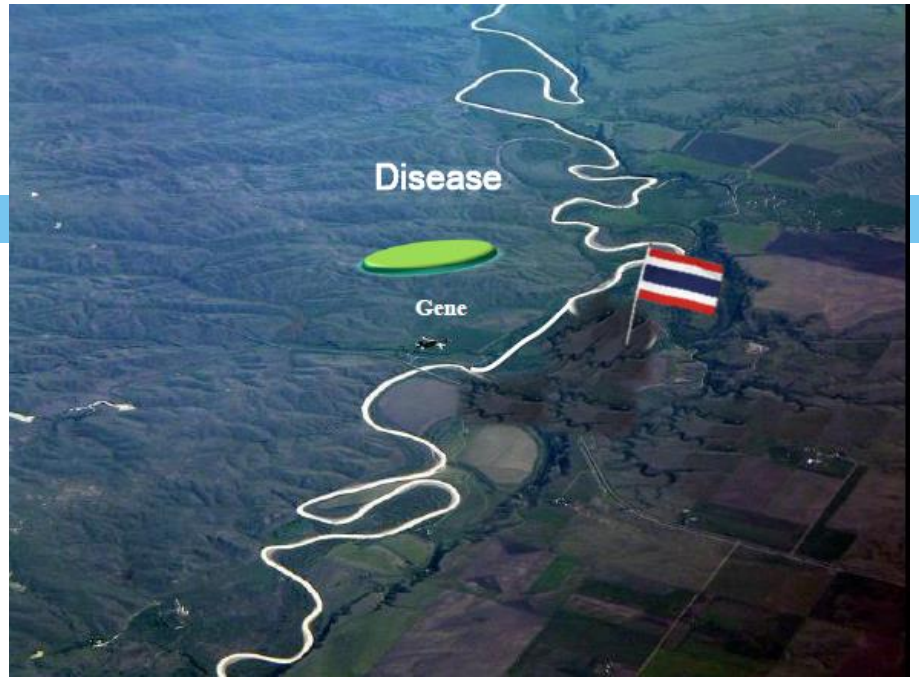
The more distant two genes, the more frequently separated by recombination

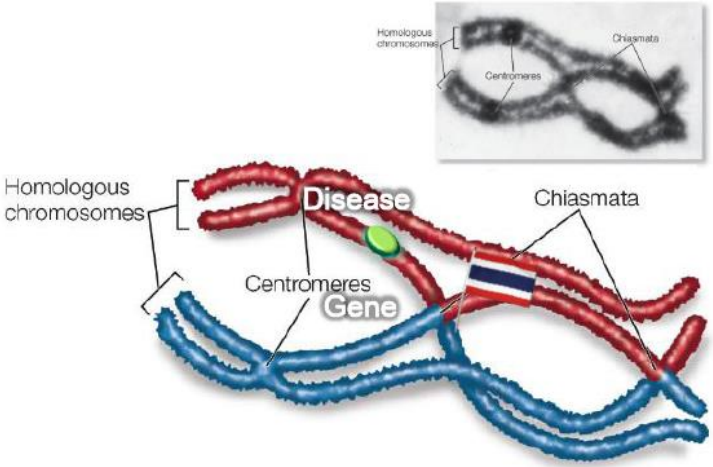
Each 1% recombination frequency constitutes 1 centimorgan (cM)

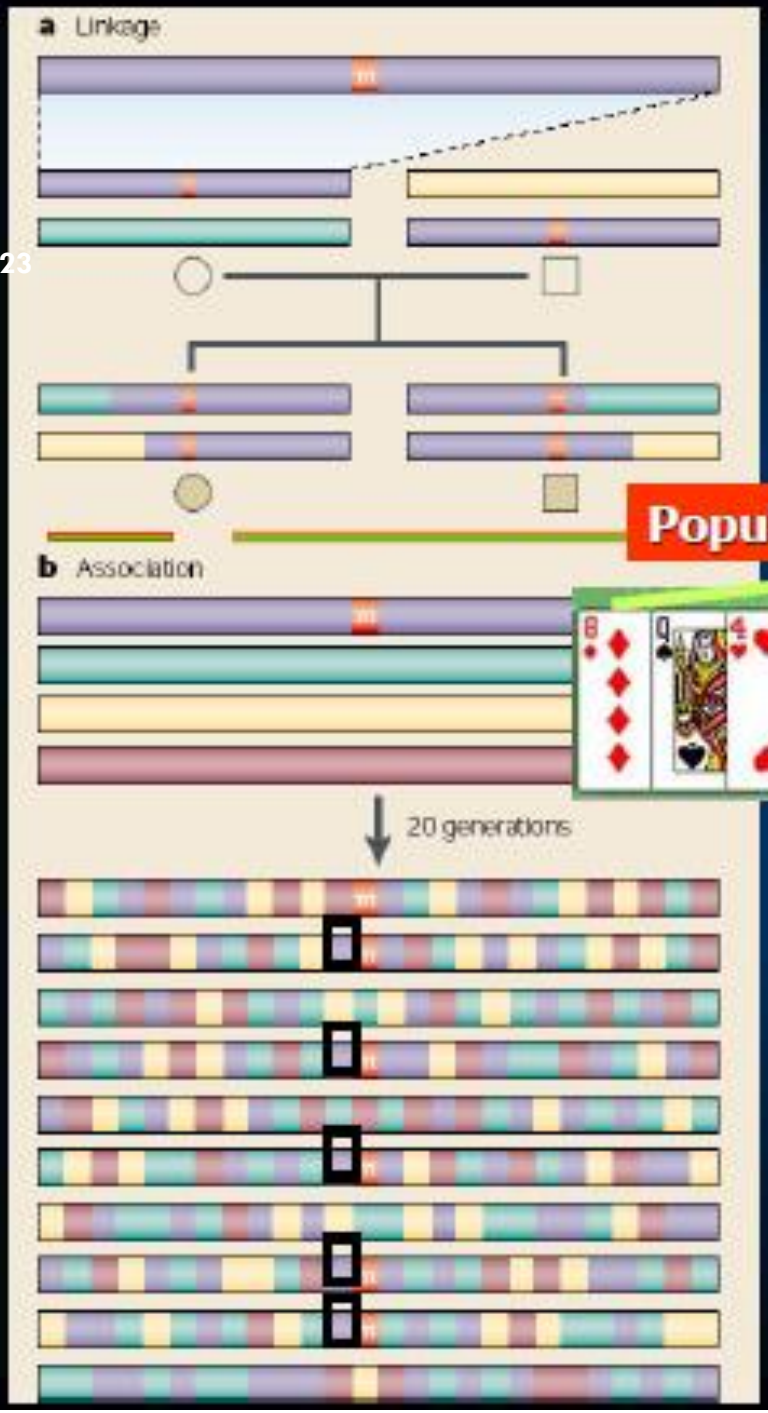


Linkage Disequilibrium Between: Allele D and Marker 4


Disease
Gene


Disease
Gene


Homologous chromosomes
Chiasmata
Centromeres
Disease
Gene



Ethnic

Linkage analysis focuses only on recent, usually observable ancestry, in whom there have been relatively few opportunities for recombination to occur, disease gene regions that are identified by linkage will often be large, and can encompass hundreds or even thousands of possible genes across many megabases of DNA. Crossing over has less influence in relatively inbred community.

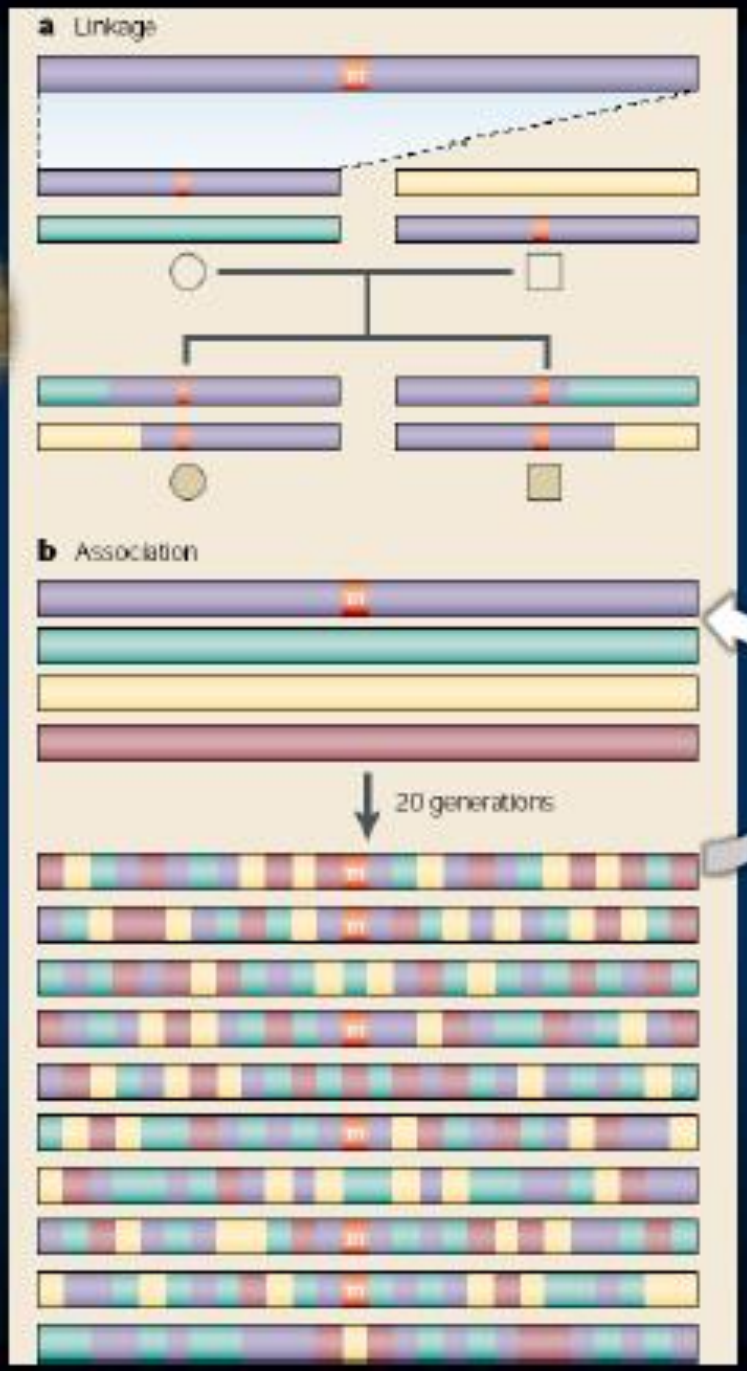
linkage Equilibrium

Population/Race

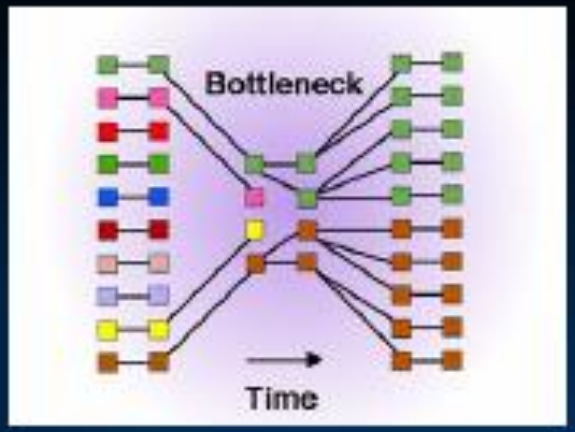
Association studies draw from historic recombination so disease-associated regions are (theoretically) extremely small in outbred random mating populations, encompassing only one gene or gene fragment.

