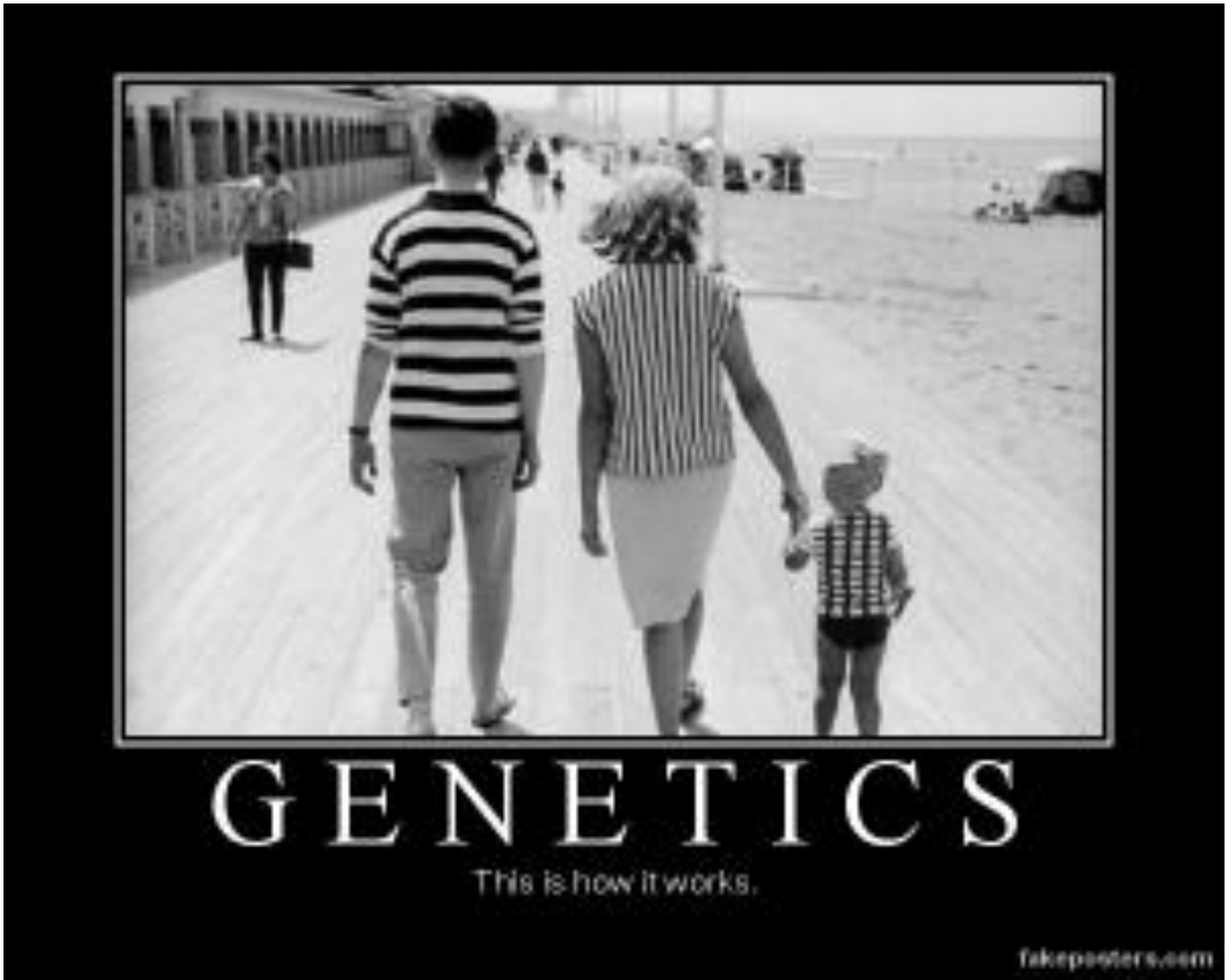
Epidemiology 217
Molecular and Genetic Epidemiology



Outline

- **Course structure**
- Overview of genetics
- Overview of genetic epidemiology and the rest of the course

Training
In
Clinical
Research

- Home
- Overall Schedule
- Courses
- Programs
- Rosters

Molecular and Genetic Epidemiology I
EPI 217 Winter 2016 (2 units)

Course Directors: John S. Witte, PhD
Professor, Department of Epidemiology & Biostatistics
Thomas J. Hoffmann, PhD
Assistant Professor, Department of Epidemiology & Biostatistics

Course Goals

- Develop a framework for interpreting and incorporating genetic information in your research

- **Learn:**
  - Common genetic measures.
  - Some population genetics.
  - Approaches to search for and interpret disease-causing genes.

Course Details

Tuesdays: 1/5/2016–3/15/2016. 1:10-3:00 pm, MH 1406
Website: https://courses.ucsf.edu/course/view.php?id=2591

Course Directors:
John Witte
Professor, UCSF
jwitte@ucsf.edu

Thomas Hoffmann
Asst Professor, UCSF
HoffmannT@humgen.ucsf.edu

Teaching Assistant:
Caroline Tai
PhD Student
caroline.tai@ucsf.edu

Lecturers

Joe Wiemels
Professor, UCSF
joe.wiemels@ucsf.edu

Laura Fejerman
Asst Professor, UCSF
laura.fejerman@ucsf.edu

Eric Jorgenson
Research Scientist, Kaiser
eric.jorgenson@kp.org

Assignments

- **Problem sets (50%)**

Due at noon on Mondays to Caroline Tai

- **Reading / class participation (20%)**

Genetic Epidemiology: Methods and Applications, by Melissa A. Austin. CABI, 2013. ISBN: 978-1780641812.

Books may be purchased either through the publisher or Amazon.com, where you can also read the beginning of the first chapter.

Students may be called upon during class to answer questions about the assigned chapters.

- **Final project (design study)**
  - 30% of grade (due Friday, 3/11 at Noon)
  - Present to class

Syllabus - https://courses.ucsf.edu/course/view.php?id=1862

Topic / Content	Lecturer	Required Reading (pre-lecture)	Assignment Due (Monday @ noon)
Introduction: The Big Picture The process of genetic epidemiology; general approaches to assess the genetic basis of disease.	J Witte	Austin Chapter 1	
Population Genetics, Modeling Genetic Inheritance Basics of population genetics; Hardy-Weinberg Equilibrium; aggregation; heritability.	J Witte	Austin Chapter 2	Assignment 1
Mendel's Laws and Molecular Genetics Mendel's laws (segregation, assortment); molecular measures; genotyping; arrays; sequencing.	J Wiemels	Austin Chapter 3	Assignment 2
Family-based Studies Linkage analysis, Family-based association studies, TDT, FBAT	T Hoffmann	Austin Chapter 4	Assignment 3
Genetic Association Studies General principles; candidate gene studies.	T Hoffmann	Austin Chapter 5	Assignment 4
Genome-wide Association Studies (GWAS) Agnostic searches across genome for associated SNPs; multiple testing; Imputation; Consortia.	T Hoffmann	Austin Chapter 6	Assignment 5
Admixture Analysis Population substructure, admixture mapping.	L Fejerman	Winkler et al.,, Ann Rev Genom Human Genet 2010 11:65-89,.	
Gene-Environment Interactions and Epistasis Interactions; pharmocogenetics; less common & rare variants	J Witte	Austin Chapter 7	Assignment 6
Non-mendelian genetics Mitochondrial (MT), De novo, Parent-of-origin, Methylation	E Jorgenson	Austin Chapter 8	Assignment 7
Next Generation Sequencing Full genome and exome sequencing studies. Software for genetic analysis.	E Jorgenson	Austin Chapter 9	Assignment 8
Putting it all Together: Incorporating Molecular and Genetic Measures into Your Research Final Project presentations	J Witte, T Hoffmann	Austin chapters 10 and 11	Final Project Due **Friday 3/11 @ noon**

Molecular & Genetic Epidemiology

Distinction

- Molecular: molecular, cellular, and other biologic measurements, on disease [e.g., biomarkers - selenium in toe nails, proteins, hormones]

- Genetic: role of inherited factors in disease (encompassed within molecular)

Genetic Epidemiology

Focus of course

- Genetic epidemiology
  - Initially studied single gene disorders
  - Now more complex genetic disorders and environment
- Many designs same as epidemiology (e.g., case-control)
- Some specialized analysis methods.
- Population genetics increasingly important

Aims

- Detect genetic causes of disease
- Understand biological process
- Prevention strategies, lifestyle intervention
- Improved therapeutic strategies, personalized medicine

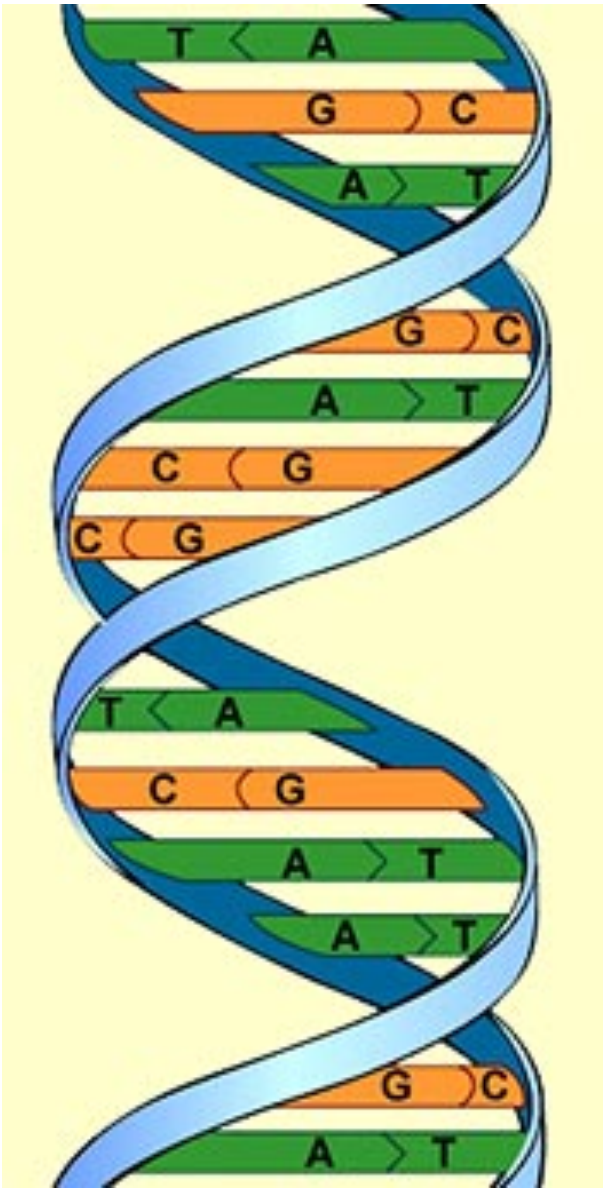
	Genetics	Epidemiology
Origins	Biological science of inheritance Experimental/breeding	Methodological sciences / public health Observational/analytic studies
Focus	Mechanisms of inheritance, often on rare diseases	Etiologic factors, primarily on common diseases
Tools	Family studies Laboratory methods Statistical models Population models	Observational studies Case-control studies Cohort designs Clinical trials
Goals	Understand mechanisms of inheritance	Understand distribution etiology, and progression of diseases
Overlap	Diseases of interest Population studies Ultimate goal of disease prevention	

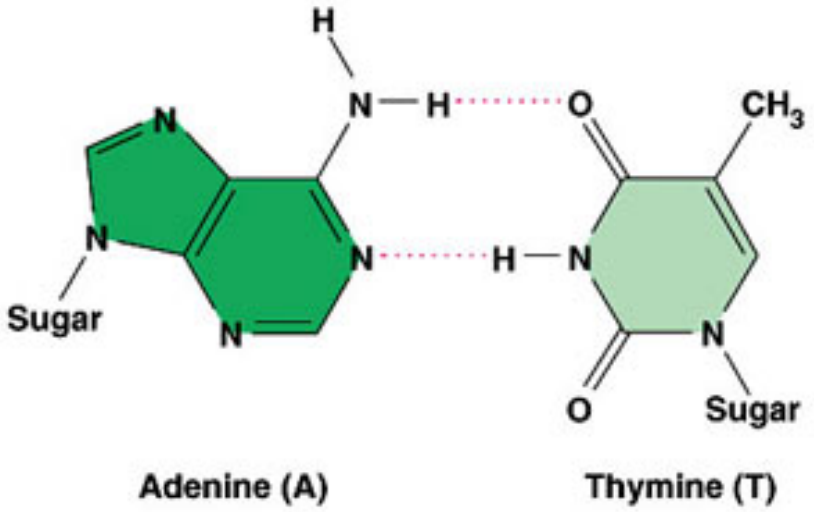
Austin, Ch 1, Table 1.1. Comparison of genetics and epidemiology.