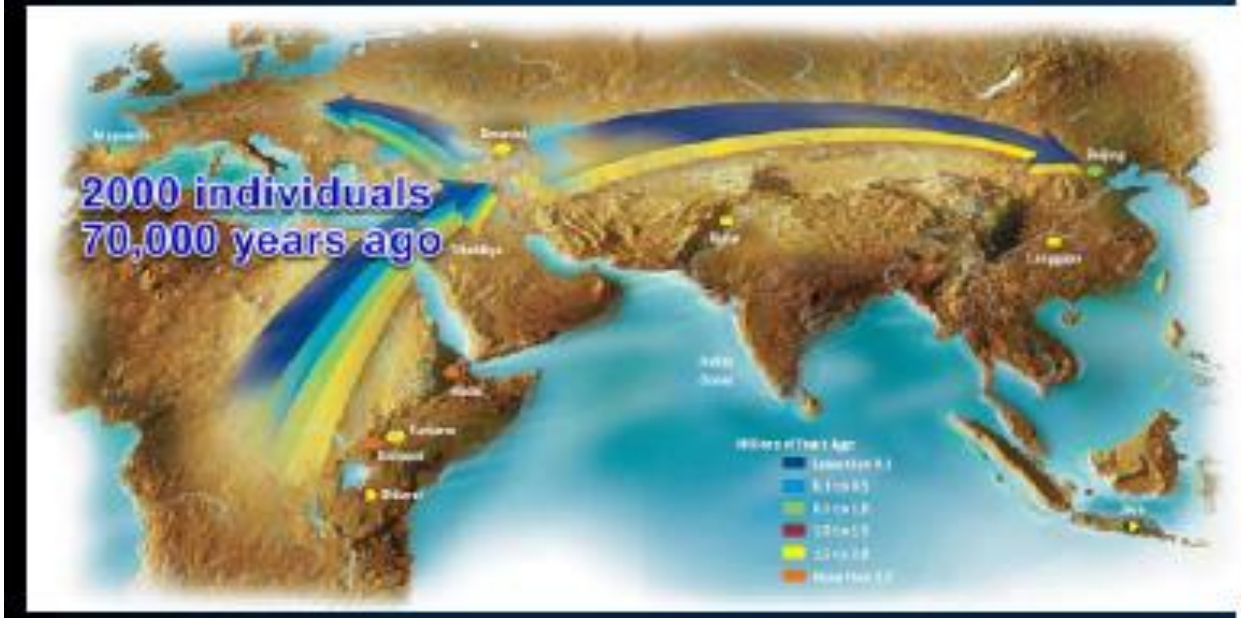
Thus, association studies can be regarded as very large linkage studies of unobserved, hypothetical pedigrees.

Particular DNA variants can remain together on ancestral haplotypes for many generations. This type of non-random association of alleles is known as linkage disequilibrium. It is **linkage disequilibrium** that provides the genetic basis for most association strategies.



Intermarriages:
Many generations

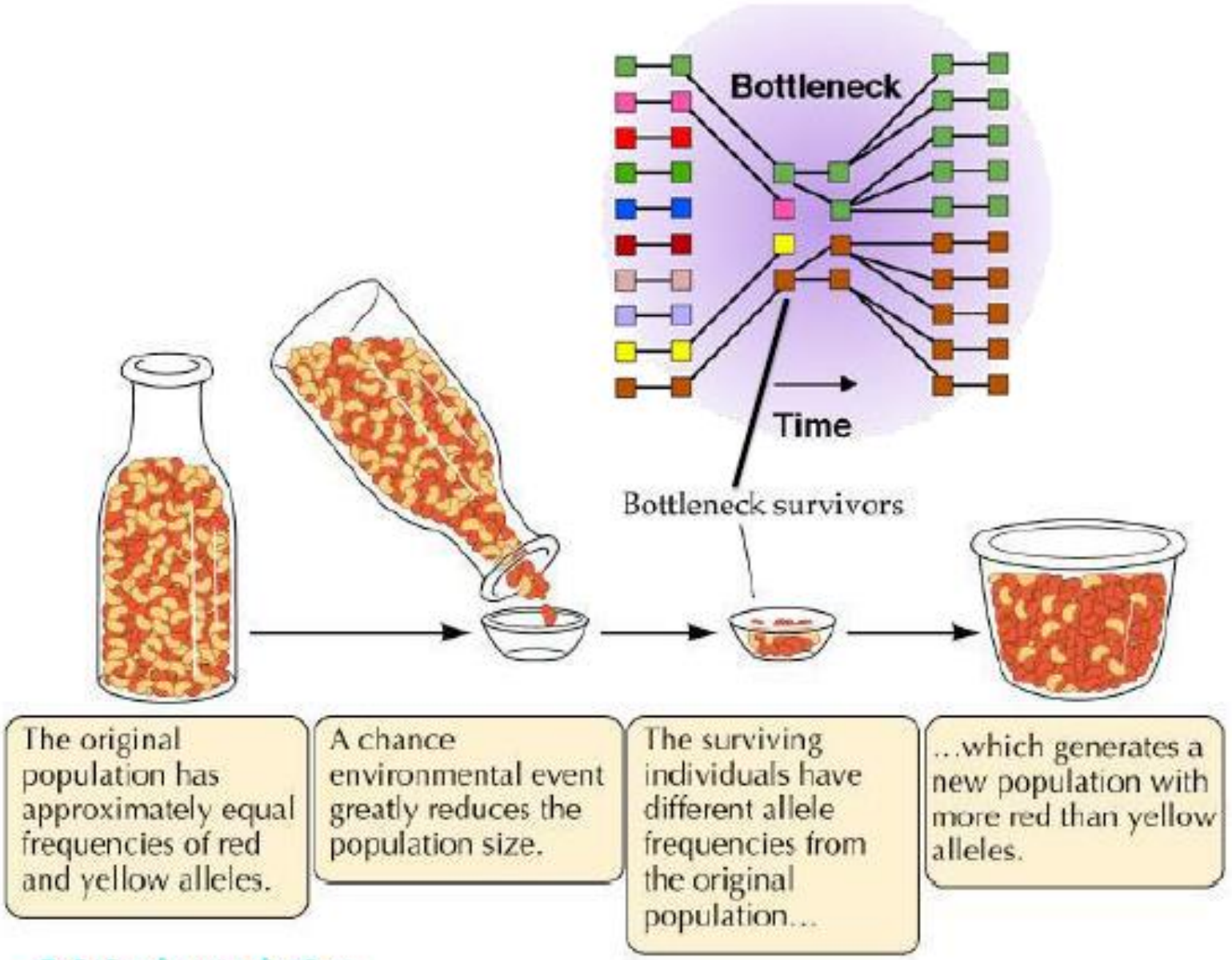




Immigrants are at risk of Skin cancer or vitamin D deficiency.

Why we're slightly different.

Differ by 0.1%


Bottleneck
Time
Bottleneck survivors
The original population has approximately equal frequencies of red and yellow alleles.
A chance environmental event greatly reduces the population size.
The surviving individuals have different allele frequencies from the original population…
…which generates a new population with more red than yellow alleles.
Original population

Genome-Wide Association Studies (GWAS)

GWAS are large-scale genetic investigations of human disease that measure 100 of 1000 of genetics variants scattered throughout the human genome

Advance technologies and falling prices

Search for genetic influences on common diseases of major public health significance.

Since 2005: > 1,600 publications have identified 2000 replicated genetics association with >300 common human diseases and traits.

Genome-Wide Association Study (GWAS)

An approach that involves rapidly **scanning markers across the complete sets of DNA**, **or genomes**, of many people to find genetic variations associated with a particular disease.

Once new genetic associations are identified, researchers can use the information

- develop better strategies to detect
- treat and prevent the disease

Particularly useful in finding genetic variations that contribute to common, complex diseases

: **asthma, cancer, diabetes, heart disease and mental illnesses**

Why are such studies possible now?

With the completion of the Human Genome Project in 2003 and the International HapMap Project in 2005, researchers now have a set of research tools that **make it possible to find the genetic contributions to common diseases.**

The tools include

- computerized databases that contain the reference human genome sequence,
- a map of human genetic variation
- a set of new technologies

quickly and accurately analyze whole-genome samples for genetic variations that contribute to the onset of a disease.

How will genome-wide association studies benefit human health?

The impact on medical care from genome-wide association studies could potentially be substantial.

- the groundwork for the era of **personalized medicine**, in which the current one size-fits-all approach to medical care will give way to more customized strategies.

In the future, after improvements are made in the cost and efficiency of genome-wide scans and other innovative technologies, health professionals will be able to use such tools.

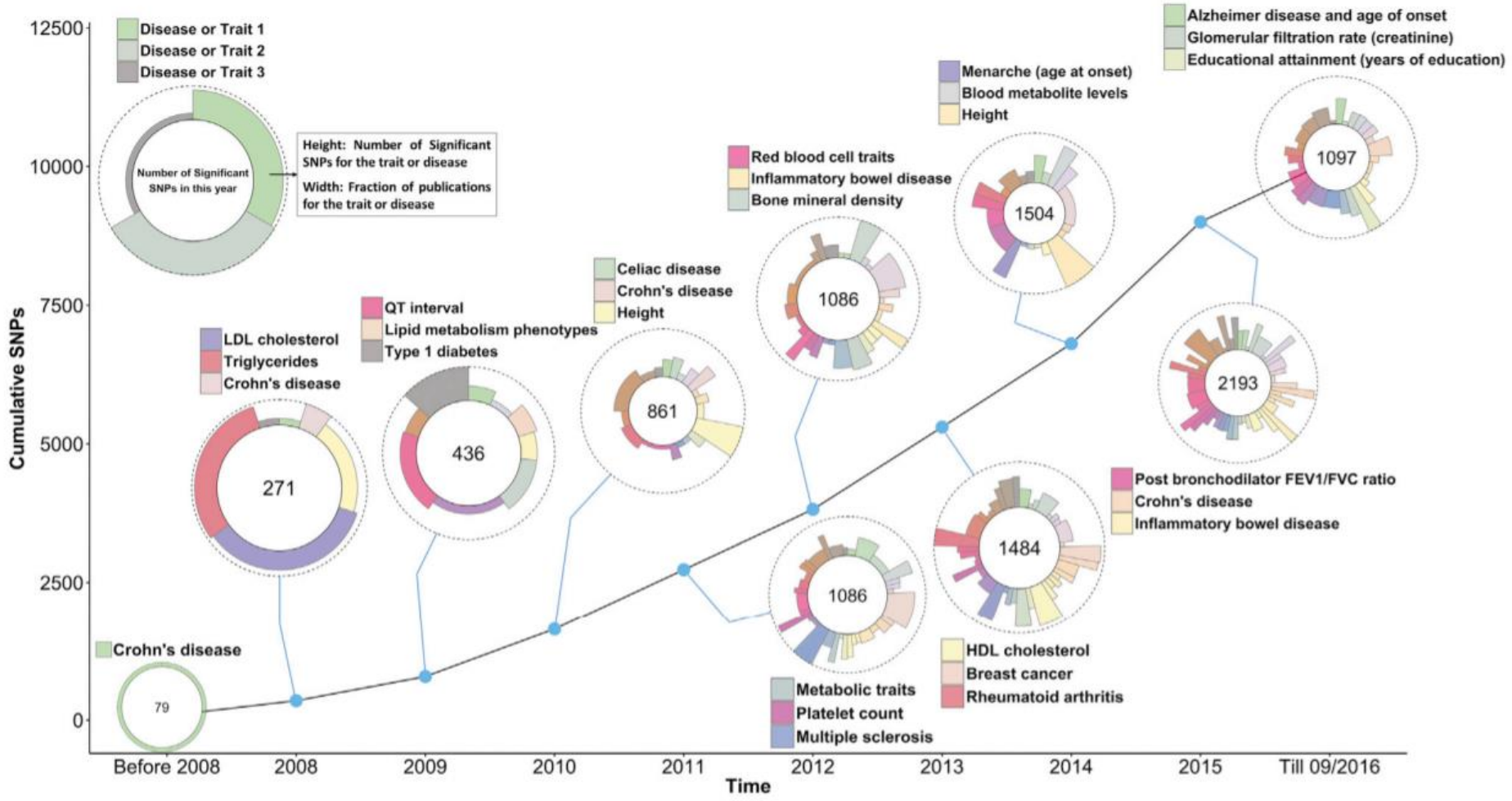
- provide patients with **individualized information about their risks of developing certain diseases.**
- **enable health professionals** to tailor prevention programs to each person's unique genetic makeup.
- If a patient does become ill, the information can be used to **select the treatments most likely to be effective and least likely to cause adverse reactions** in that particular patient.

What have genome-wide association studies found?

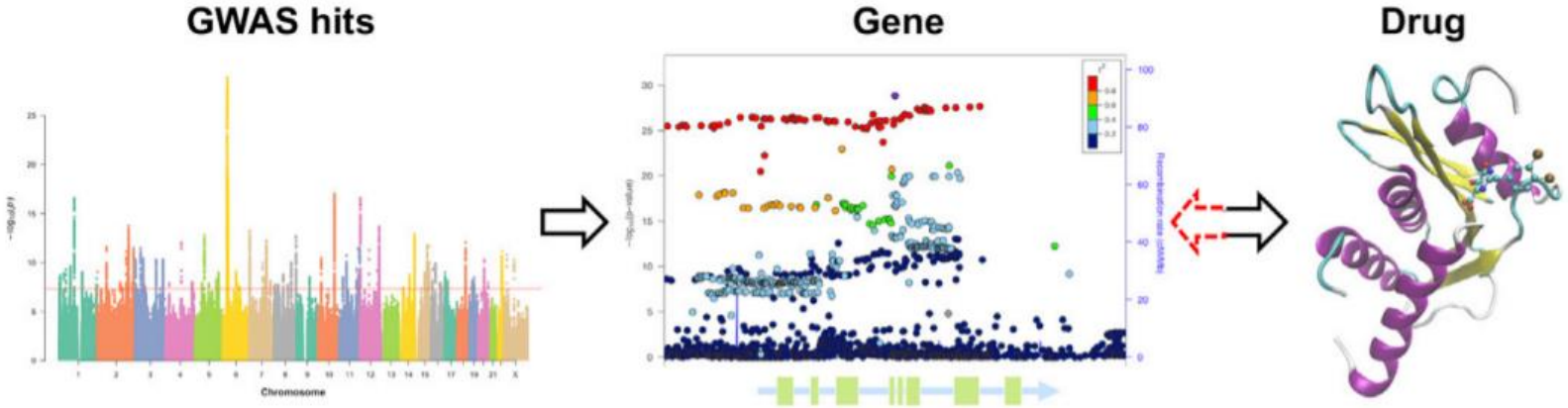
Researchers already have reported considerable success using this new strategy.

- In 2005, three independent studies found that a **common form of blindness is associated with variation in the gene for complement factor H**
- Genome-wide association studies to identify genetic variations that contribute to risk of
  - type 2 diabetes
  - Parkinson's disease
  - heart disorders, obesity
  - Crohn's disease
  - prostate cancer
  - genetic variations that influence response to anti-depressant medications

GWAS SNP-Trait Discovery Timeline



Links between GWAS Discoveries and Drugs



Trait	Gene with GWAS hits	Known or candidate drug
Type 2 Diabetes	*SLC30A8/KCNJ11*	ZnT-8 antagonists/Glyburide
Rheumatoid Arthritis	*PADI4/IL6R*	BB-Cl-amidine/Tocilizumab
Ankylosing Spondylitis(AS)	*TNFR1/PTGER4/TYK2*	TNF-inhibitors/NSAIDs/fostamatinib
Psoriasis(Ps)	*IL23A*	Risankizumab
Osteoporosis	*RANKL/ESR1*	Denosumab/Raloxifene and HRT
Schizophrenia	*DRD2*	Anti-psychotics
LDL cholesterol	*HMGCR*	Pravastatin
AS, Ps, Psoriatic Arthritis	*IL12B*	Ustekinumab

Genome-Wide Association Study (GWAS)

Genome-wide markers (>100,000 markers)
>100,000 tables of 2 by 3 table of case vs control genotype frequencies

Multiple tests of table (Hypotheses)
- Significant threshold at 9.37×10^{-8} (533,252 SNPs)

How are genome-wide association studies conducted?

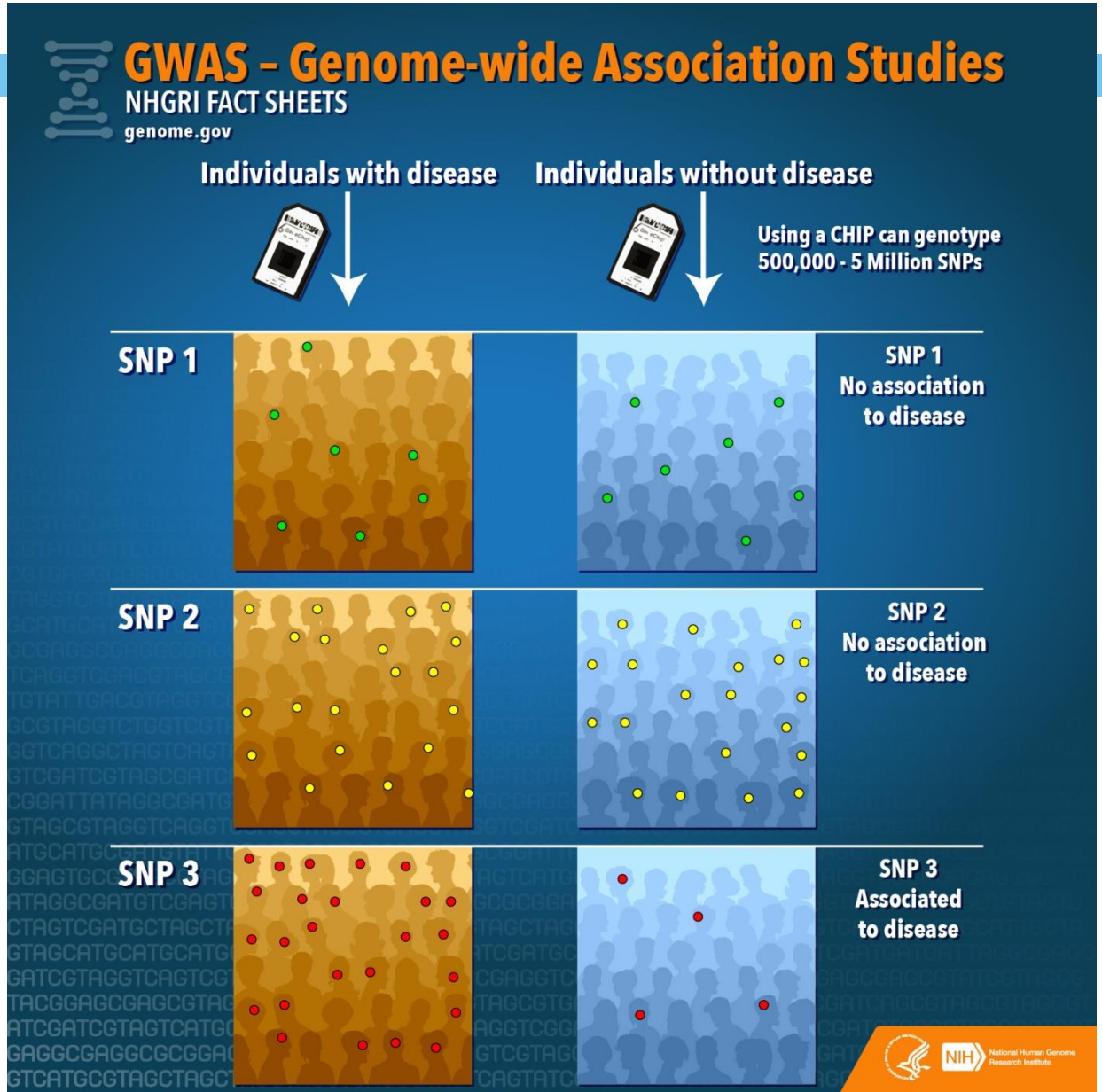
These studies compare the DNA of two groups of participants: people with the disease/ADRS (cases) and similar people without (controls).

DNA is extracted from these participants, and spread on DNA chips, which can read millions of SNPs that are markers for DNA variations.

If genetic variations are more frequent in people with the disease/ADRs, the variations are said to be "associated" with the disease/ADRS (P value should be at least $=10^{-8}$)

The associated genetic variations are pointers to the region of the human genome where the disease/ADRs-causing problem is likely to reside.

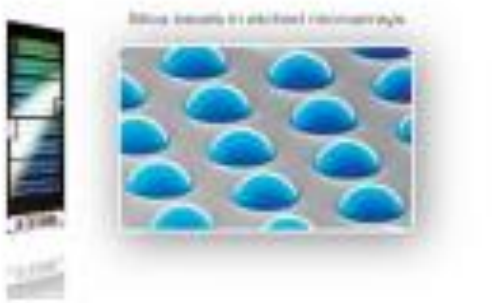
Genome-Wide Association Study (GWAS)

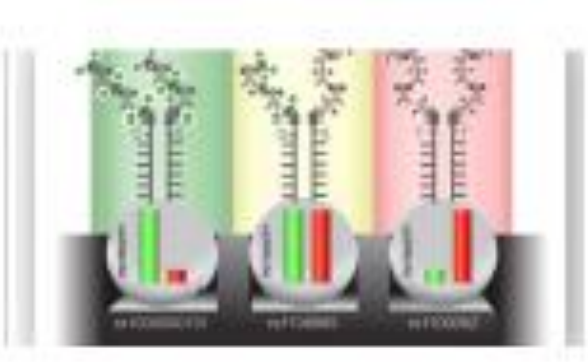


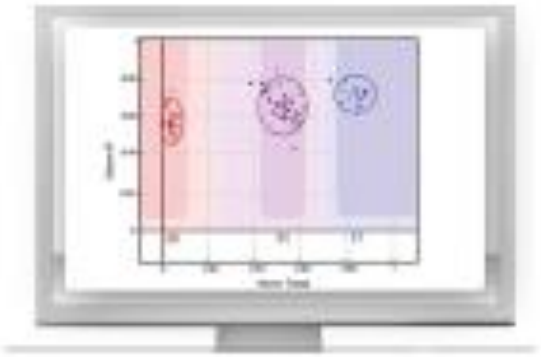
Genotyping technology

> 100,000 SNPs genotyped at once by microarray technology or next generation sequencing

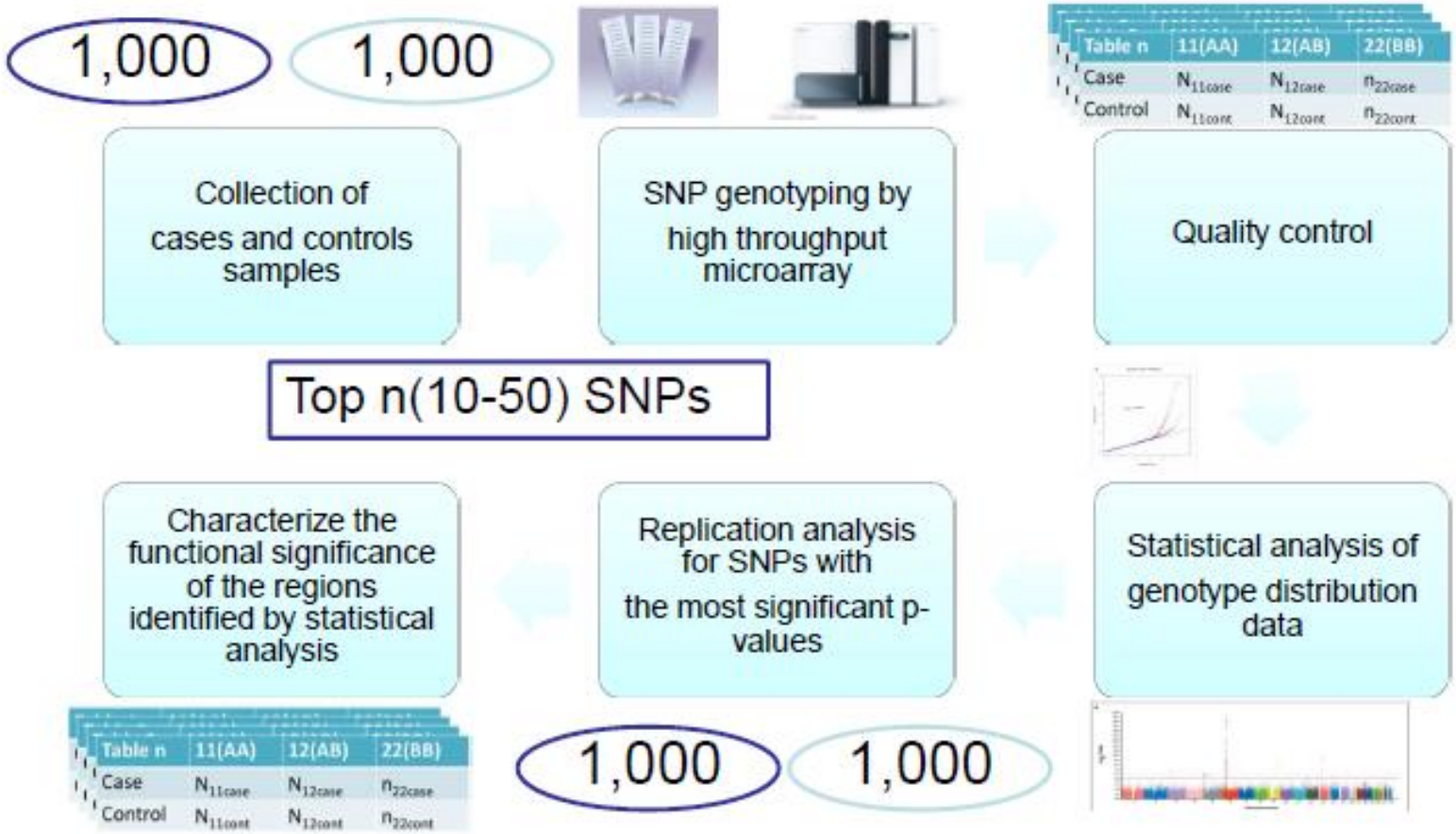
Genome-Wide Genotyping







Genome-Wide Association Study (GWAS)