Outline

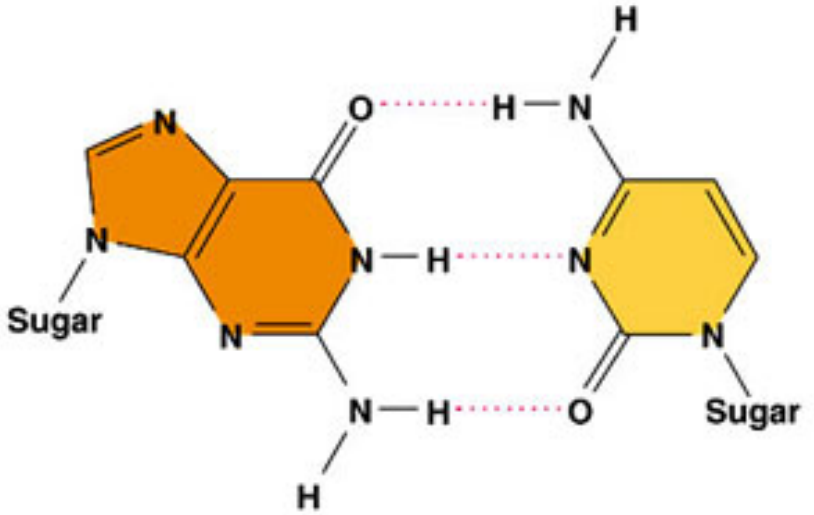
- Course structure
- **Overview of genetics**
- Overview of genetic epidemiology and the rest of the course

DNA





Adenine (A) Thymine (T)



Guanine (G) Cytosine (C)

Chromosome Bands

- Stain chromosomes so they can be seen by microscope
  - e.g., Giesma stain (G-banding).
- Appear as alternating bands
  - e.g., dark/G-band and light band.
  - Specific to phosphate groups of DNA.
  - Attaches to DNA regions with high adenine-thymine (A-T) bonding.

- With low resolution, few bands seen:
  - … p2, p1 centromere q1, q2, … (count out from centromere).
- With higher resolution sub bands seen:
  - … p12, p11 centromere q11, q12 …

Human Chromosome 21

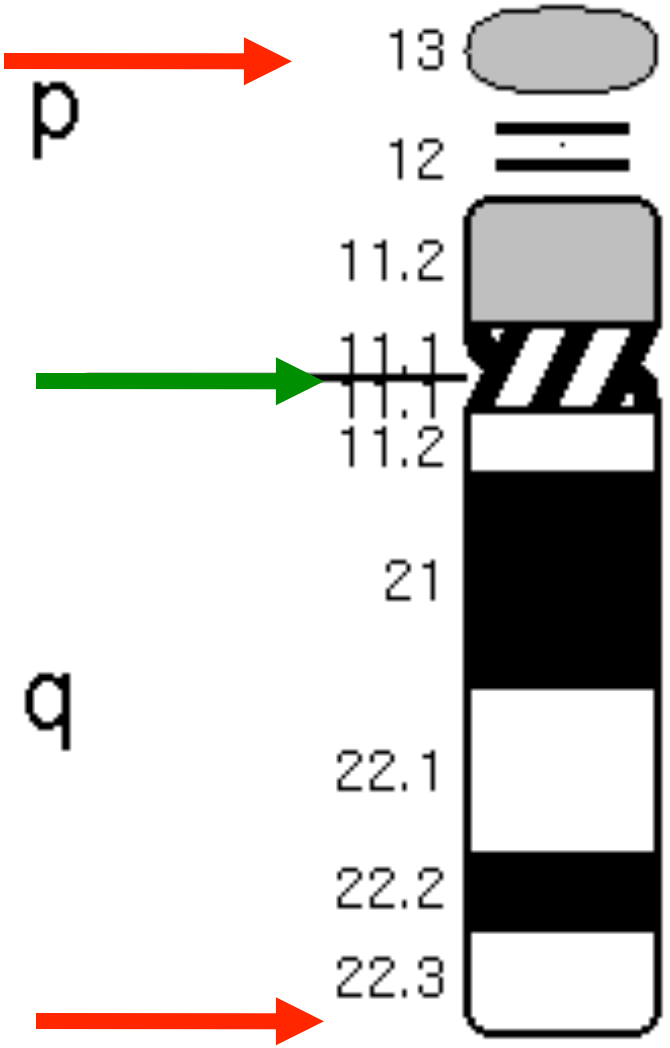
Telomeres

Centromere

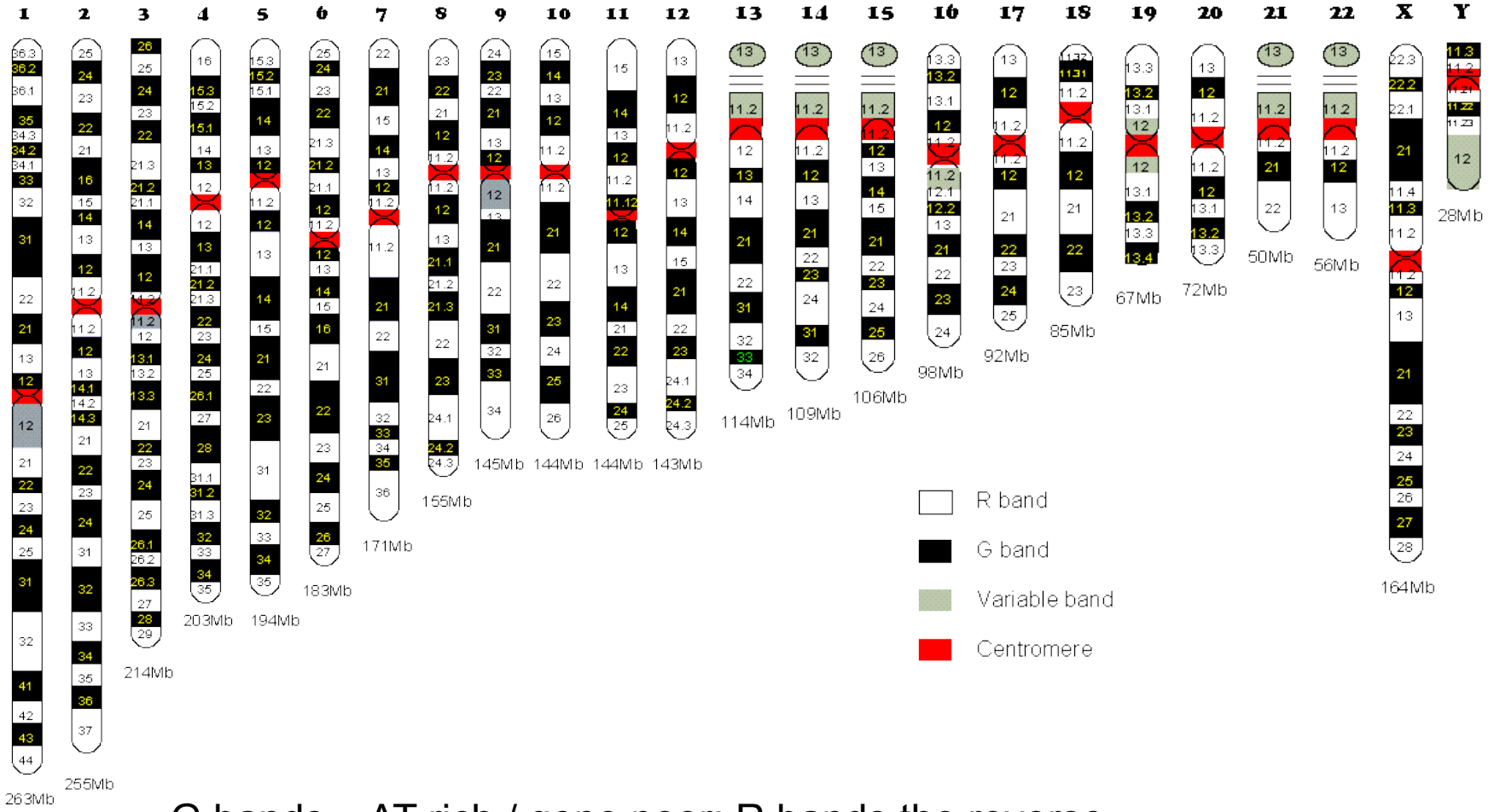
p: petit arm

q: queue (tail) or long arm

21q22.1 is pronounced twenty-one q two two point one



Human Chromosomes



G bands – AT rich / gene poor; R bands the reverse
Look for broad structural abnormalities from these stains

Variation in Genome

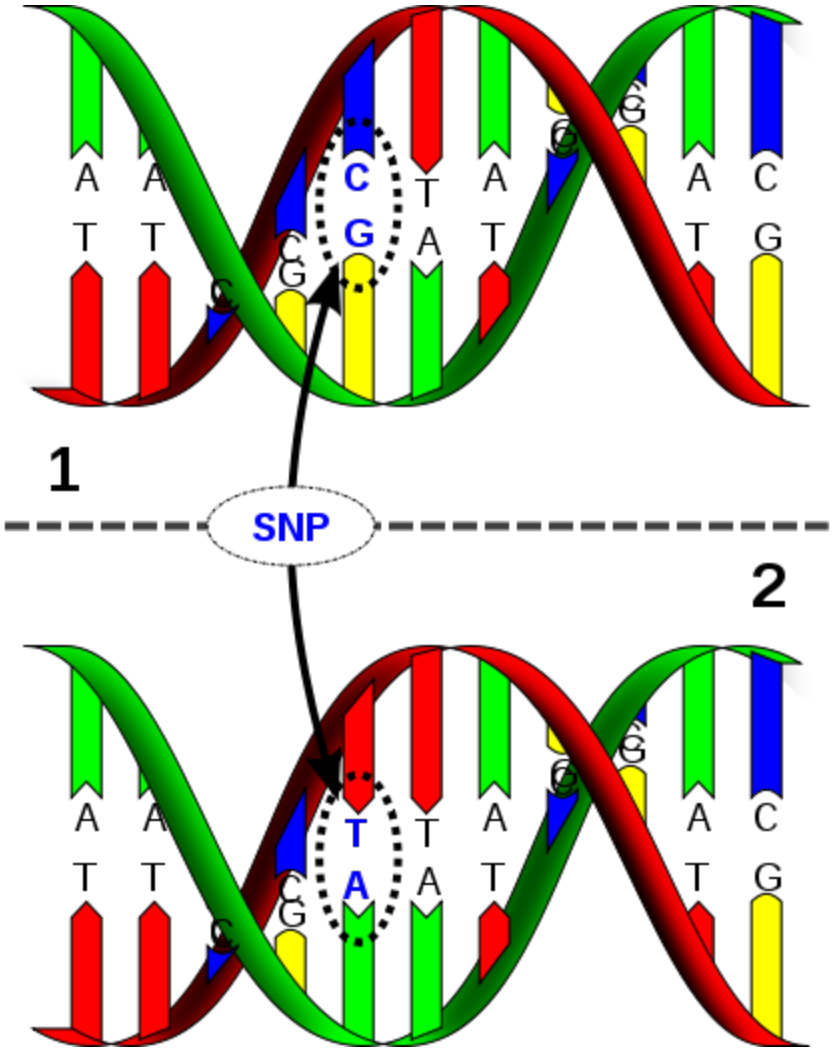
- Mutation
  - When event first occurs in an individual:
    - genetic change due to internal events (e.g., copy errors during cell division) or external agents (e.g., radiation, mutagens).
  - Can end with one generation, or be passed on (germline mutations)

- Polymorphism
  - Means “many forms”
  - Minor allele frequency (MAF) > 1%
  - Generated by old mutations (random, or fitness benefit).

Single Nucleotide Polymorphism (SNPs)

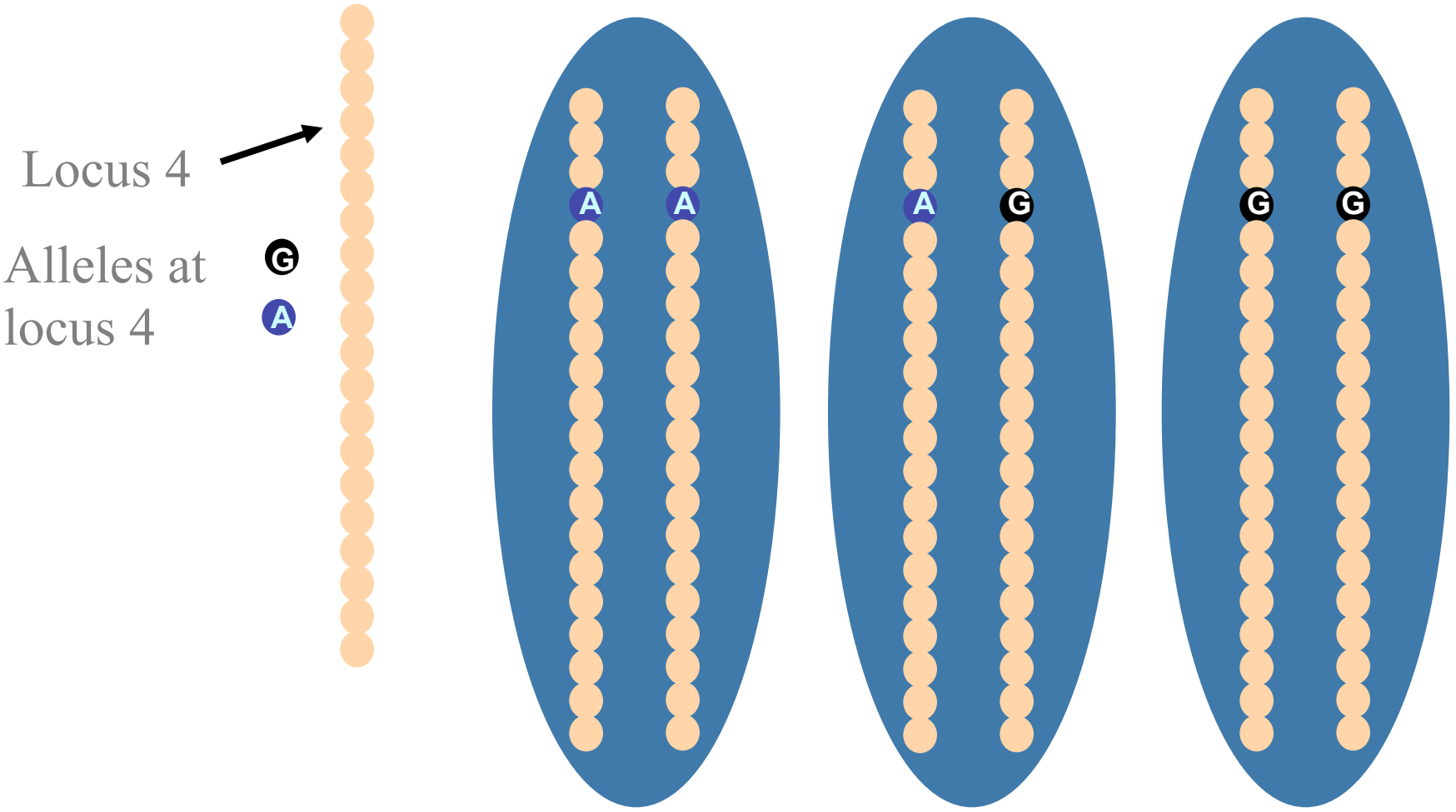
- Change a single DNA letter
- Most frequent genetic variant

- 1 per 300 base pairs
  - Common (MAF>5%)
  - Less common (1-5%)
  - Rare ‘variants’ (<1%)



David Hall

Genotypes



Locus: chromosomal location that's polymorphic.
Alleles: different variants @ locus

- Each somatic cell is diploid (two copies of each autosome)
- Thus 3 genotypes at locus 4 (use only one strand, often forward): AA, AG, GG

Types of Variants in Genes

Noncoding

Coding

- **Synonymous = no change in amino acid**
- **Nonsynonymous/nonsense = change to stop codon**
- **Nonsynonymous/missense = change amino acid**

e.g., **MTHFR C677T SNP**

Normal ('wild-type') allele

Gene sequence	.....GCG GGA GCC GAT.....
Protein Sequence	.....Ala Gly **Ala** Asp.....

Variant allele

Gene Sequence	.....GCG GGA GTC GAT.....
Protein Sequence	.....Ala Gly **Val** Asp .....

Human Genome Statistics

- Genome Reference Consortium Human Build 38 (updated 10/15)

- 3,547,121,844 basepairs (bp)

- 20,313 known protein coding genes
- Short variants (SNPs, indels, somatic mutations): 149,490,457

- Mutation rate ≈ 10^{-8} per bp per generation

- <u>In each person:</u>
  ~ 65 new mutations expected
  ~ 1 variant per 1.3K bps
  > 2M variants

http://www.ensembl.org/Homo_sapiens

Course Goals – Email to Tom & me

Name: John Witte

Department/Program: Epidemiology & Biostatistics, Urology, Cancer Center, and Institute for Human Genetics

Background in Genetics: Post-doc training in statistical genetics / genetic epi.

Background in Statistics: PhD in Epidemiologic Methods.

Research Interests: Methods for and analysis of the underlying genetic architecture of many diseases, esp. prostate cancer, birth defects, and pharmacogenomics.

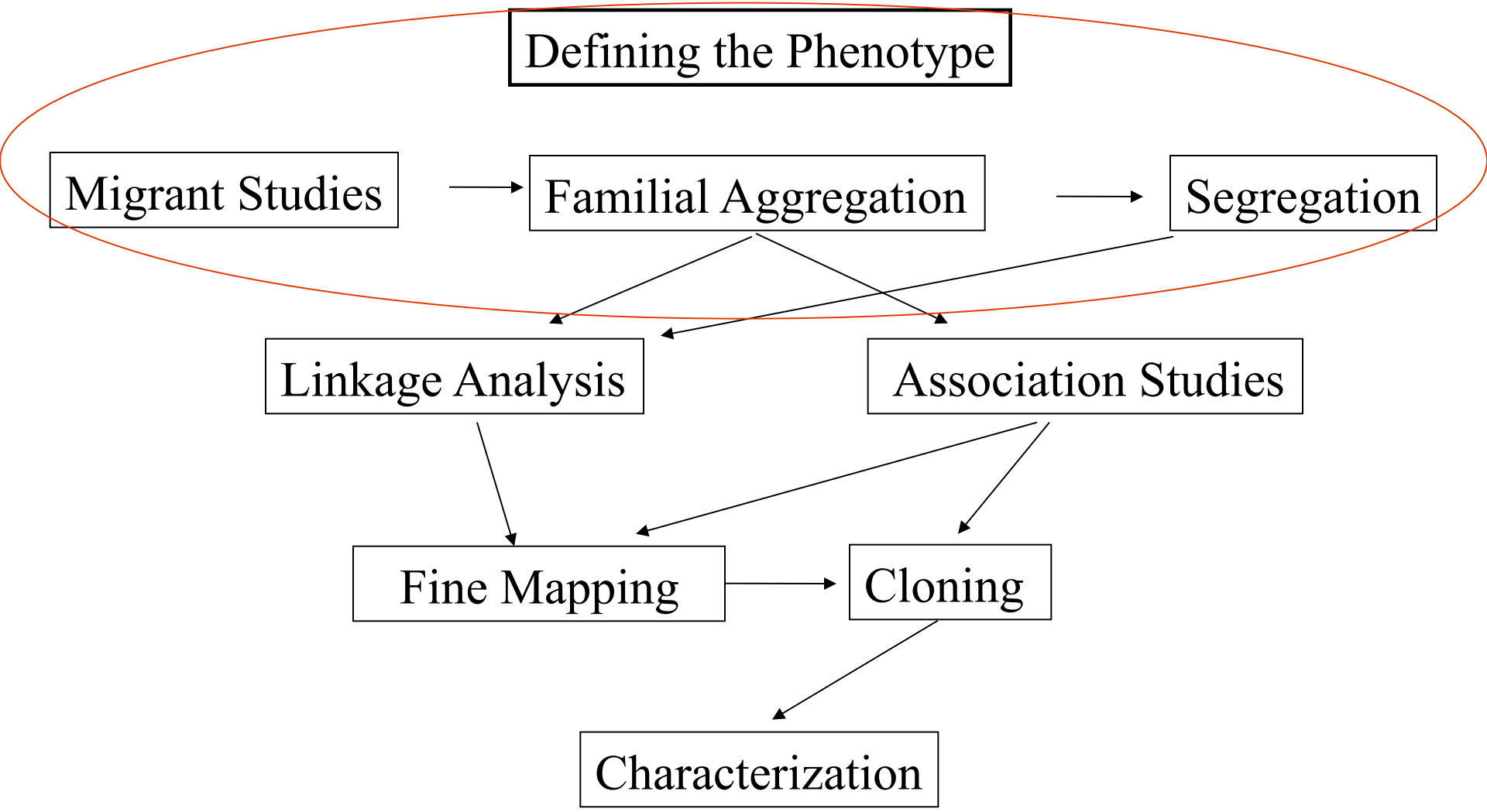
What do you want to get out of this course?

Be able to use genetic information in my research.

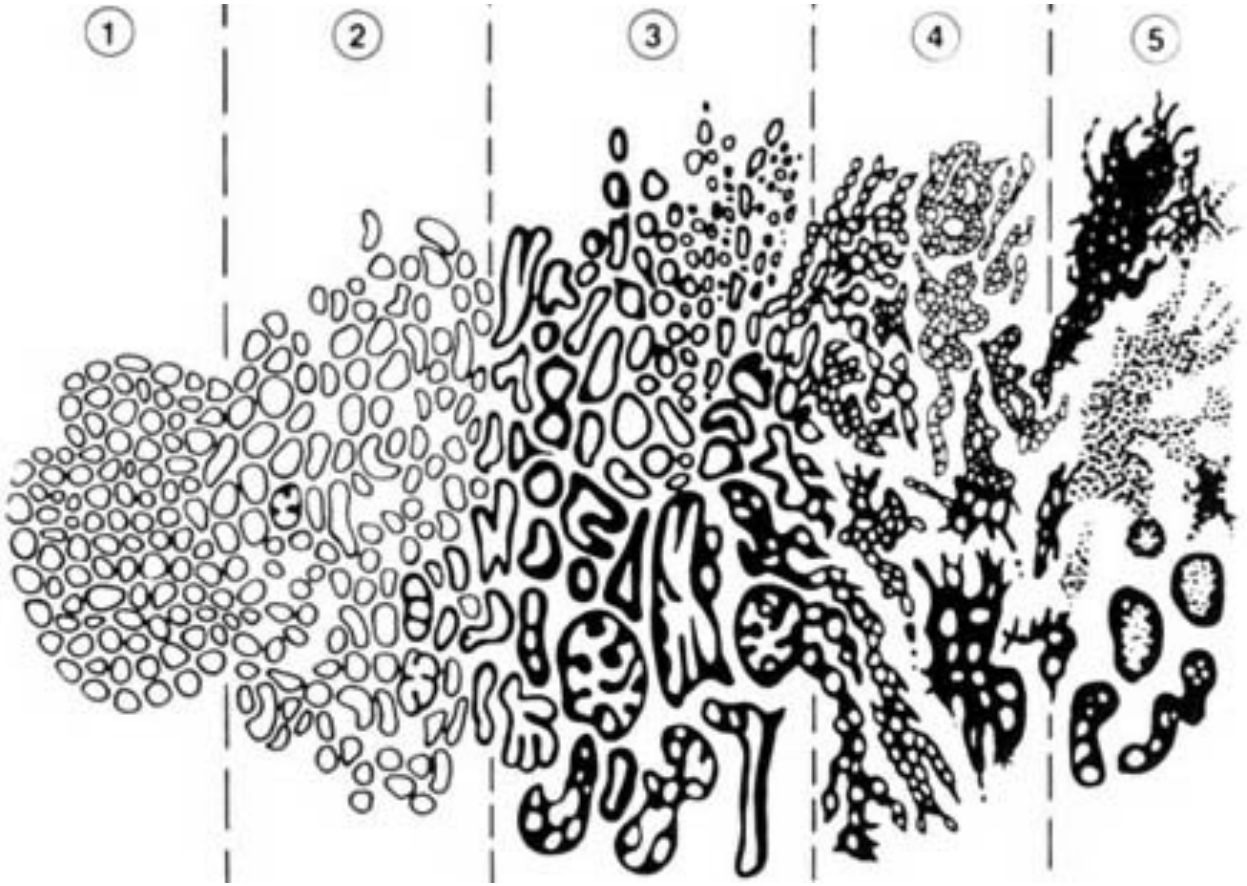
Outline

- Course structure
- Overview of genetics
- **Overview of genetic epidemiology and the rest of the course**