The NEW ENGLAND JOURNAL of MEDICINE

REVIEW ARTICLE

GENOMIC MEDICINE

W. Gregory Feero, M.D., Ph.D., and Alan E. Guttmacher, M.D., *Editors*

Genomewide Association Studies and Assessment of the Risk of Disease

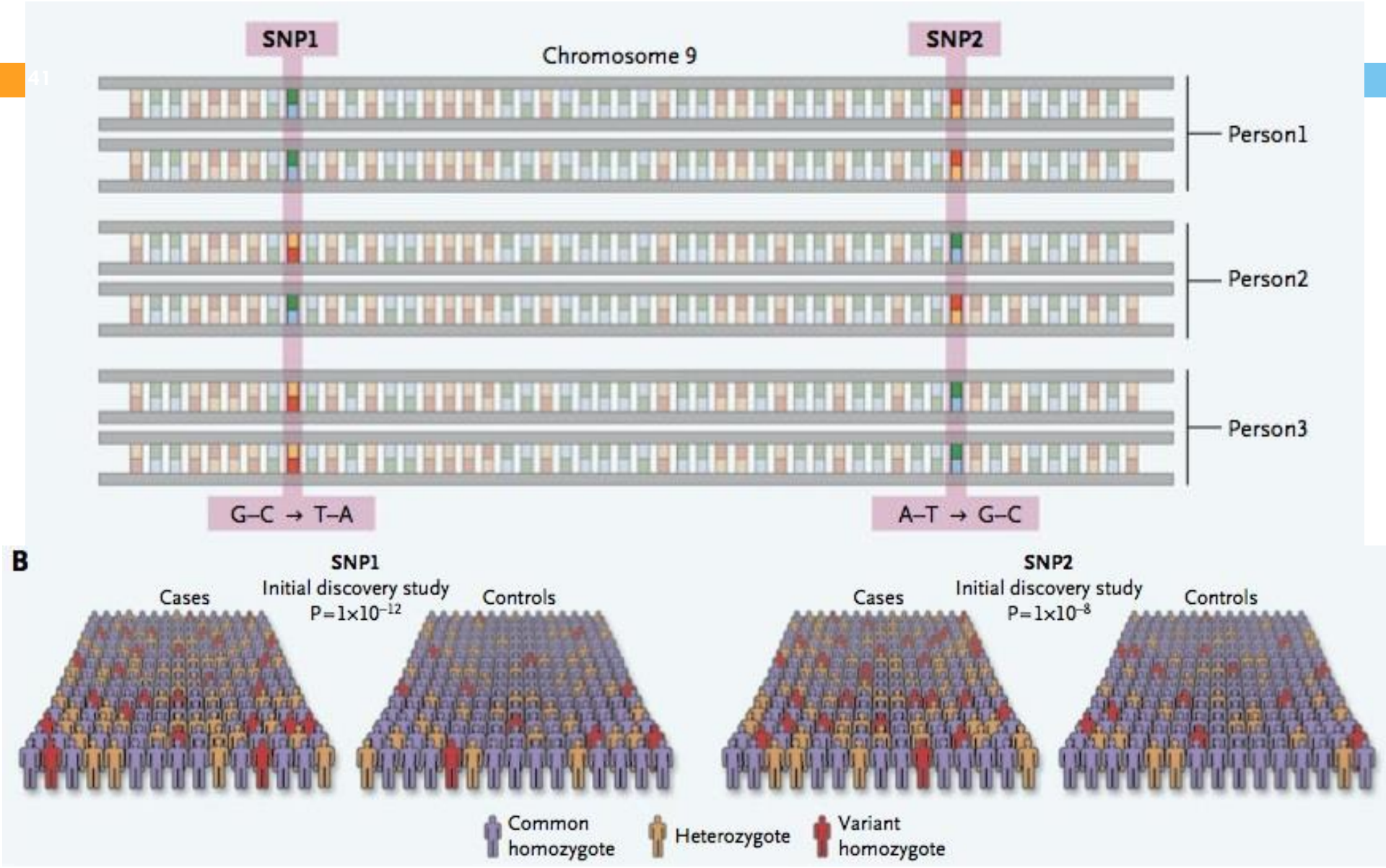
Teri A. Manolio, M.D., Ph.D.

From the Office of Population Genomics, National Human Genome Research Institute, Bethesda, MD. Address reprint requests to Dr. Manolio at the Office of Population Genomics, National Human Genome Research Institute, Bldg. 31, Rm. 4B-09, 31 Center Dr., MSC 2152, Bethesda, MD 20892, or at manolio@nih.gov.

N Engl J Med 2010;363:166-76.

GENOMEWIDE ASSOCIATION STUDIES — IN WHICH HUNDREDS OF THOUsands of single-nucleotide polymorphisms (SNPs) are tested for association with a disease in hundreds or thousands of persons (Fig. 1) — have revolutionized the search for genetic influences on complex traits.[1,2] Such conditions, in contrast with single-gene disorders, are caused by many genetic and environmental factors working together, each having a relatively small effect and few if any being absolutely required for disease to occur. Although complex conditions have been referred to as the geneticist's nightmare,[3] in the past 5 years genomewide association studies have identified SNPs implicating hundreds of robustly replicated loci (i.e., specific genomic locations) for common traits.[4]

100 of 1,000 of SNPs are tested for association with the disease in 100 or 1,000 of person



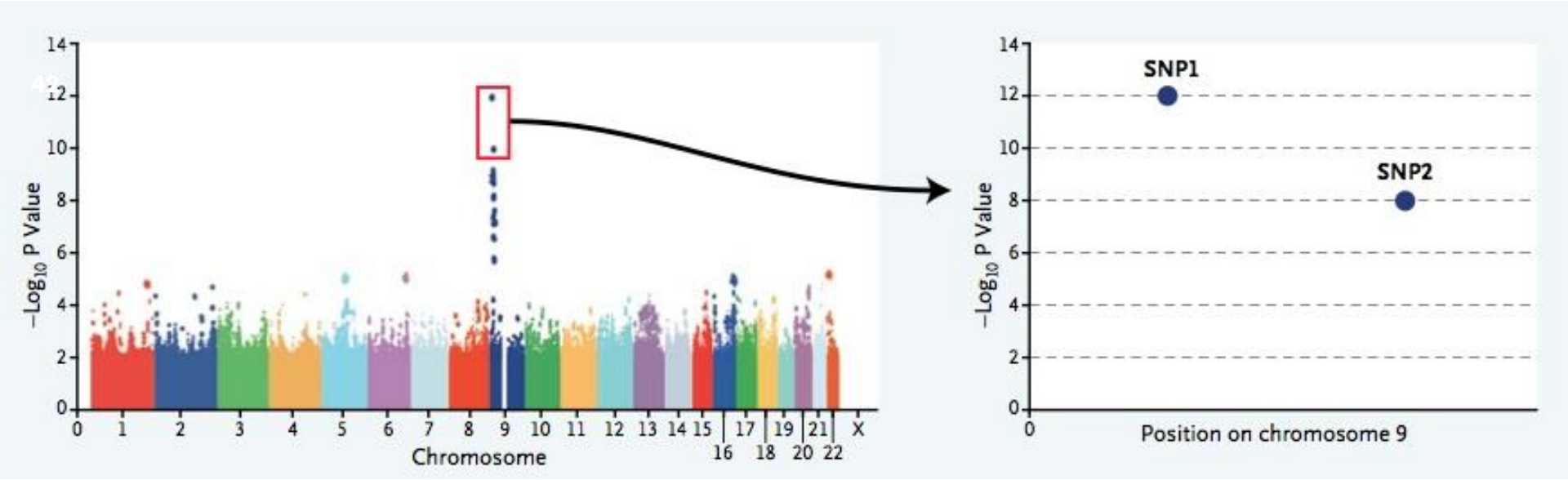


Figure 1. The Genomewide Association Study.
The genomewide association study is typically based on a case–control design in which single-nucleotide polymorphisms (SNPs) across the human genome are genotyped. Panel A depicts a small locus on chromosome 9, and thus a very small fragmentCOoLOfRthFleGUgREenome. In Panel B, the strength of association between each SNP and disease is calculated on the basis of the prevalence of each SNP in cases and Rev4 06/22 /10 controls. In this example, SNPs 1 and 2 on chromosome 9 are associated with disease, with P values of 10−12 and 10−8, respectively. The Author Dr. Manolio plot in Panel C shows the P values for all genotyped SNPs that have survived a quality-control screen, with each chromosome shown in
Fig # 1 a different color. The results implicate a locus on chromosome 9, marked by SNPs 1 and 2, which are adjacent to each other (graph at right), and other neighboring SNPs.

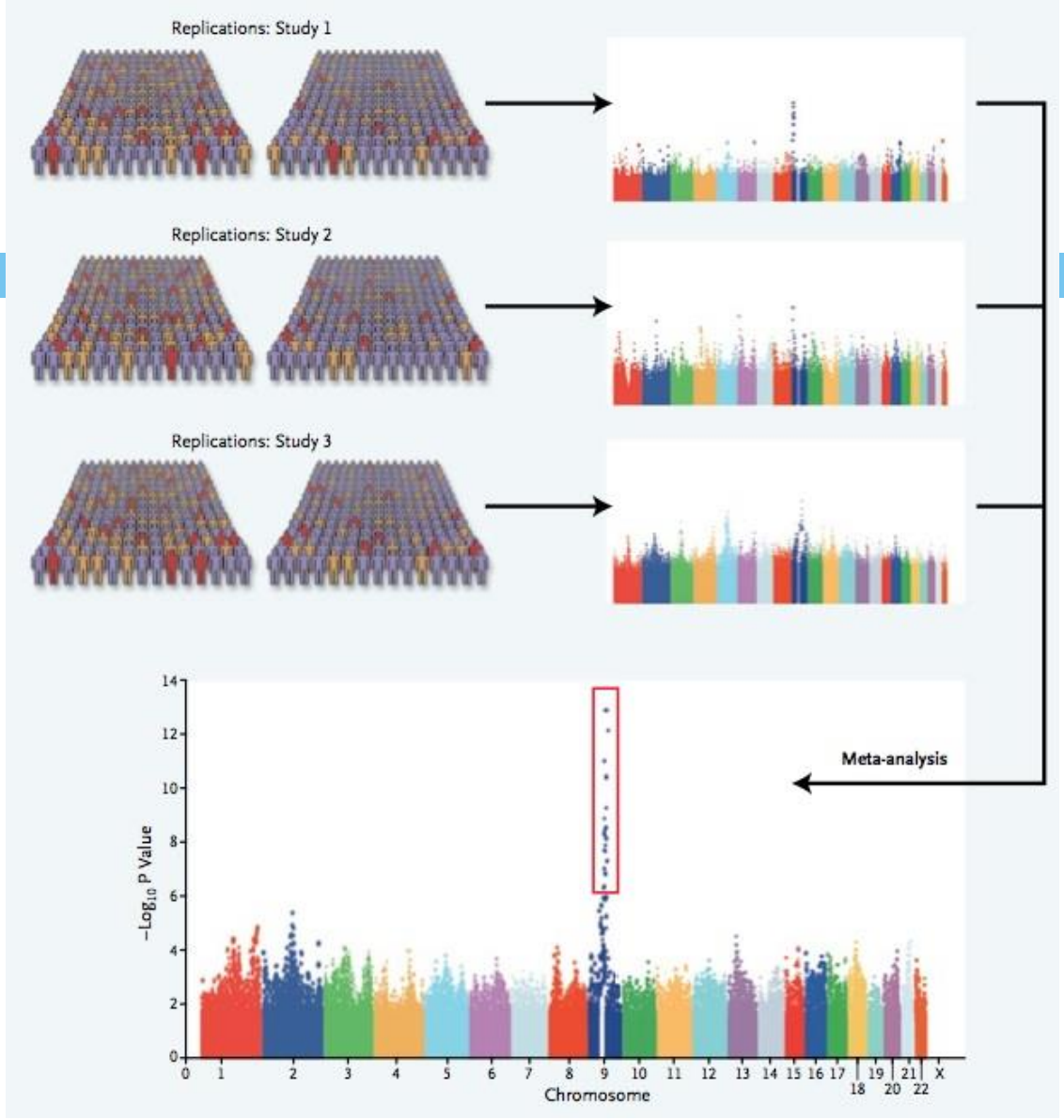


Figure 2. Meta-Analysis of Genomewide Association Studies.
The results of genomewide association studies can be evaluated in a meta-analysis, which combines the results of multiple studies to improve the power for detecting associations. In this example, the results of three studies, none COLOR FIGURE
Author Dr. Manolio of which may show genomewide significance individually, are combined in a meta-analysis to reveal a strong, signifi- cant signal on chromosome 9.