Process of Genetic Epidemiology

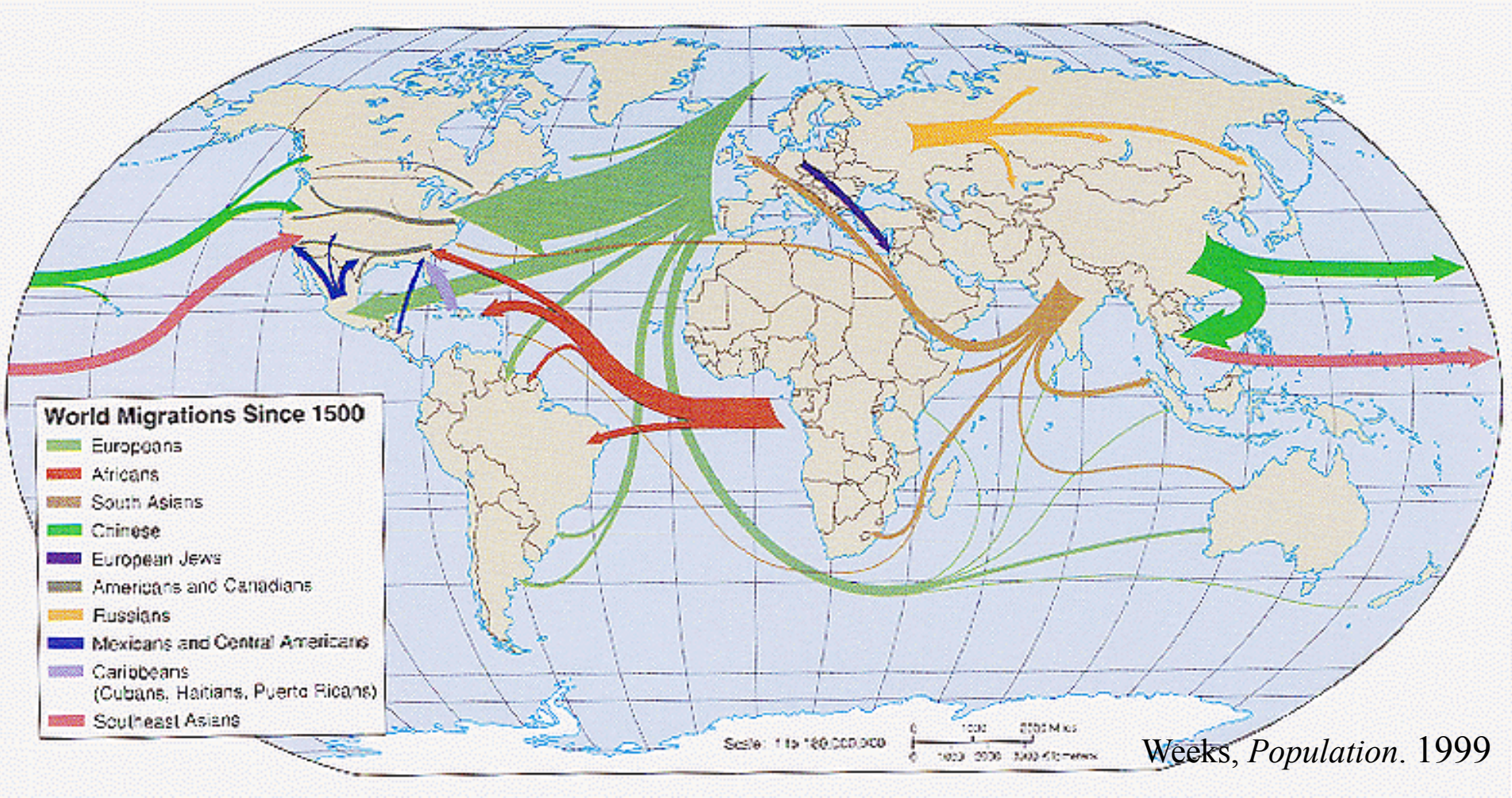


First: Define the Phenotype!



Gleason DF. In Urologic Pathology: The Prostate. 1977; 171-198.

Migrant Studies

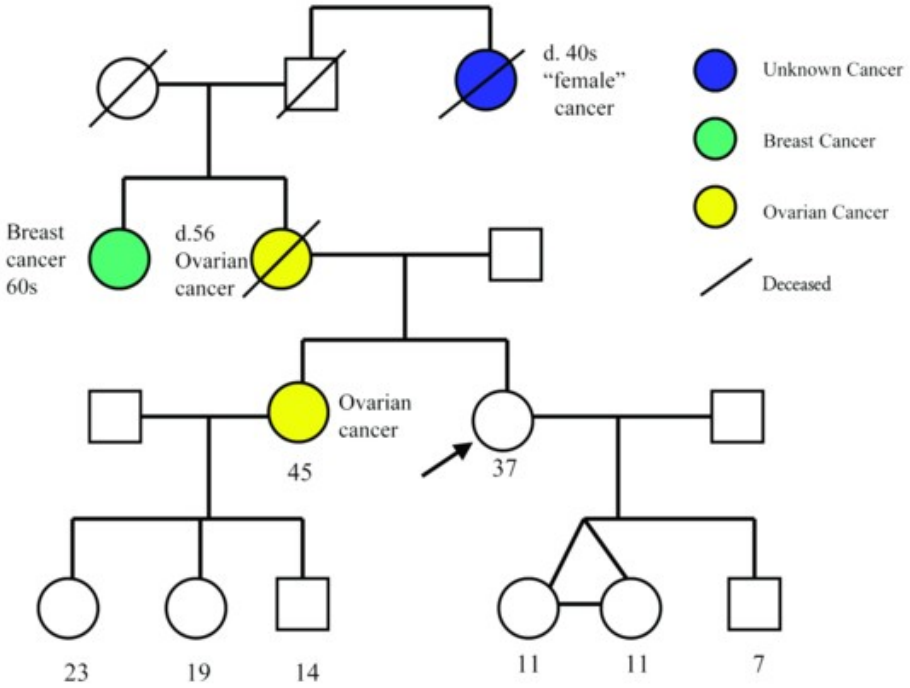


Studies based on the assumption that migrants carry a risk that reflects their country of origin, rather than the host country.

"Migrant Studies". Encycolpedia of Biostatistics, John Wiley and Sons 2005.

Familial Aggregation

- Does the phenotype tend to run in families?



Covered more in Austin Ch 2 & Lecture 2: Population Genetics, Modeling Genetic Inheritance

Analysis of Twin Studies

- Compare the disease concordance rates of MZ (identical) and DZ (fraternal) twins.

Twin 1

Twin 2	Disease	Yes	No
	Yes	A	B
	No	C	D

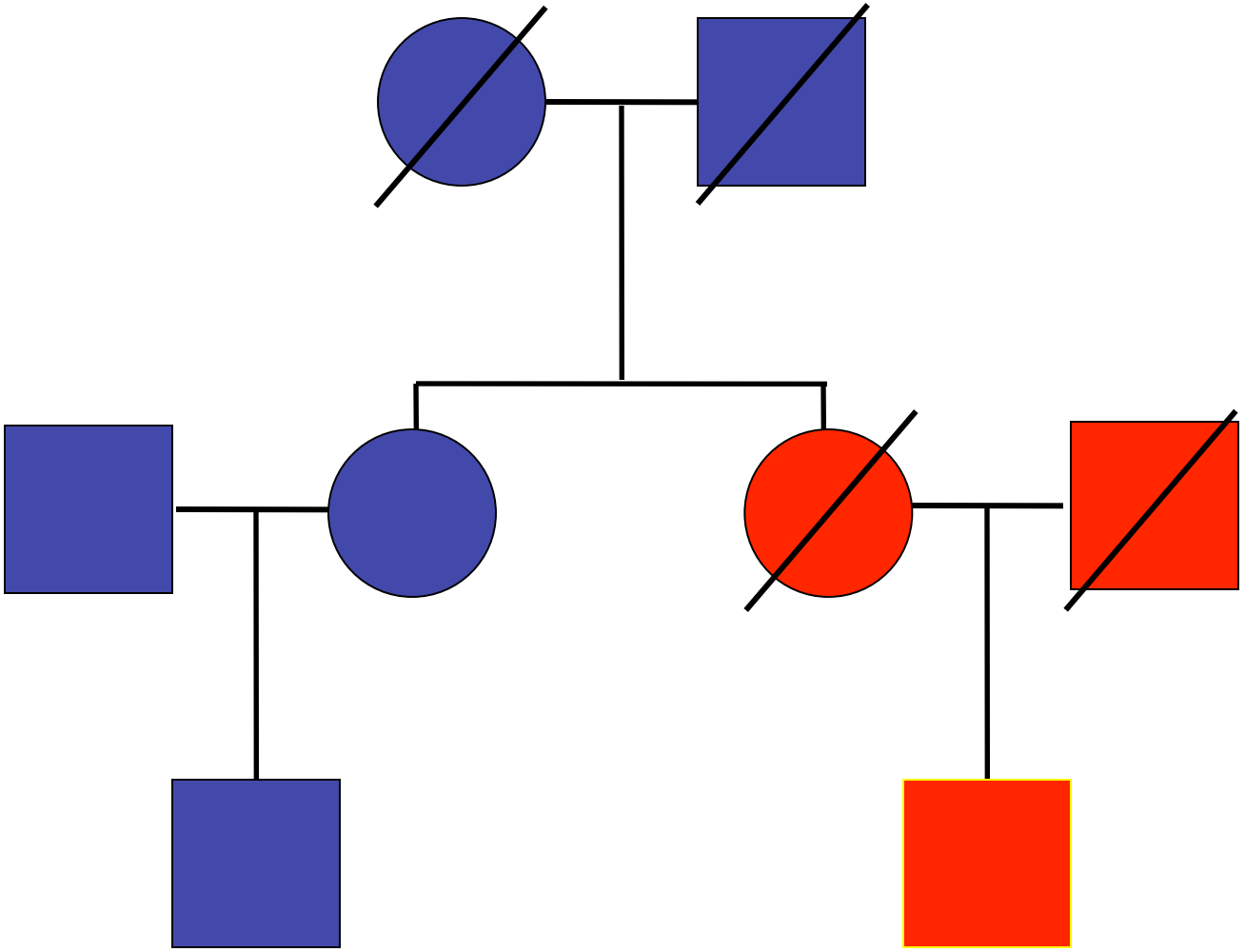
Concordance = 2A/(2A+B+C)

Then one can estimate heritability of a phenotype.