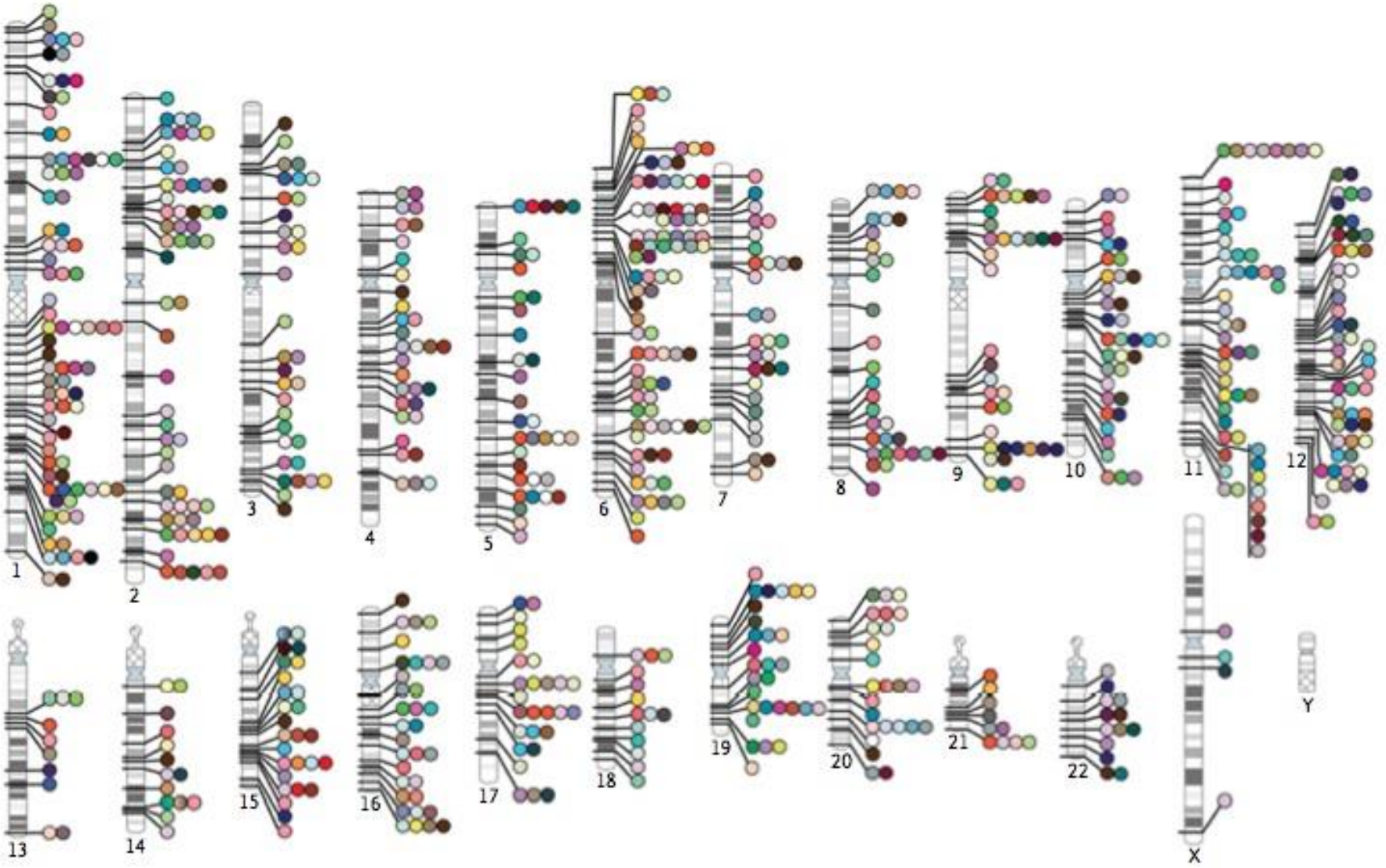


Figure 3. Genomewide Associations Reported through March 2010.

Circles indicate the chromosomal location of nearly 800 single-nucleotide polymorphisms (SNPs) significantly associated ($P<5\times10^{-8}$) with a disease or trait and reported in the literature (545 studies published through March 2010 yielded the associations depicted). Each disease type or trait is coded by color. Adapted from the National Human Genome Research Institute.[4]

Some GWAS examples

Science News

Colorectal Cancer Risk Increased By Single-Base Change In The Human Genome

ScienceDaily (June 29, 2009) — Finnish Academy Professors Lauri Aaltonen and Jussi Taipale have identified and described a mechanism whereby a single-base change in the human genome increases the risk of colorectal cancer.

nature genetics VOLUME 41 | NUMBER 8 | AUGUST 2009 NATURE GENETICS

The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling

Sari Tuupanen[1], Mikko Turunen[2,3], Rainer Lehtonen[1], Outi Hallikas[2,3], Sakari Vanharanta[1,12], Teemu Kivioja[2–4], Mikael Björklund[2,3], Gonghong Wei[2,3], Jian Yan[2,3], Iina Niittymäki[1], Jukka-Pekka Mecklin[5], Heikki Järvinen[6], Ari Ristimäki[7–9], Mariachiara Di-Bernardo[10], Phil East[11], Luis Carvajal-Carmona[11], Richard S Houlston[10], Ian Tomlinson[11], Kimmo Palin[4,12], Esko Ukkonen[4], Auli Karhu[1], Jussi Taipale[2,5] & Lauri A Aaltonen[1]

Homozygosity for the G allele of rs6983267 at 8q24 increases colorectal cancer (CRC) risk ~1.3 fold. We report here that the risk allele G shows copy number increase during CRC development. Our computer algorithm, Enhancer Element Locator (EEL), identified an enhancer element that contains rs6983267. The element drove expression of a reporter gene in a pattern that is consistent with regulation by the key CRC pathway Wnt. rs6983267 affects a binding site for the Wnt-regulated transcription factor TCF4, with the risk allele G showing stronger binding *in vitro* and *in vivo*. Genome-wide ChIP assay revealed the element as the strongest TCF4 binding site within 1 Mb of *MYC*. An unambiguous correlation between rs6983267 genotype and *MYC* expression was not detected, and additional work is required to scrutinize all possible targets of the enhancer. Our work provides evidence that the common CRC predisposition associated with 8q24 arises from enhanced responsiveness to Wnt signaling.

Some GWAS example

Nat Genet. 2016 May 2. doi: 10.1038/ng.3562. [Epub ahead of print]

Five endometrial cancer risk loci identified through genome-wide association analysis.

Cheng TH[1], Thompson DJ[2], O'Mara TA[3], Painter JN[3], Glubb DM[3], Flach S[1], Lewis A[1], French JD[3], Freeman-Mills L[1], Church D[1], Gorman M[1], Martin L[1]; National Study of Endometrial Cancer Genetics Group (NSECG), Hodgson S[4], Webb PM[3]; Australian National Endometrial Cancer Study Group (ANECS), Attia J[5,6], Holliday EG[5,6], McEvoy M[6], Scott RJ[5,7,8,9], Henders AK[3], Martin NG[3], Montgomery GW[3], Nyholt DR[3,10], Ahmed S[11], Healey CS[11], Shah M[11], Dennis J[2], Fasching PA[12,13], Beckmann MW[13], Hein A[13], Ekici AB[14], Hall P[15], Czene K[15], Darabi H[15], Li J[15], Dörk T[16], Dürst M[17], Hillemanns P[18], Runnebaum I[17], Amant F[19], Schrauwen S[19], Zhao H[20,21], Lambrechts D[20,21], Depreeuw J[19,20,21], Dowdy SC[22], Goode EL[23], Fridley BL[24], Winham SJ[23], Njølstad TS[25,26], Salvesen HB[25,26], Trovik J[25,26], Werner HM[25,26], Ashton K[5,8,9], Otton G[27], Proietto T[27], Liu T[28], Mints M[29], Tham E[28,30]; RENDOCAS; CHIBCHA Consortium, Li MJ[31], Yip SH[31], Wang J[23,32], Bolla MK[2], Michailidou K[2], Wang Q[2], Tyrer JP[11], Dunlop M[33,34], Houlston R[35], Palles C[1], Hopper JL[36]; AOCS Group, Peto J[37], Swerdlow AJ[35,38], Burwinkel B[39,40], Brenner H[41,42,43], Meindl A[44], Brauch H[43,45,46], Lindblom A[28], Chang-Claude J[47,48], Couch FJ[23,49], Giles GG[36,50,51], Kristensen VN[52,53,54], Cox A[55], Cunningham JM[49], Pharoah PD[11], Dunning AM[11], Edwards SL[3], Easton DF[2,11], Tomlinson I[1], Spurdle AB[3].

⊕ **Author information**

Abstract

We conducted a meta-analysis of three endometrial cancer genome-wide association studies (GWAS) and two follow-up phases totaling 7,737 endometrial cancer cases and 37,144 controls of European ancestry. Genome-wide imputation and meta-analysis identified five new risk loci of genome-wide significance at likely regulatory regions on chromosomes 13q22.1 (rs11841589, near KLF5), 6q22.31 (rs13328298, in LOC643623 and near HEY2 and NCOA7), 8q24.21 (rs4733613, telomeric to MYC), 15q15.1 (rs937213, in EIF2AK4, near BMF) and 14q32.33 (rs2498796, in AKT1, near SIVA1). We also found a second independent 8q24.21 signal (rs17232730). Functional studies of the 13q22.1 locus showed that rs9600103 (pairwise r^2 = 0.98 with rs11841589) is located in a region of active chromatin that interacts with the KLF5 promoter region. The rs9600103[T] allele that is protective in endometrial cancer suppressed gene expression in vitro, suggesting that regulation of the expression of KLF5, a gene linked to uterine development, is implicated in tumorigenesis. These findings provide enhanced insight into the genetic and biological basis of endometrial cancer.

Some GWAS example

Pathogens. 2016 Apr 26;5(2). pii: E39. doi: 10.3390/pathogens5020039.

A Cellular GWAS Approach to Define Human Variation in Cellular Pathways Important to Inflammation.

Miller SI[1], Chaudhary A[2].

Author information

[1]Department of Microbiology, Department of Immunology, Department of Medicine, Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. millersi@uw.edu.
[2]Department of Microbiology, University of Washington, Seattle, WA 98195, USA. anuc@uw.edu.

Abstract

An understanding of common human diversity in innate immune pathways should be beneficial in understanding autoimmune diseases, susceptibility to infection, and choices of anti-inflammatory treatment. Such understanding could also result in definition of currently unknown components of human inflammation pathways. A cellular genome-wide association studies (GWAS) platform, termed Hi-HOST (High-throughput human in vitro susceptibility testing), was developed to assay in vitro cellular phenotypes of infection in genotyped lymphoblastoid cells from genetically diverse human populations. Hi-HOST allows for measurement of multiple host and pathogen parameters of infection/inflammation including: bacterial invasion and intracellular replication, host cell death, and cytokine production. Hi-HOST has been used to successfully define a significant portion of the heritable human diversity in inflammatory cell death in response to Salmonella typhimurium. It also led to the discovery of genetic variants important to protection against systemic inflammatory response syndrome (SIRS) and protection against death and bacteremia in individuals with SIRS. Our laboratory is currently using this platform to define human diversity in autophagy and the NLPR3 inflammasome pathways, and to define new components that can impact the expression of phenotypes related to these pathways.