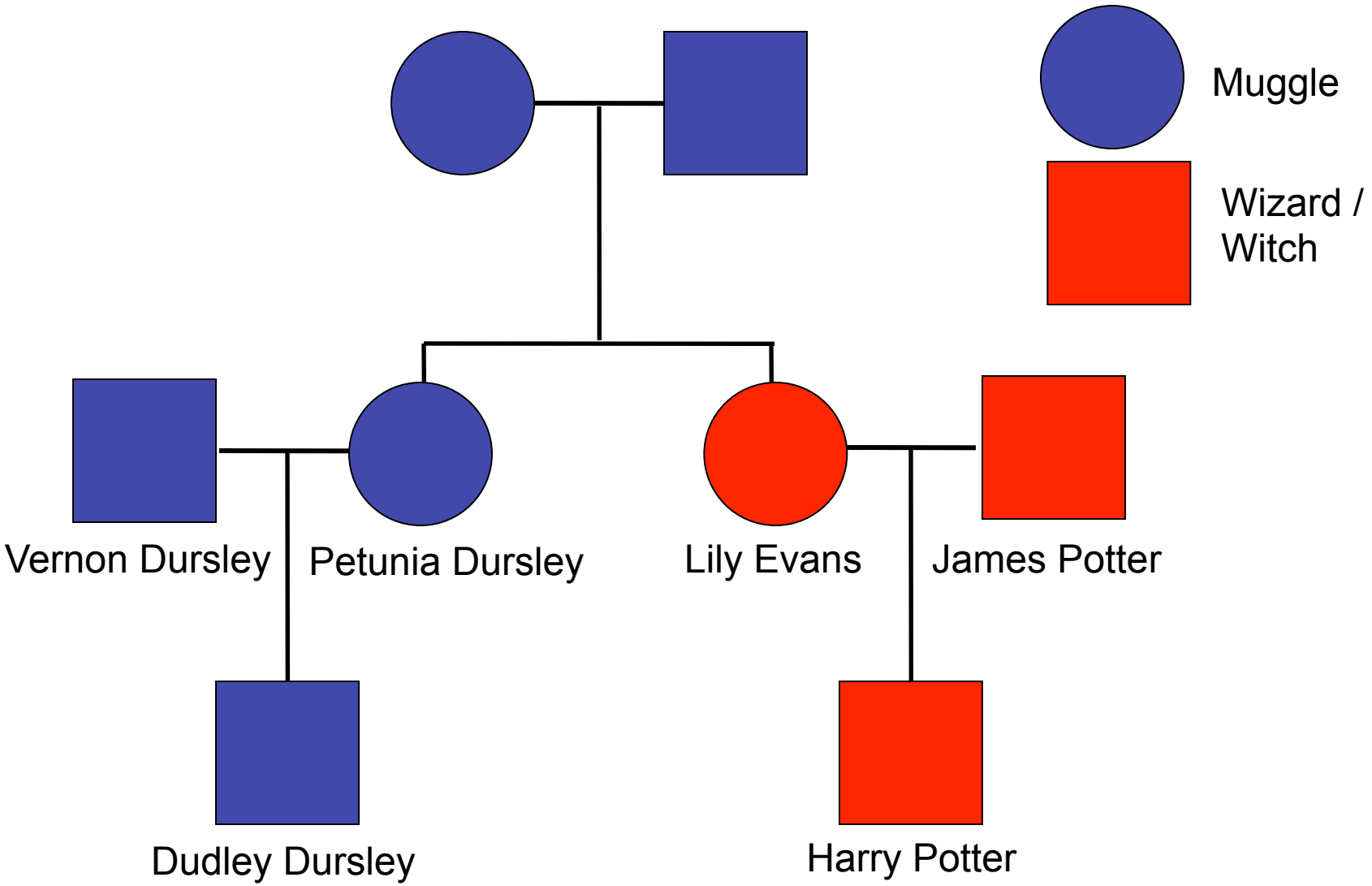
Models of Genetic Susceptibility

- Study families.
- Estimate ‘mode of inheritance’ & what type of genetic variant might be causal.
- Determine whether the disease appears to follow particular patterns across generations.
- Estimate whether variants are rare or common, etc.

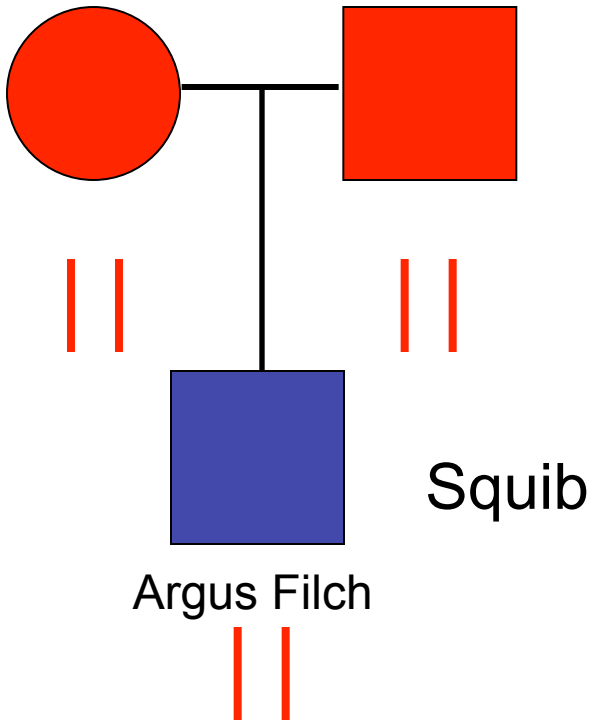
Segregation



Segregation: Harry Potter's Pedigree



Filch?

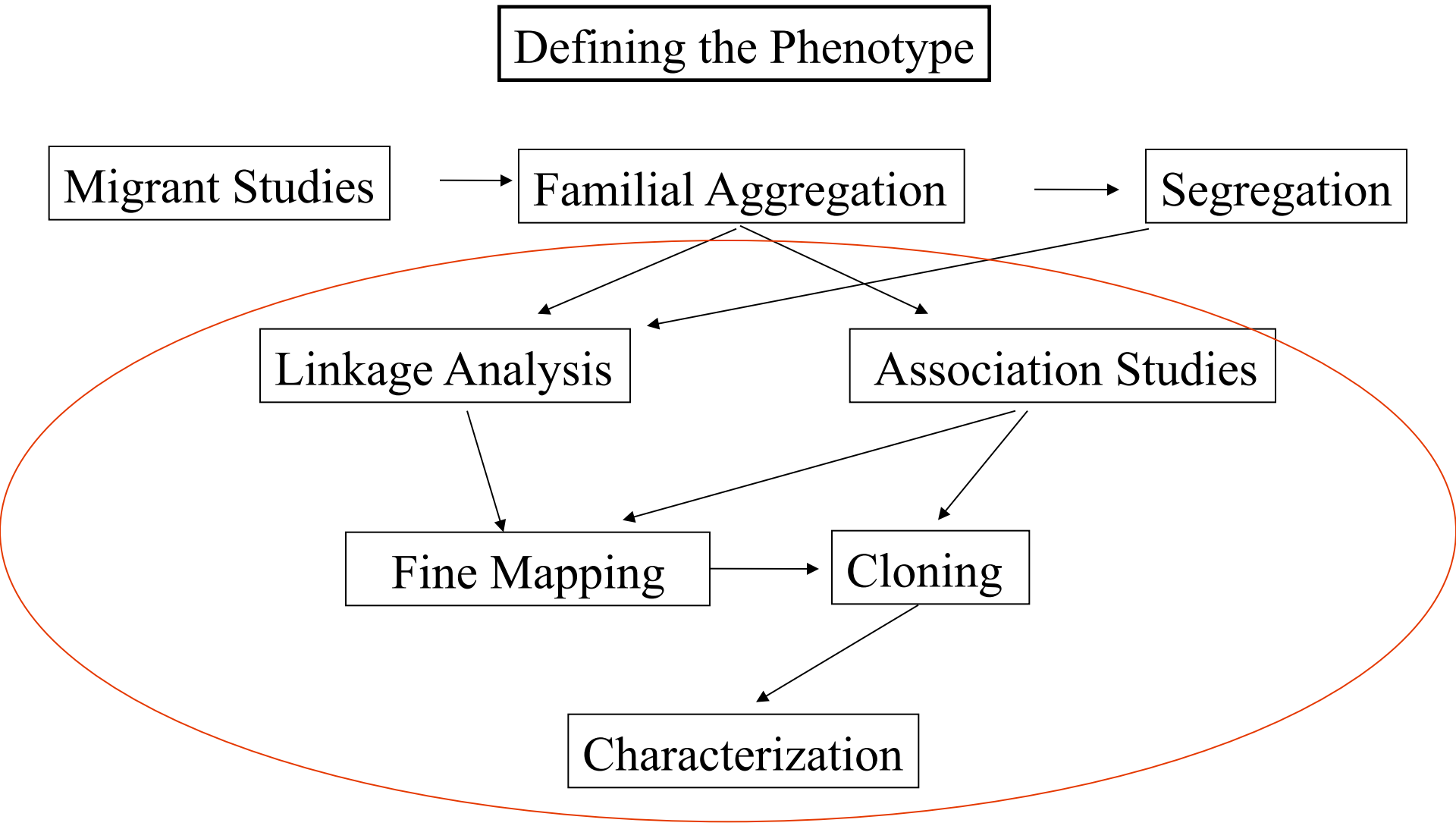


Segregation Analysis

- What is the best model of inheritance for observed families?
    - Dominant
    - Recessive
    - Additive
  - Disease allele frequency?
  - Magnitude of risk?

- Fit formal genetic models to data on disease phenotypes of family members.
- The parameters of the model are generally fitted finding the values that maximize the probability (likelihood) of the observed data.
- This information is useful in parametric linkage analysis, which assumes a defined model of inheritance.