KEYWORDS: GWAS; Hi-HOST; SIRS; host; lymphoblastoid; pathogen; pyroptosis

Implication: Disease prevention

- Increased cost-effectiveness of prevention strategy by gene-based risks stratifying
  - *Streptococcus pneumoniae* vaccination
  - Latent TB treatment
  - HPV vaccination
- Increased surveillance in high genetic risk group
  - Familial cancer syndrome
  - Diseases with high heritability; Type I DM
  - Mendelian screening at birth

Implication: Personalized medicine

- US FDA recommended to test HLA-B for
  - Carbamazepine in Asian (HLA-B*1502)
  - Abacavir in Caucasian (HLA-B*5707)
  - TPMT genetic test for 5-FU
  - CYP2*9 and VKORC for Warfarin dosing
- Selection of effective drug
- Avoidance of side effect

Human Genome variation database

IMS-JST FROM JAPANESE JSNP DATABASE | Japanese Single Nucleotide Polymorphisms

Japanese | Site Map | Text Size S M L

HOME | Keyword Search | Integrated Map | Homology Search | Genotype Summary Data | Download | Help

Search by: Search | Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y Unknown

Search JSNP

JSNP® Search Guide

- Integrated Map Browser
- Search Genes
- Search SNPs
- SNPs in DME
- Show List of SNPs
- BLAT Genome
- Search Example
- Help

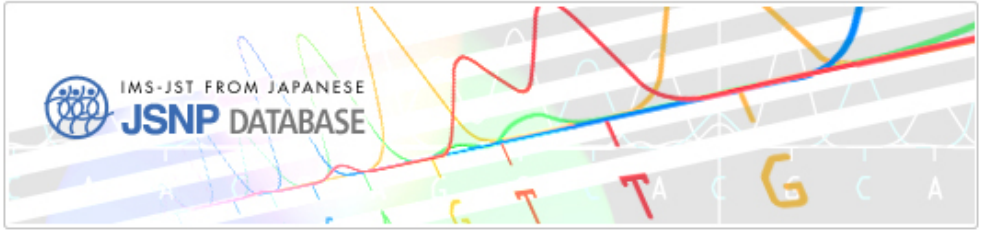
Download

Download Guide

- Dump JSNP®
- FASTA files(FTP)
- XML files

Notice

JSNP database will not be accessible



What's New

2012-04-04

The 40th data release consists of 197,195 SNPs ; 84,651 SNPs with allele frequency.(release 40)

2012-04-04

The Integrated mapping information is based on NCBI's build 37 version 3, Ensembl Human release 65 and rs number of dbSNP

RSS You can use the RSS feed to keep updated about JSNP releases.

Citing JSNP® database

When referring to JSNP® database in scientific communications, please use the following reference;

JSNP: a database of common gene variations in the Japanese population

Nucleic Acids Research, 30:158-162, 2002 PubMed

Genotype Summary Data

- Genotype Summary Data Access
- Control
- Case
- Search Tag SNP

Documents

- What's New
- HapMap Related News
- Release Notes
- Statistics
- About JSNP®
- Publications
- Data Contributions
- List of SNP-trait associations

Perlegen Sciences, Inc. is working to provide safe and effective medicines to patients worldwide. The company quickly and cost effectively analyzes more than one million genetic variations in clinical trial participants to explain and predict the efficacy and adverse effect profiles of prescription drugs. Perlegen also applies this expertise to discovering genetic variants associated with disease in order to pave the way for new therapeutics and diagnostics. For years, scientists and drug manufacturers have been eager to comprehensively examine entire genomes; through Perlegen, their moment has come. Perlegen is able to bring drugs to the market whose clinical development would have otherwise been discontinued.

We're changing forever the way medicine is practiced.

SNPs on the Mapping 100K; 50% from public databases, and 50% from SNP database created by Perlegen Sciences, Inc.

International HapMap Project

The International HapMap Project is a partnership of scientists and funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States to develop a public resource that will help researchers find genes associated with human disease and response to pharmaceuticals. See "**About the International HapMap Project**" for more information.

Project Information

About the Project
HapMap Publications
HapMap Tutorial
HapMap Mailing List
HapMap Project Participants
HapMap Mirror Site in Japan

Project Data

HapMap Genome Browser (Phase 1, 2 & 3 - merged genotypes & frequencies)
HapMap Genome Browser (Phase 3 - genotypes, frequencies & LD)
HapMap Genome Browser (Phase 1 & 2 - full dataset)
GWAs Karyogram

News

- 2009-09-09: **Website Maintenance**
  The Hapmap website will be undergoing maintenance from 1pm-5pm EDT on September 9th. We apologize for the disruption.
- 2009-04-02: **HapMap3 CEL files available**

  Raw signal intensity data from HapMap3 genotypes on the Genome-Wide Human SNP Array 6.0 are now **available for bulk download**.
- 2009-02-09: **HapMap3 Phased Haplotypes available**

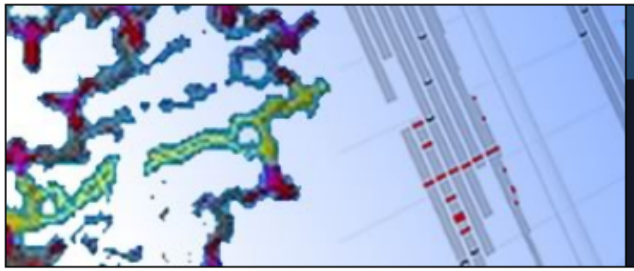
  Phased haplotypes for consensus HapMap3 release 2 data has been phased for autosomes are now **available for bulk download**.
- 2009-02-06: **HapMap Public Release #27 (merged II+III)**

  Genotypes and frequency data for the three phases of the project (I+II: rel #24 and III: release #2), were combined in NCBI build 36 (dbSNP b126) coordinates. Data is **available for downloading** and also **available for browsing**. Click here to read the latest **release notes**.

NCBI Resources How To Sign in

dbSNP SNP Search

Limits Advanced



dbSNP

Database of single nucleotide polymorphisms (SNPs) and multiple small-scale variations that include insertions/deletions, microsatellites, and non-polymorphic variants.

Getting Started

- Overview of dbSNP
- FAQ

Submit Data

- Clinically Associated Human Variations
- All Other Variations
- Hold Until Published (HUP) Policies

Access Data

- Important RefSNP (RS) Attributes
- Web Search
- Batch Query
- FTP Download

หน้าแรก
เกี่ยวกับโครงการเภสัชพันธุศาสตร์
- เภสัชพันธุศาสตร์ในประเทศกำลังพัฒนา
- แผนผังองค์กร
- เป้าหมาย
- โครงสร้างพื้นฐาน
- บุคลากร
- ความร่วมมือระหว่างหน่วยงาน

งานวิจัยเภสัชพันธุศาสตร์
ผลงานวิชาการ
บริการ
เภสัชพันธุศาสตร์ - วิกิพีเดีย
ข่าวของโครงการเภสัชพันธุสาสตร์
ติดต่อกับเรา
ENGLISH
LINK

หน้าแรก ▸ เกี่ยวกับโครงการเภสัชพันธุศาสตร์ ▸ โครงสร้างพื้นฐาน

PHARMACOGENOMICS COLLABORATIVE PROJECT

THAILAND CENTER OF EXCELLENCE FOR LIFE SCIENCES (TCELS)

AND

RAMATHIBODI HOSPITAL (MAHIDOL UNIVERSITY)

Vision:

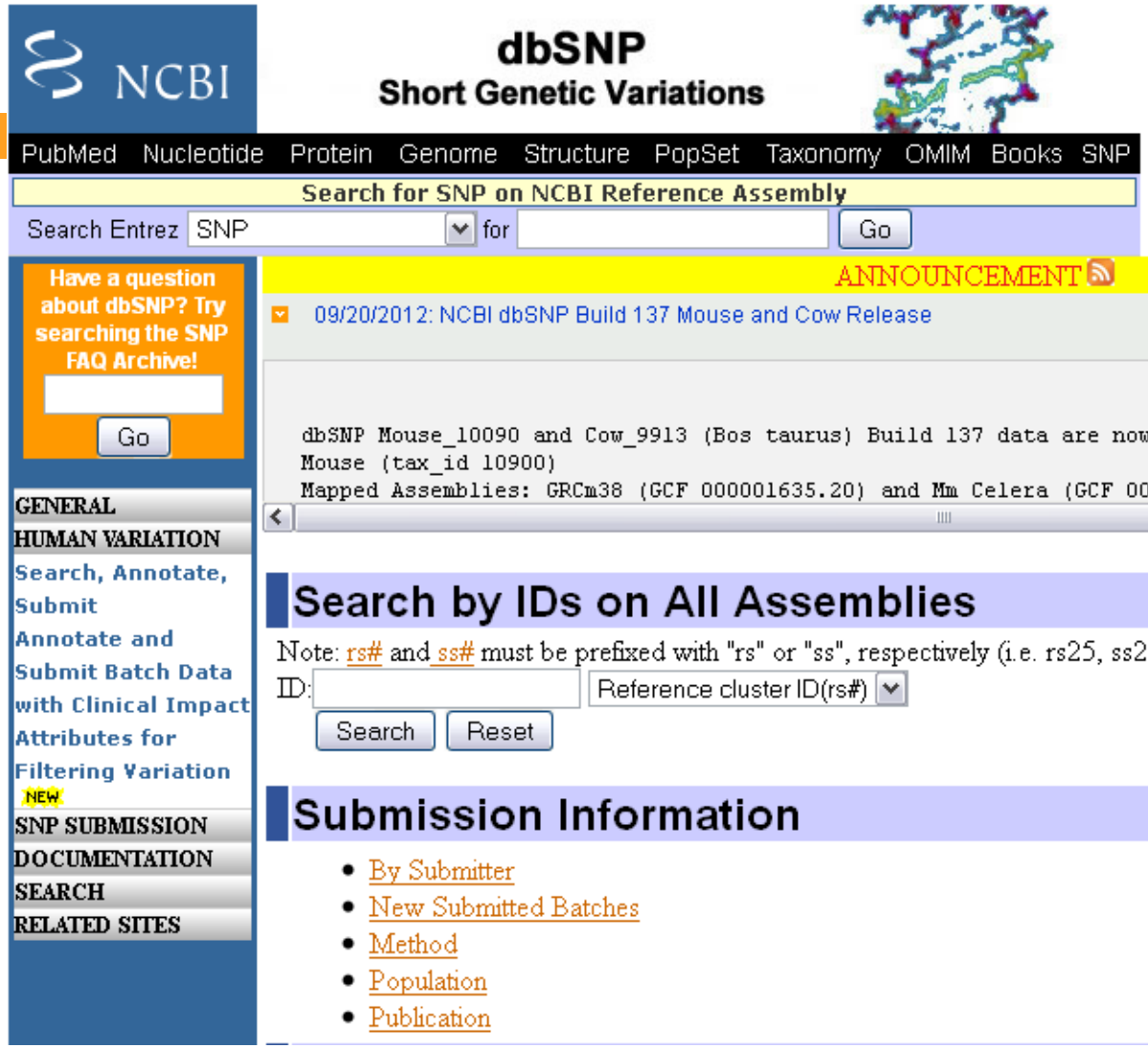
The discovery of genetic variances that affect drugaction will lead to the development of new diagnostic procedures andtherapeutic products, which will enable selective prescription of safe andeffective drugs. This advancement will save time, resources and money, andimprove the quality of life and well-being of patients.

Mission:

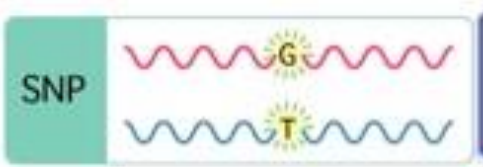
To identify genetic variances that contributes to thevariable effects of drugs in different individuals.

Strategy:

Aclinical and genomic industrial quality database and tool software will be setup based on pharmacogenomics datum standards and will involve collaborativeresearch with industrial partners: such as, Advanced BiologicalLaboratories (ABL) S.A. the European academic research center (CRP-Santé,Luxembourg) and Oracle to verify their association with phenotypic and genotypic data. Pharmacogenomic researchers from targeted hospitals participating inthe project will then utilize bioinformatics tools to search for and collectvast clinical data involving the response, metabolism and side effects ofdrugs, toxicity phenotype, etc., which will be accompanied by genomicfingerprints to identify similar patients. This information will be used todetermine or predict the very best course of treatment and identify potentialclinical trials or pre-clinical testing groups of patients.

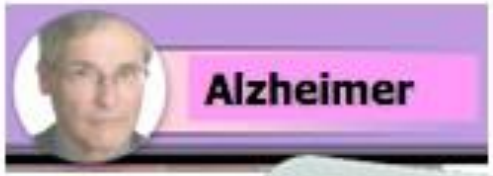

NCBI
dbSNP
Short Genetic Variations
PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP
Search for SNP on NCBI Reference Assembly
Search Entrez SNP for Go
Have a question about dbSNP? Try searching the SNP FAQ Archive!
Go
GENERAL
HUMAN VARIATION
Search, Annotate, Submit
Annotate and Submit Batch Data with Clinical Impact Attributes for Filtering Variation
NEW
SNP SUBMISSION
DOCUMENTATION
SEARCH
RELATED SITES
ANNOUNCEMENT
09/20/2012: NCBI dbSNP Build 137 Mouse and Cow Release
dbSNP Mouse_10090 and Cow_9913 (Bos taurus) Build 137 data are now available:
Mouse (tax_id 10900)
Mapped Assemblies: GRCm38 (GCF 000001635.20) and Mm Celera (GCF 000002165.2)
Search by IDs on All Assemblies
Note: rs# and ss# must be prefixed with "rs" or "ss", respectively (i.e. rs25, ss25)
ID:
Reference cluster ID(rs#)
Search
Reset
Submission Information
By Submitter
New Submitted Batches
Method
Population
Publication

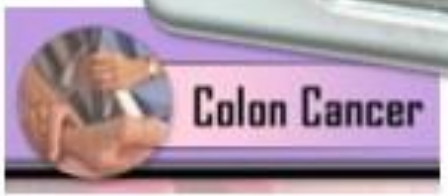
Genetic variations



Making differences in quantity and/or quality of gene products (proteins, functional RNAs)

Consequences of disease prediction: Public may not welcome.