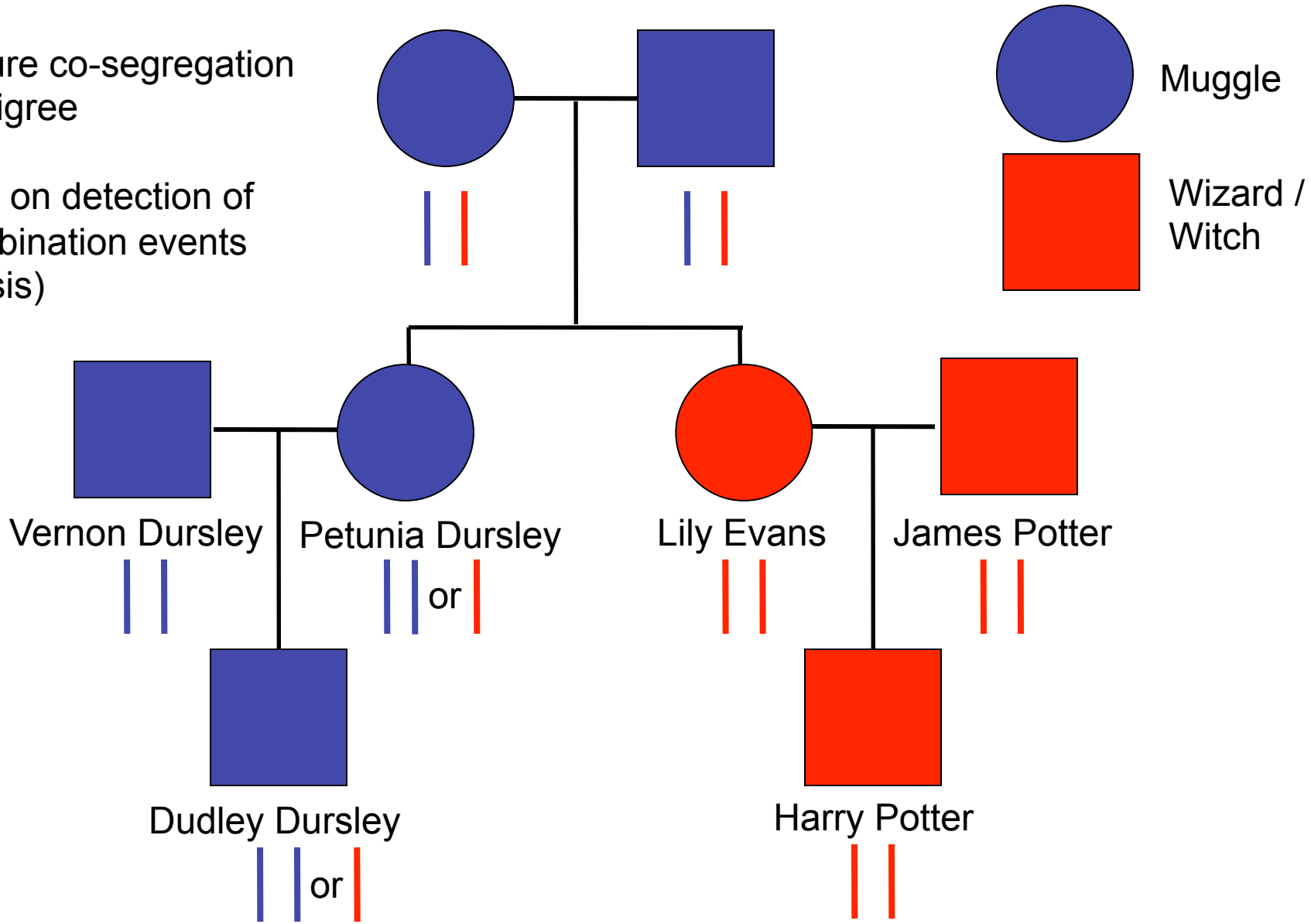
Process of Genetic Epidemiology



Linkage: Harry Potter's Pedigree

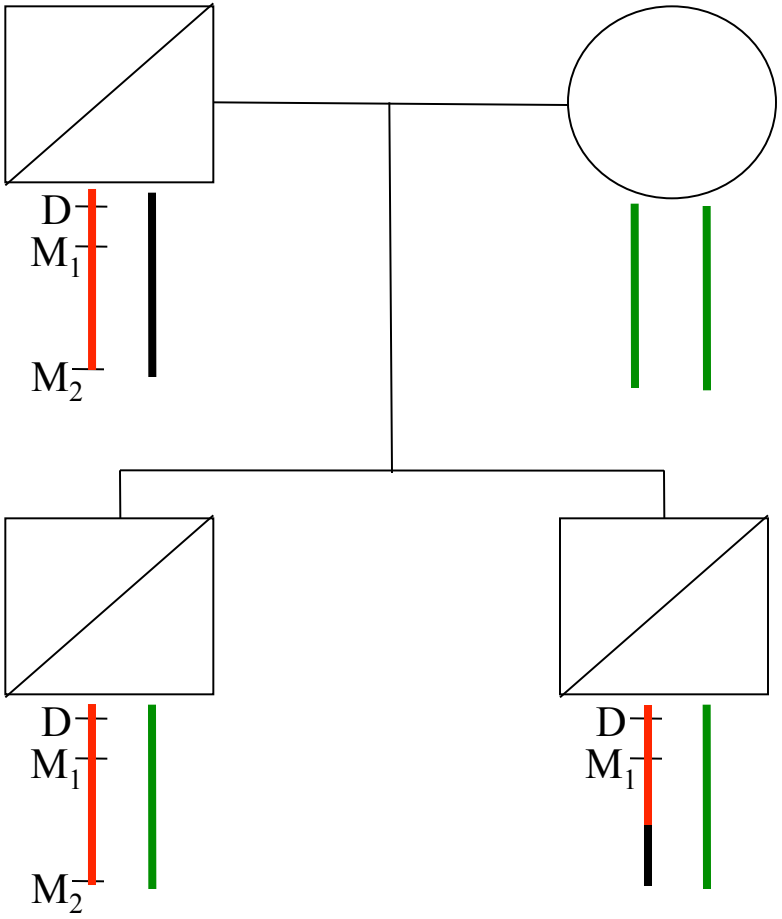
Measure co-segregation in pedigree

Based on detection of recombination events (meiosis)

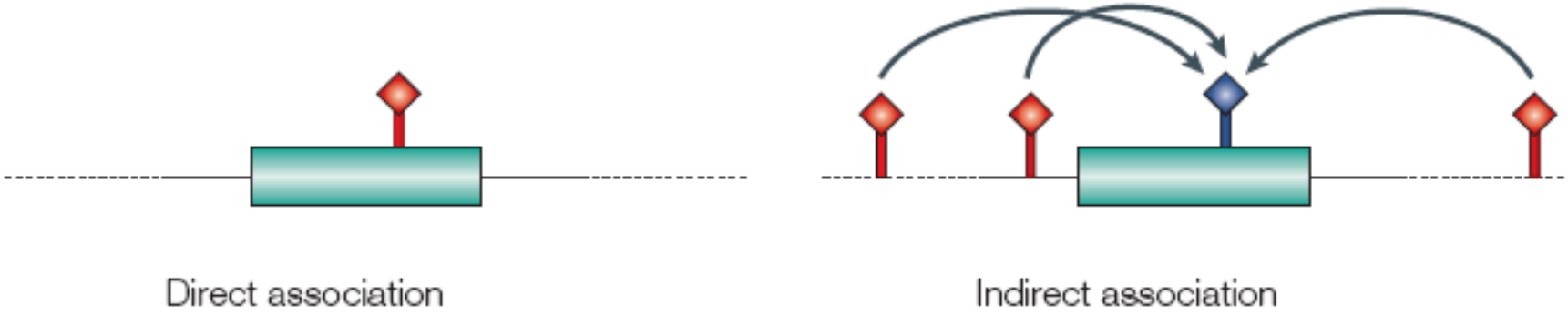


Covered more in Austin Ch 4 & Lecture 4: Family-based studies

Affected sib-pair Linkage



Linkage Disequilibrium (LD)

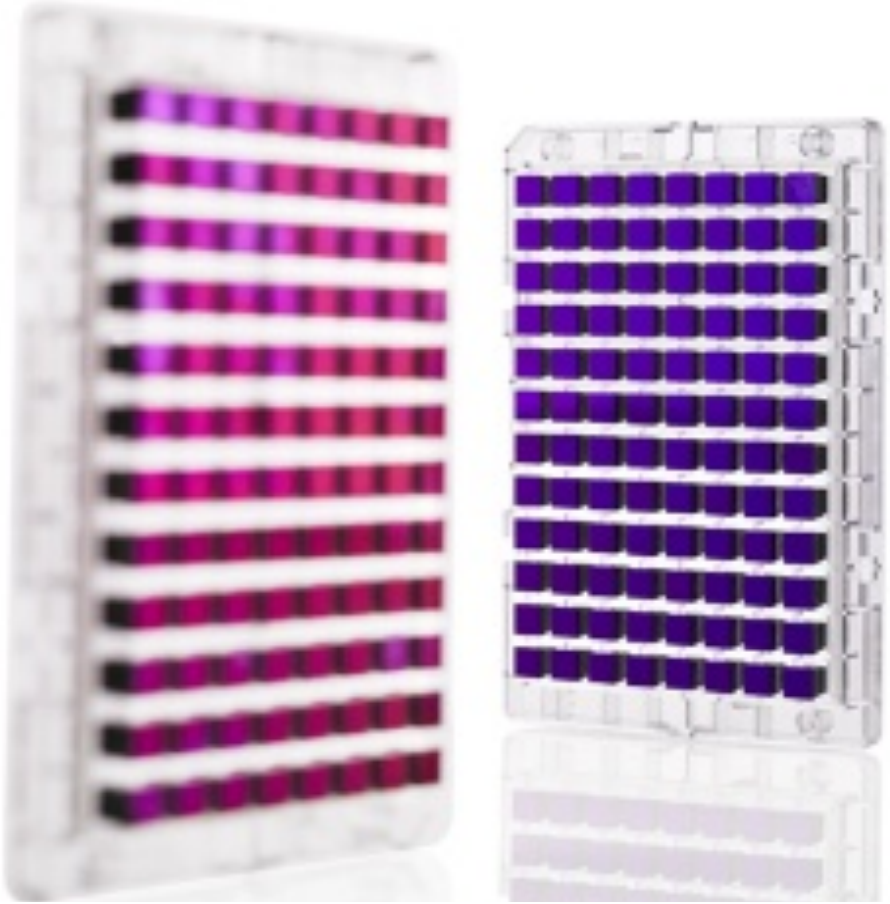


LD: non-random association of alleles at two or more loci that descend from a single, ancestral chromosomes.

Important fundamental concept *will come up repeatedly.*

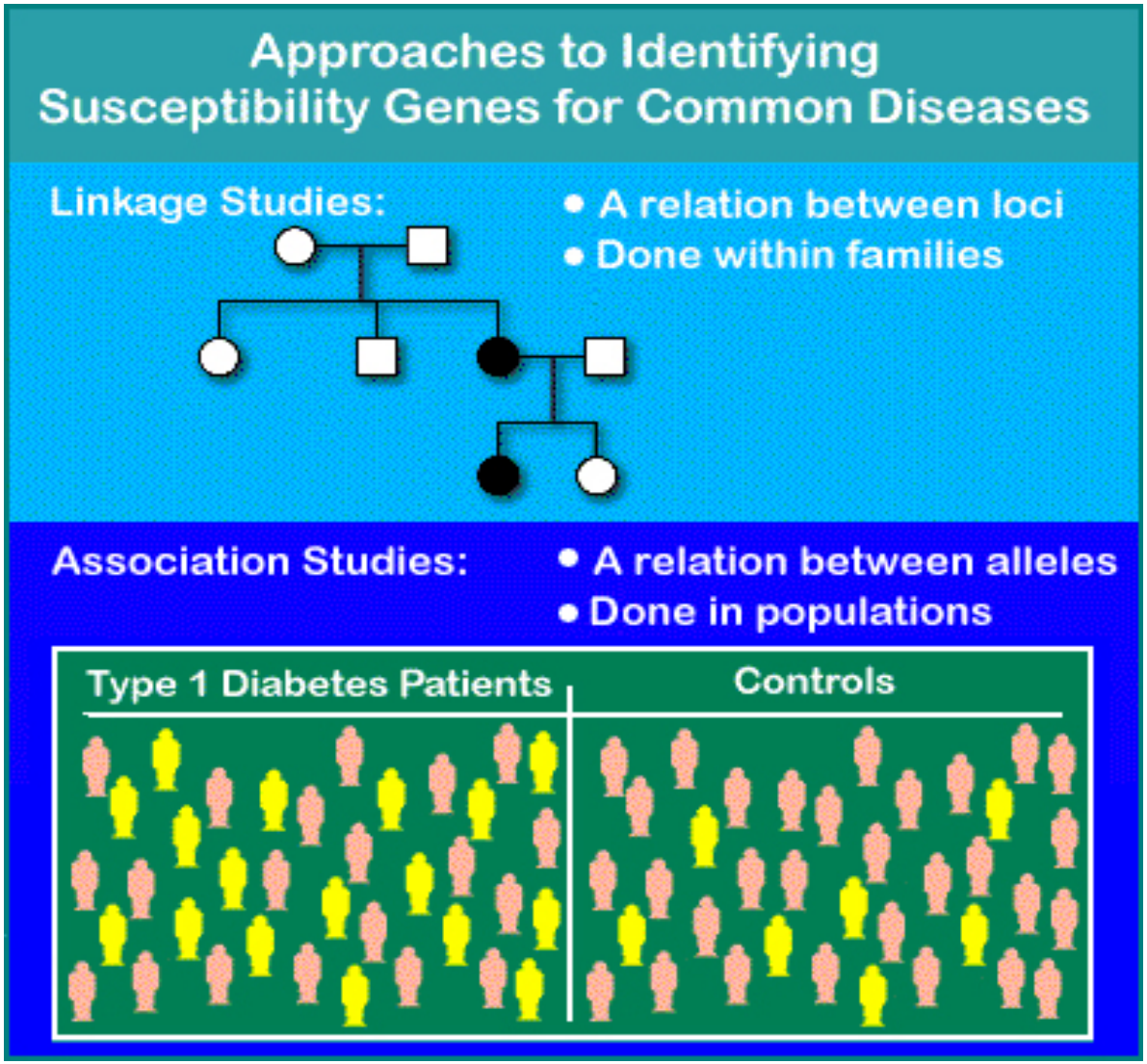
Hirschhorn & Daly, Nat Rev Genet 2005

Genotyping technologies



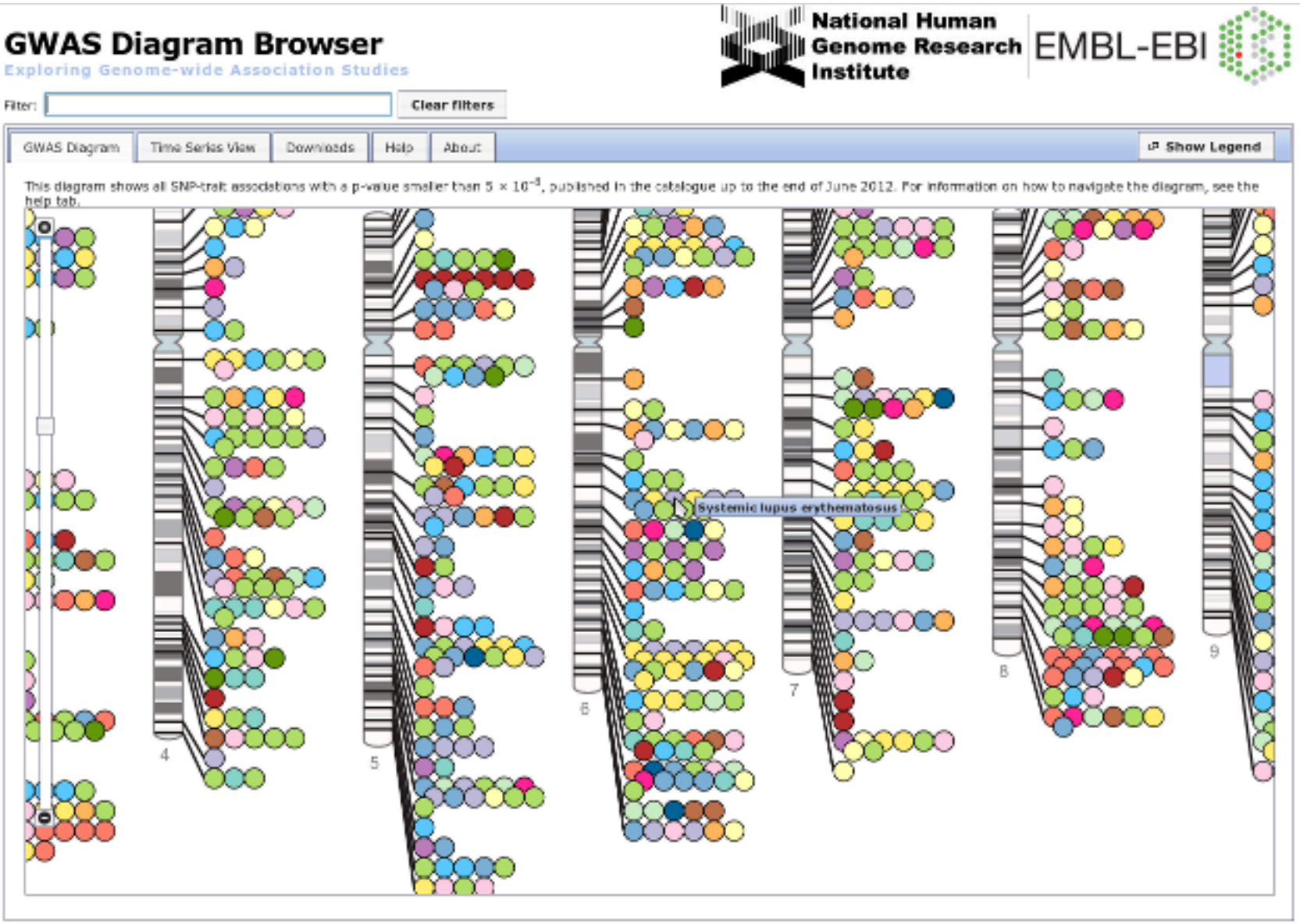
Covered more in Austin Ch 3 & Lecture 3: Mendel's laws and Molecular Genetics

Association Studies



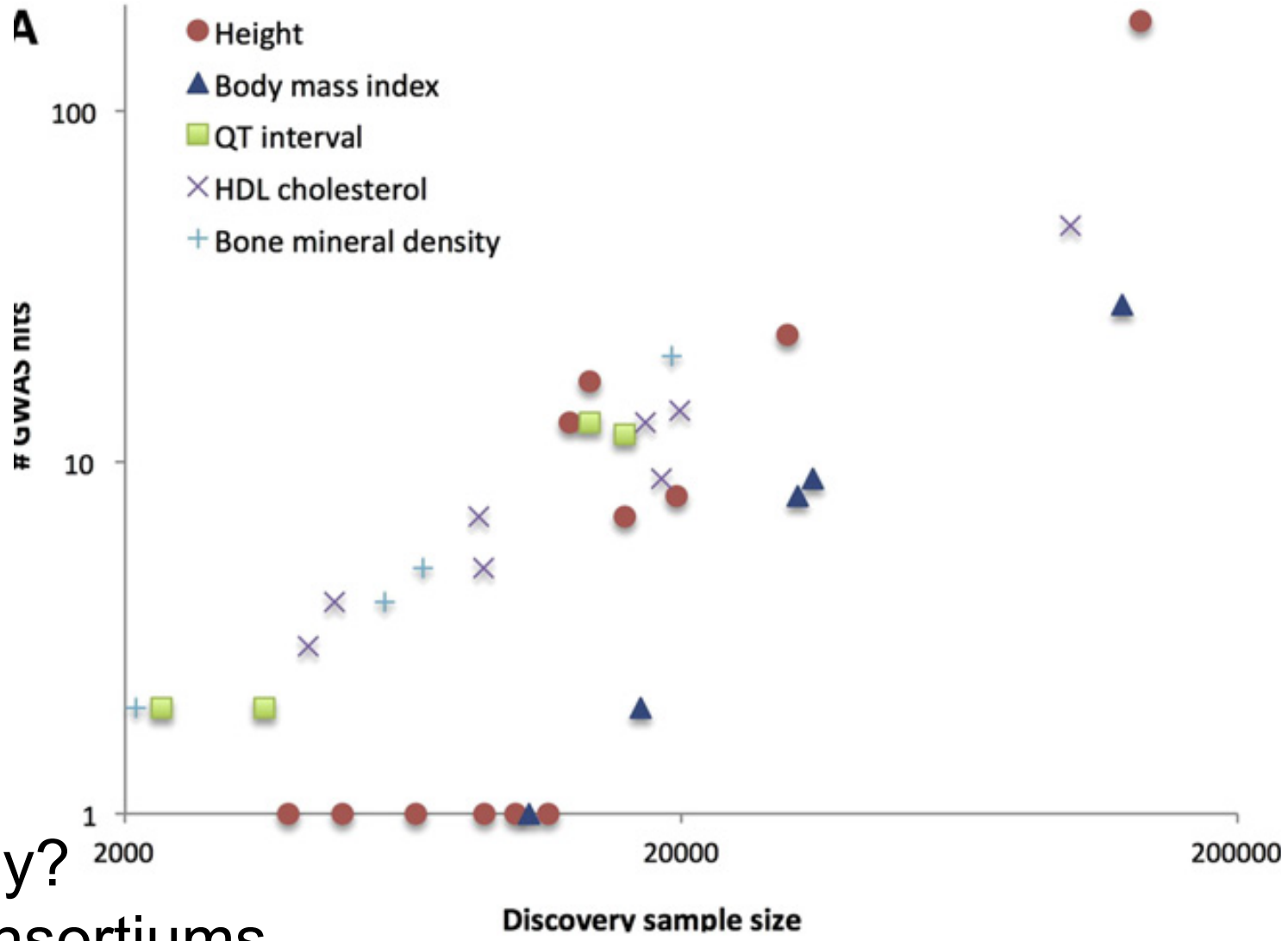
Covered more in Austin Ch 5 &
Lecture 5: Genetic Association Studies ROCHE Genetic Education (www)

Many GWAS hits: Polygenic Model



http://www.ebi.ac.uk/fgpt/gwas.

Sample Size Matters



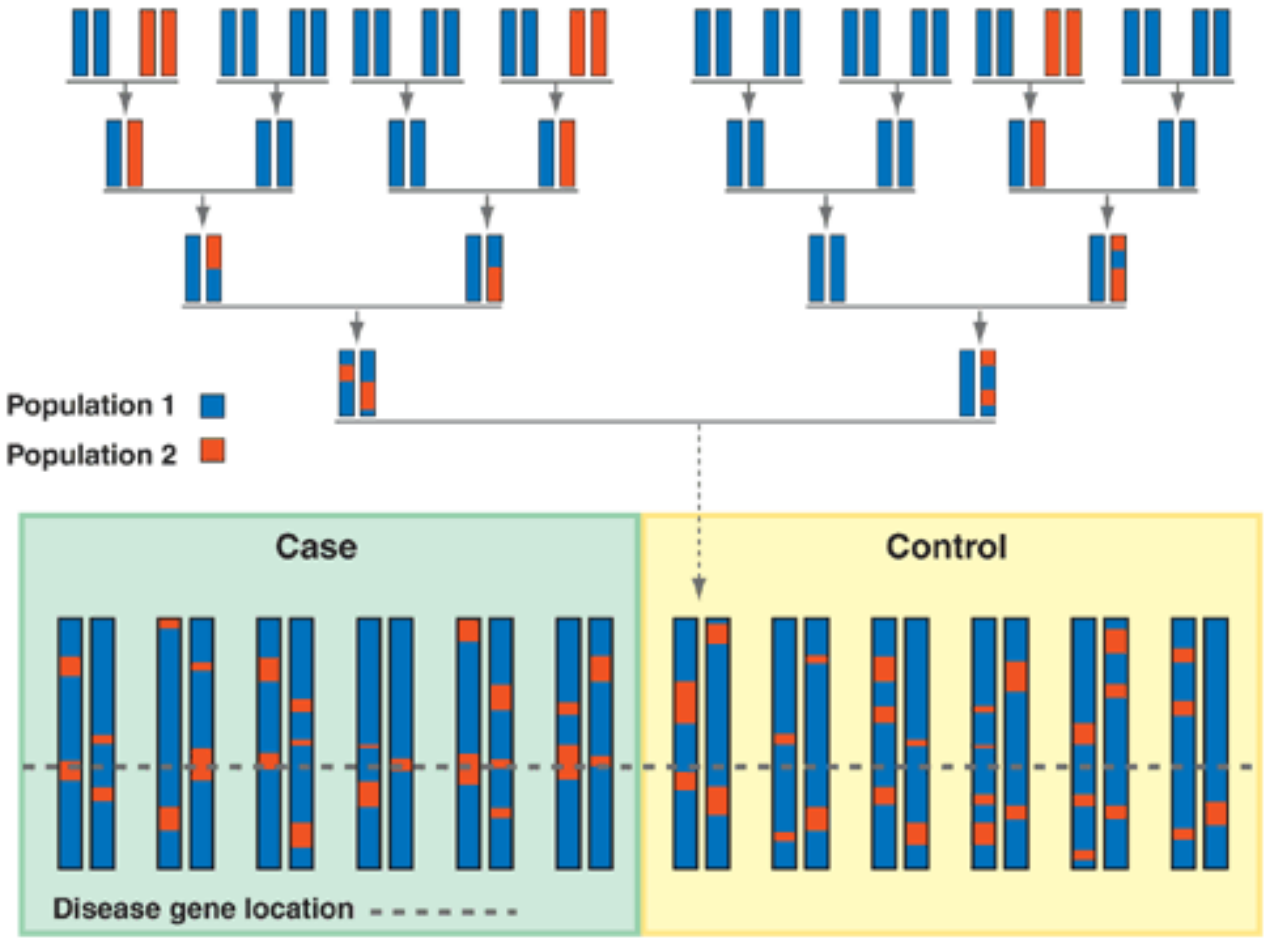
- Why?
- Consortiums

Visscher, AJHG 2012

Admixture Mapping

- Potentially powerful approach to searching for disease-causing genes
- Requires:
  1. Two populations with naturally occurring phenotypic and genetic differences.
  2. Recent gene flow between the populations (e.g., within 10 generations).
- Markers in the vicinity of the trait locus will also show excess ancestry from the population with the higher allele frequency

Admixture Mapping



Covered more in Lecture 7: Admixture Analysis

Nature Genetics 37, 118 - 119 (2005)

Summary of Main Mapping Approaches

	Linkage Analysis	Admixture	Association Study
Power*	Low	Moderate/High	High
# SNPs required	Low	Low	High
Sensitivity to genetic heterogeneity	Low	Moderate	High
Mapping resolution	Poor	Intermediate	Good

Next-Generation Sequencing (Ozzy Osbourne?)

cientists to map Ozzy Osbourne's
enetic code to find out how he survived
o much substance abuse

NICK KLOPSIS
ILY NEWS WRITER

nday, June 14th 2010, 1:39 PM

Sayles/AP

zzy Osbourne has had many issues relating to drug and alcohol abuse, so scientists are apping out his genetic code to find out how his body can take it.

u can't kill rock and roll, but it's not ually this hard to kill a rocker.

RELATED NEWS

"Sequencing and analysing individuals with extreme medical histories provides the greatest potential scientific value."

Nathan Pearson, Director of Research Knome

Covered more in Lecture 9: Next-generation sequencing

"Look, Mr. Osbourne, after studying your history, taking your blood, extracting your genes from the white cells, making them readable, sequencing them, analysing and interpreting the data using some of the most advanced technology available in the world today–and of course comparing your DNA against all the current research in the US National Library of Medicine, not to mention the 18th revision of the public human reference genome–I think I can say with a good deal of confidence why you're still alive."

I looked at him. He looked at me.

"Go on, then," I said. "Spit it out."

"Sharon," he replied.

TEDMED Talk:
https://www.youtube.com/watch?v=ZX0culYjU_A

Cloning and Characterizing a Gene

- Showing that it is clearly causal for disease.
- Numerous basic science experiments.
- Once genes are identified, molecular methods are used to determine the structure of the gene, ID regulatory elements, etc.
- Use epidemiologic studies to distinguish public health implications:
  1. Determine frequencies of causal alleles; and
  2. Characterize their effects—and interacting environmental factors—on disease rates.