Consequences of phamacogenetics testing: Public may accept.

An attempt to use common SNPs that show association with diseases, as " a tool for screening assay" mostly fail

- However, the bad news is GWAS are not useful in finding genes or genomic markers that predict risks of disease.
- Therefore, general, screening tests, based on most SNPs detected in GWAS to date, are likely to have low positive (and negative) predictive value for disease and limited usefulness in a diagnostic setting.
- This observation has led many to question of the common disease-common variant hypothesis and has contributed to growing interest in evaluating the roles of rare genetic variants in common diseases (using whole genome sequencing).
- However, the good news is GWAS have discovered pharmacogenomic related genes for some traits, such as severe adverse reactions to certain drugs, which are essentially monogenic and already used clinically.

Personal (whole) genome sequencing

Do we really need it or why do we need it?

- After completing the human genome project in 2003, Francis Collins predictive genetic tests will exist for many common conditions: within ten years (2013).

- The deadline for his prediction is fast approaching (2013), but how close to these tests are we?

The price of whole genome sequencing at $100 (not $1,000) will be hit within a decade, and $1,000 by July 2013

Jonathan M. Rothberg

Growing List Of Genetic Tests

Gene Rearrangement Molecular Detection | Acute Myeloid Leukemia: Molecular Detection of the AML1/ETO Chimeric Transcript | ADmark ApoE Genotype Analysis & Interpretation (Symptomatic) | Alpha-1-Antitrypsin and Alpha-1-Antitrypsin Genotype with Reflex to Alpha-1- Antitrypsin Phenotype | Amyloidosis Evaluation | Anaplastic Large Cell Lymphoma: Molecular Detection of the NPM/ALK Chimeric Transcript APRTAT Aprataxin DNA Sequencing TEST |

Utilizing Whole Genome Sequencing;
We Can do All The Tests at Once (All-In-One)

1. **Genetic disorders, or inherited diseases.**
(โรคพันธุกรรม): monogenetic (Mendelian) disorder.
2. **Genetic risk for common, complex disease.**
(ความเสี่ยงที่จะเกิดโรคพันธุกรรมที่ซับซ้อนและพบบ่อย)
3. **Personalized Therapy**
(การรักษาเฉพาะบุคคล)
3.1 **A genetic drug response or pharmacogenetics**
(เภสัชพันธุศาสตร์).

Clinical genetic screening test
(FDA and non-FDA approved)

- Carrier testing
- Preimplatation
- Prenatal testing
- Newborn screening
- Pharmacogenomic testing
- Diagnostic & Monitoring testing
- Predictive testing.

Test | PINK1 DNA Sequencing Test | PKU Mutation Detection by DNA Sequencing LSMISC Platelet Antigen
Detection of Acute Promyelocytic Leukemia | PMP22 DNA Sequencing Test | Prader Willi/Angelman Syndrome Analysis | Prader-Willi/Angelman Syndrome Molecular Detection | SCA1 DNA Test | SCA10 DNA Test |
SCA14 DNA Test SCA17 SCA17 DNA Test SCA2B SCA2B DNA Test | SCA3 DNA Test | SCA5 DNA Test | SCA6 DNA Test |
SCA7 DNA Test | SCA8 DNA Test | SETX DNA Sequencing Test SMACT Spinal Atrophy LSMISC Spinal Muscular Atrophy
DNA Test | Spinal Muscular Atrophy Test | TPMT Genetics (genotype) | Warfarin Sensitivity by CYP2C9 and VKORC1 Genotyping

All-in-One Assay

WGS; Public health benefit?

NEWS & EVENTS

BGI Unveils Significant New Global Research Collaborations at The 6th International Conference on Genomics

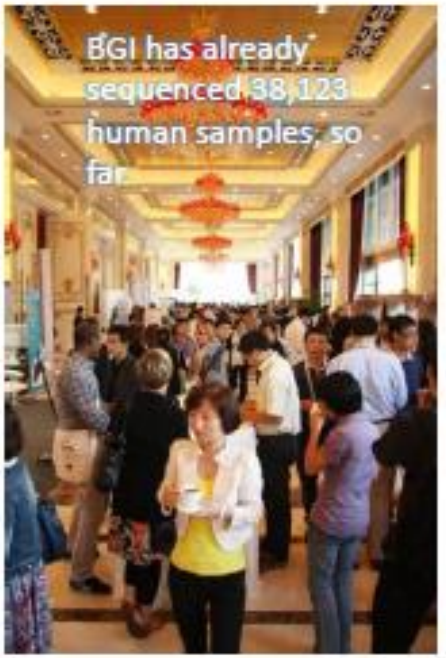
-- Collaborative Initiatives to include Three Million Genomes, 10,000 Rice Genomes and 1% Danes' Genome Project --

- "Three Million Genomes Projects"- M & M & M Projects

- introduced by Dr Jun Wang, Executive Director of BGI, will consist of

1. "Million Plant and Animal Genomes Project,"
2. "Million Human Genomes Project" and
3. "Million Micro-Ecosystem Project."

- They will provide a clear classification on the studied species, and advance the understanding of the species genome and the application of genome-based research for different objectives.



http://www.bgisequence.com/eu/newsandevents/news/bgi-unveils-significant-new-global-research-collaborations-at-th

New vaccine introduction for public health benefits: HPV case study

Yot Teerawattananon, M.D., Ph.D.
Health Intervention and Technology Assessment (HITAP)

3rd Vaccinology course

14 July 2011

A few targets assay, such as real time PCR will soon be replaced by next generation sequencer.

The human genome sequencing

The human genome interpretation

Whole human genome re-sequencer in one day @ $1,000 and dropping to $100

Finding the genetic basis of human disease and drug response

Rare Disease
Identify the variants that underlie rare disease.

Common Disease
Unravel the complex basis of common disease.

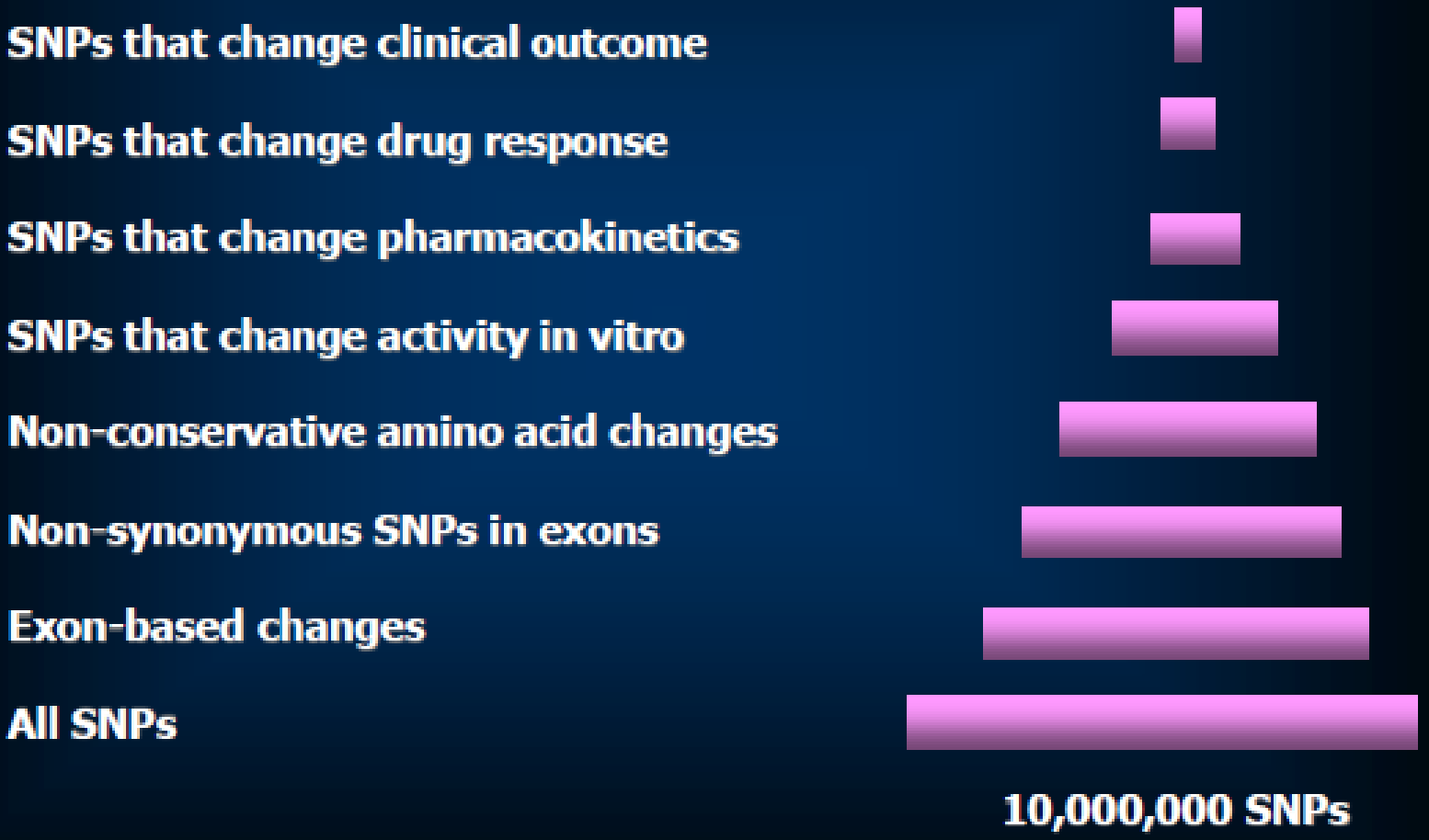
Cancer
Find and understand the drivers of tumor growth.

Drug Response
Pinpoint the genetic variants that govern drug response.

Predictive capacity of whole genome/whole exome sequencing for disease prediction in healthy adults;

1. Complex (common) diseases w/o lifestyle/environmental data =>low
2. Cancer w/o strong family history=> low, hereditary cancers (rare) => high
3. Mendelian (a mutation in a single gene)-many are rare diseases => high
4. Pharmacogenomics => high

Not all SNPs that affect phenotype or clinical outcome but it may be in linkage disequilibrium with other SNP that do so



1000 Genomes

A Deep Catalog of Human Genetic Variation

The 1,000 (Anglo-Saxon) genome project

The 1,000 Chinese genome project

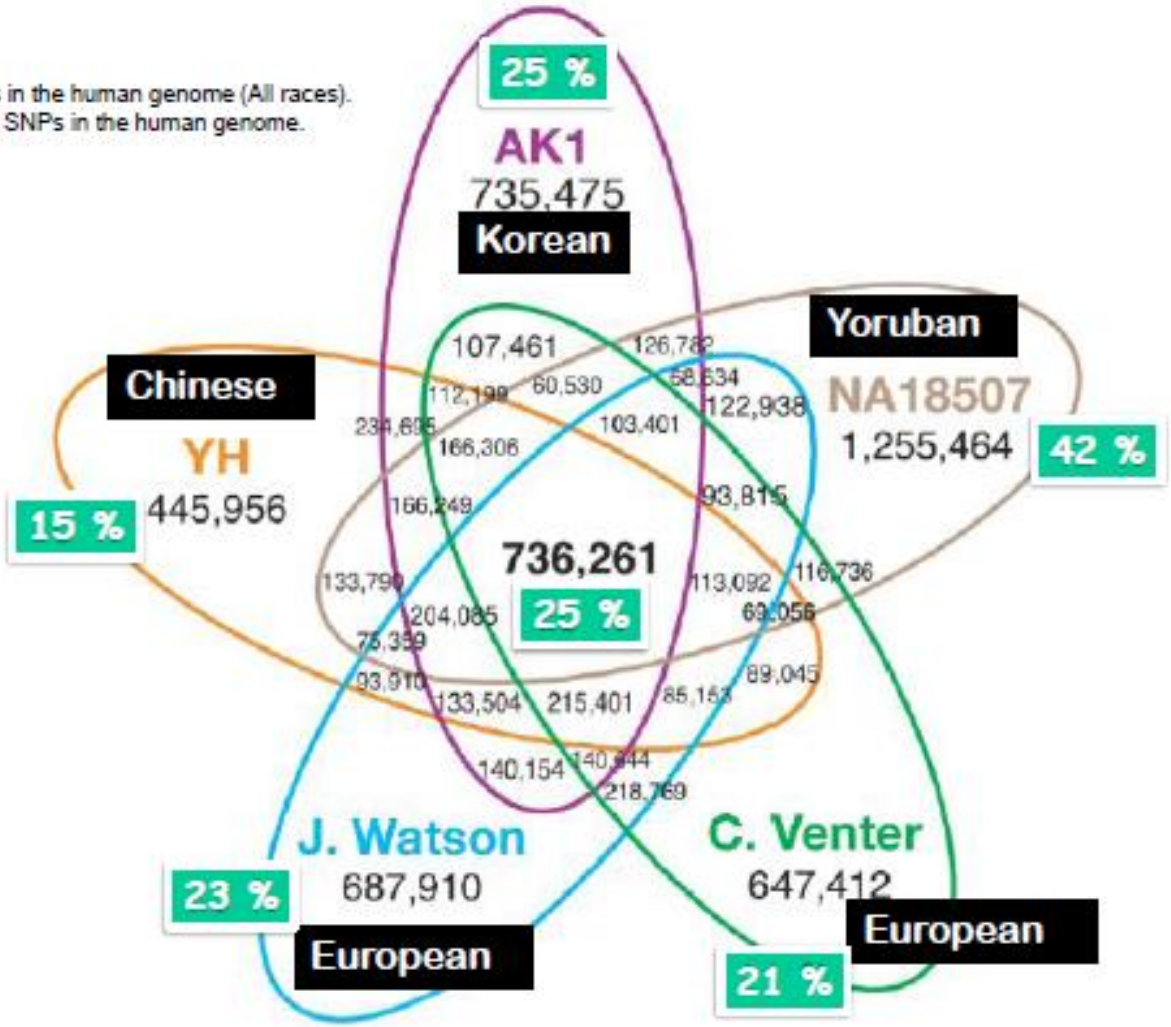
The 1,000 Korean genome project

The 1,000 Japanese genome project

will soon be available for public.

Do we need Thai genome project?

10 million SNPs in the human genome (All races).
3 million comon SNPs in the human genome.



The number of SNPs overlapping between five genomes

Thai human genome project

Prof. Wasun Chantratita lab

ThaiPR.NET

Thailand Press Release

Decoding the Thai Genome Project

วันที่ 12 พฤษภาคม 2554 08:31 น.

ที่มา TCELS

May 12 2011

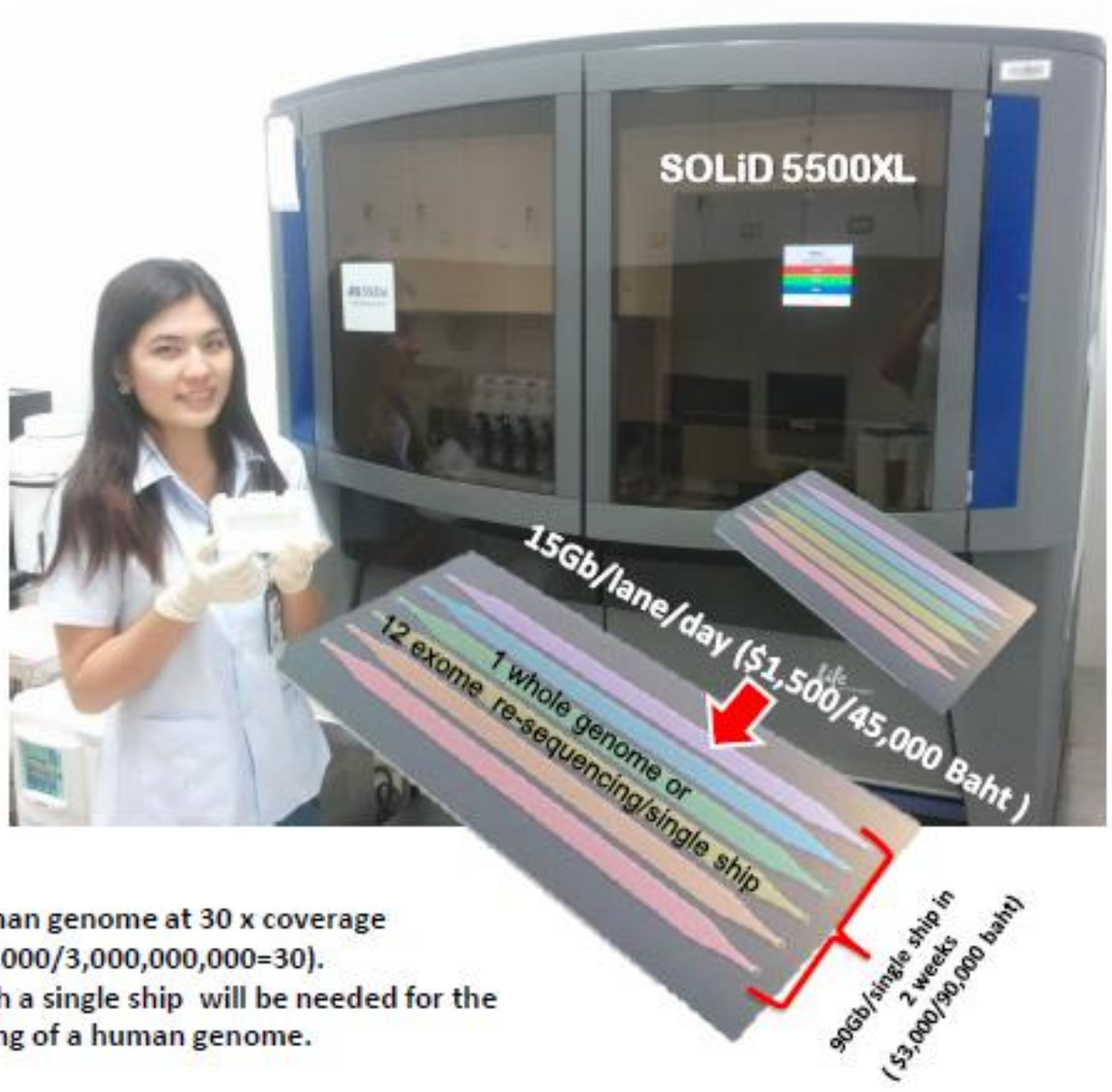
SOLiD 5500XL

From June 2011, the Virology Unit and Laboratory for Pharmacogenomics and Personalized Medicine, Department of Pathology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, with support from the Thailand Center of Excellence for Life Sciences (TCELS), will collaborate jointly to decode the human genome (whole genome sequencing) of three billion bases of an anonymous, healthy Thai male donor as a model for further study of DNA changes (DNA variant).

Dr. Wasun Chantratita, project leader made the announcement that this genome will be compared latter with genomes of various persons who have different diseases, including those that are common and rare. The two main objectives of this project are, firstly, to locate the DNA variant(s) on the genome, which is associated strongly with disease. That genomic maker can then be developed into a set of genetic screening tests at a low cost. If the test results (from the laboratory) are positive, it will mean that the person is at high risk of the diseases in the future. However, the risk may be reduced if these people get regular medical checkups, change or modify their behavior regarding diet or their environment. In children, they are often given a "failure to thrive" diagnosis for an unknown disorder. "Sequencing those genomes will be a key hint to how to treat them properly.

The second objective is to determine the genome of patients for which current therapy does not work properly (difficult to treat). By sequencing their genomes and submitting the DNA variants to the special computational biology program based on computer simulations developed by Dr. Ram Samudrala and his team from the University of Washington, United States of America to identify all approved drugs that can bind to the disease target protein structures which are caused by DNA variants and somehow malfunction in ways that lead to damage and disease in the body. The drug(s) can be picked up and used to replace the medication that did not work in the first place (personalized medicine).

The outcome of the project will finally help both the government and patients regarding effective treatment and reduction of medical expenditure and unnecessary laboratory diagnostic assays. The development of special computational biology program has been supported by the National Institutes of Health (NIH) fund, United States of America.

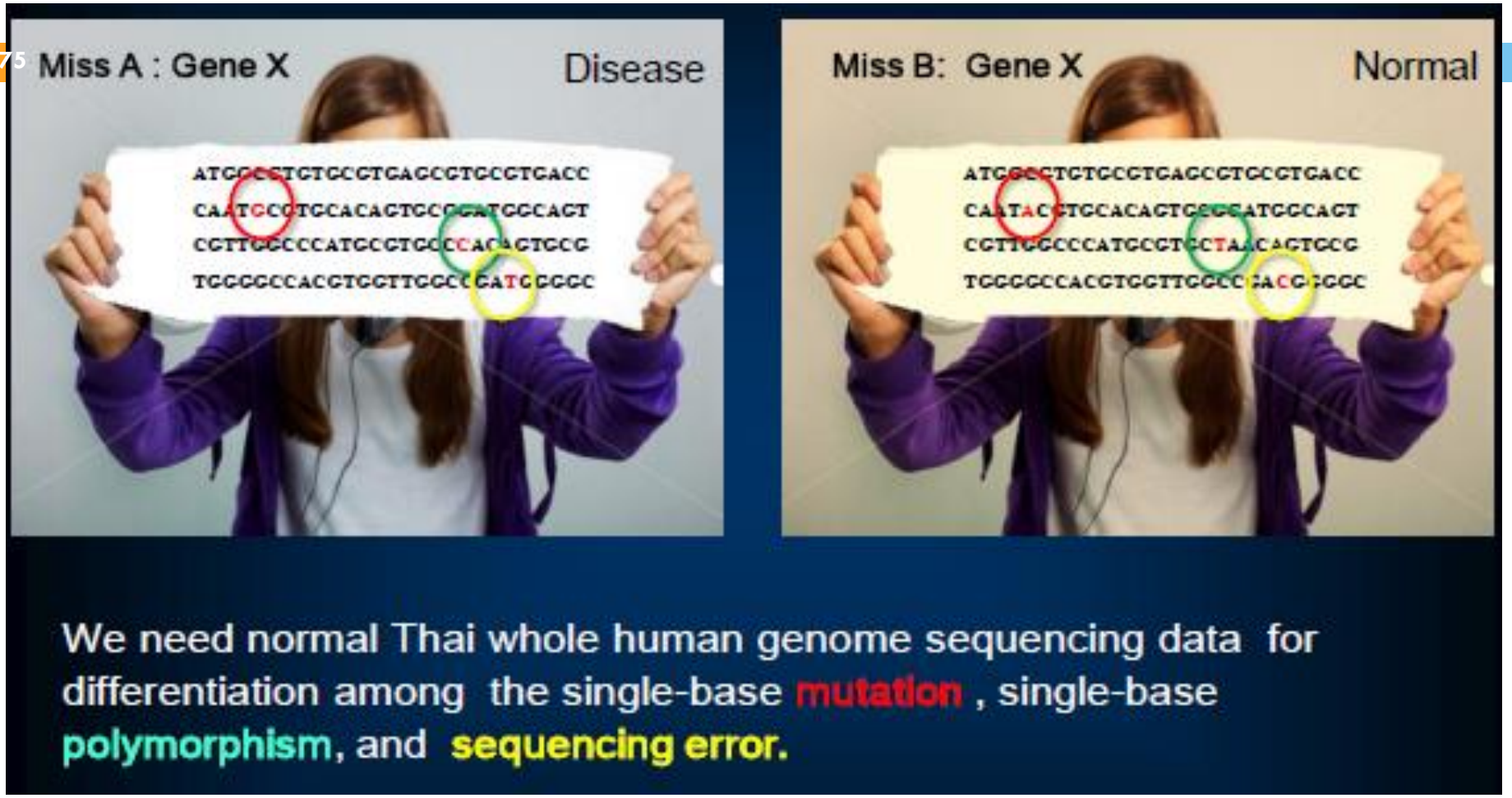


90 Gb:1 human genome at 30 x coverage
(90,000,000,000/3,000,000,000=30).
One run with a single ship will be needed for the re-sequencing of a human genome.

How many Thai human genome have been sequence?

One whole human genome and two whole human exome re-sequencings.

Aiming at 1,000 whole genome/whole exome re-sequencing.


Miss A : Gene X
Disease
Miss B: Gene X
Normal
We need normal Thai whole human genome sequencing data for differentiation among the single-base mutation , single-base polymorphism, and sequencing error.

Integrated Health Care: Disease risk prediction

today…

diagnosis → therapy → therapy monitoring

…towards the future

predisposition screening → targeted monitoring → prevention → diagnosis → therapy → therapy monitoring

Individual Choice

- lifestyle
- nutrition
- medicine

Question?