Early Precision Medicine

"All the News That's Fit to Print"

The New York Times

Late Edition

VOL. CLVII . . No. 54,391 © 2008 The New York Times NEW YORK, THURSDAY, JANUARY 17, 2008 $1.25

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Cumulative Association of Five Genetic Variants with Prostate Cancer

S. Lilly Zheng, M.D., Jielin Sun, Ph.D., Fredrik Wiklund, Ph.D., Shelly Smith, M.S., Pär Stattin, M.D., Ph.D., Ge Li, M.D., Hans-Olov Adami, M.D., Ph.D., Fang-Chi Hsu, Ph.D., Yi Zhu, B.S., Katarina Bälter, Ph.D., A. Karim Kader, M.D., Ph.D., Aubrey R. Turner, M.S., Wennuan Liu, Ph.D., Eugene R. Bleecker, M.D., Deborah A. Meyers, Ph.D., David Duggan, Ph.D., John D. Carpten, Ph.D., Bao-Li Chang, Ph.D., William B. Isaacs, Ph.D., Jianfeng Xu, M.D., D.P.H., and Henrik Grönberg, M.D., Ph.D.

$300 to Learn Risk of Cancer Of the Prostate

A New Test Is Raising Treatment Questions

By GINA KOLATA

FED'S CHAIRMAN IS SAID TO BACK AID TO ECONOMY

TALKS WITH LAWMAKERS

Crucial Support Seen on Short-Term Package for Stimulus

By EDMUND L. ANDREWS and DAVID M. HERSZENHORN

WASHINGTON — Ben S. Bernanke, chairman of the Federal Reserve, has told lawmakers that he can support the cuts in spending and measures to stimulate the economy, even if they increase the budget deficit, provided the measures are quick and temporary.

Mr. Bernanke is to testify before the House Budget Committee on Thursday. Democratic lawmakers said he had told them that he would not comment on proposals to link a stimulus package with a permanent extension of President Bush's tax cuts. That is expected to disappoint Republicans who favor such a link.

Amid with growing evidence that the economy is slipping into a recession, Congressional Democrats and President Bush are trying to come up with a package that would put more money in Americans' hands within the next few months.

The Fed's willingness to give a nod to fiscal stimulus is important. Many lawmakers will not support action without the chairman's blessing, and the double blow of knowing that the Fed and Congress are considering must be carefully calibrated.

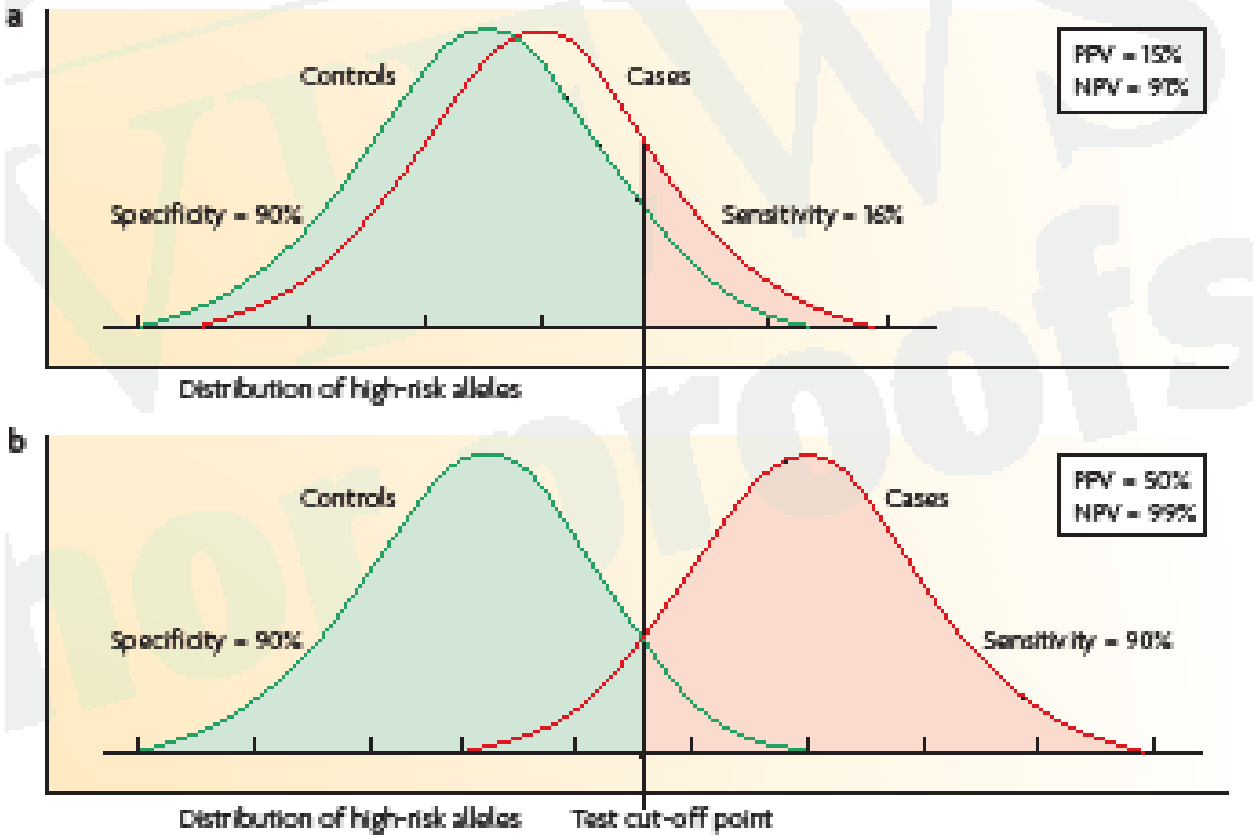
If Mr. Bernanke opposed [illegible]

About Giuliani

in South Carolina, Lesson Learned

Tactics

Large RR ≠ Good Prediction



Witte, Nat Rev Genet, 2009

Current Precision Medicine

VIEWPOINT

The Precision Medicine Initiative
A New National Effort

Euan A. Ashley, MRCP, DPhil
Departments of Medicine and Genetics, Stanford University, Stanford, California.

The announcement by President Obama of a precision medicine initiative created excitement in the medical community. The president referred not to personalized medicine but to "precision medicine," a term given profile by a recent publication from the National Research Council,[1] in which the authors explain that their tients is the hallmark of precision medicine. In fact, this particular drug development was further remarkable for the co-investment of the Cystic Fibrosis Foundation of $150 million in the development of the drug. That initial investment increased in value to $3.3 billion by the time of the sale of the royalty rights last year.

JAMA 2015: 313: 2119

DTC Genetic Testing

- Multiple companies marketing direct to consumer genetic ‘test’ kits.
- Send in spit or cheek swab.
- Array technology (Illumina / Affymetrix).
- Many results based on GWAS.

AACR Meeting Abstra...

Home Log In

HOME PRODUCTS ABOUT US COMPANY RESEARCH SHOPPING CART

ATLAS SPORTS GENETICS
ATHLETIC TALENT LABORATORY ANALYSIS SYSTEM

ATHLETIC TALENT LABORATORY ANALYSIS SYSTEM

Finding any great Olympic champion normally takes years to determine.

What if we knew a part of the answer when we were born?

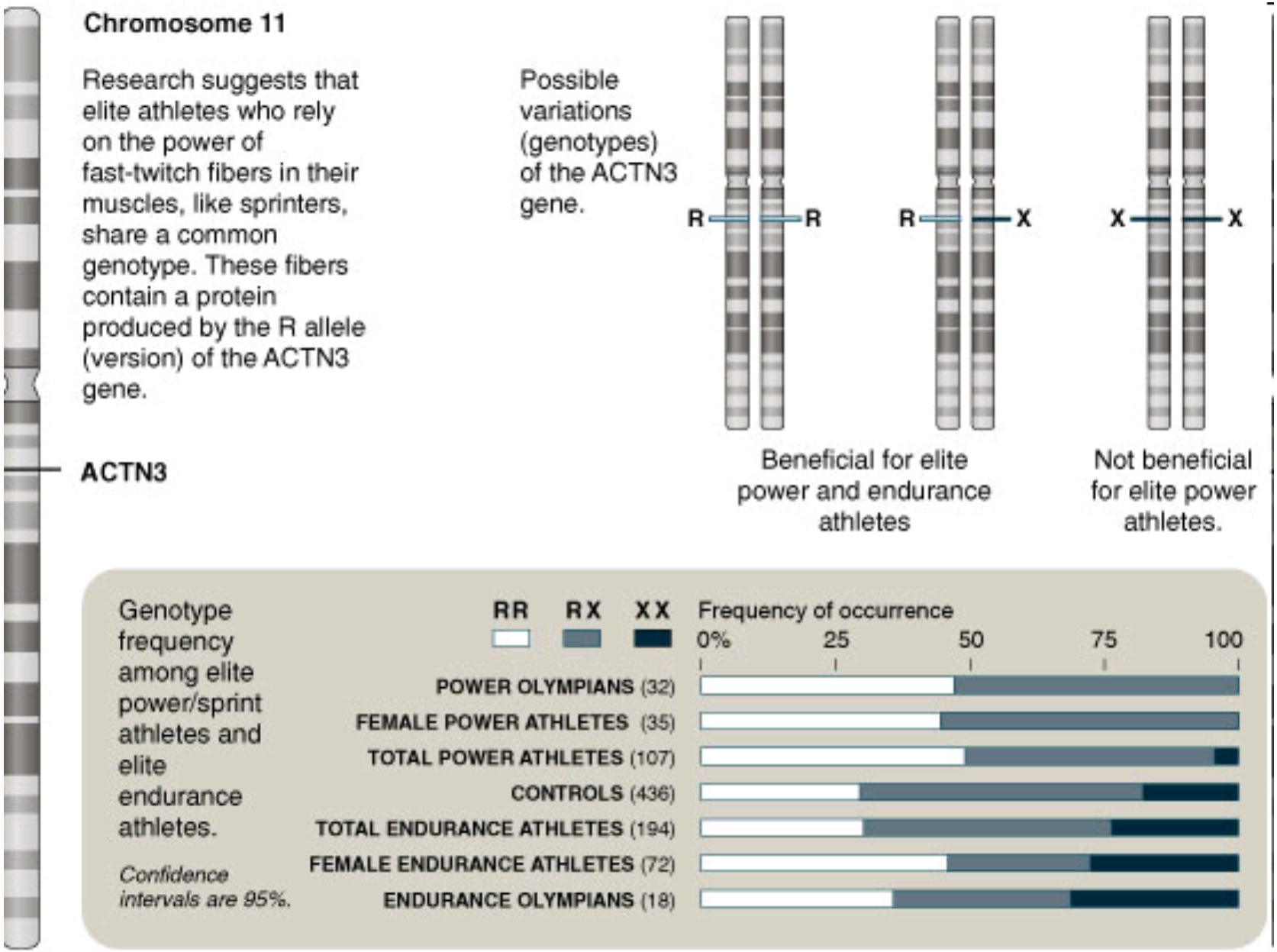
See How...

The New York Times - Born to Run? Little Ones Get Test for Sports Gene

GIST™ Sports Profile

Genetic Testing for Speed/Power and Endurance Events

Home :: Privacy Policy :: Conditions Of Use :: Disclaimer :: Links



Sources: Stephen M. Roth, Ph.D., University of Maryland; American Journal of Human Genetics

NY Times, 11/30/08

ATLAS SPORTS GENETICS

ATHLETIC TALENT LABORATORY ANALYSIS SYSTEM

larger image

ATLAS First

Recommended for ages 1 and up.
Description: Our *Atlas First* product is geared specifically at the youngest of athletes. Doing any type of performance based sport talent identification testing is very difficult below age 6 due to developmental levels of motor skills, strength and eye-hand coordination. *Atlas First* looks at only genetic markers, specifically the presence of ACTN3. Studies have found that individuals having the variant in both copies of their ACTN3 gene may have a natural predisposition to endurance events, one copy of their ACTN3 gene may be equally suited to for both endurance and sports/power event, neither copy of their ACTN3 gene may have a natural predisposition to sprint/power events. Knowing this information may be helpful, not in eliminating choices for sport activities but adding exposure to a host of team or individual sport events that may come easier to a young athlete.

The test is one of tool of many that can help children realize their athletic potential.

Other Products available through Atlas First

- Height monitoring charts
- Weight monitoring charts
- BMI charts
- Height Prediction Calculator (not genetic)

$149.00

Your Results

Who	Genotype	Genetic Result
Greg Mendel (Dad)	CC	Two working copies of alpha-actinin-3 in fast-twitch muscle fiber. **Many** world-class sprinters and some endurance athletes have this genotype.
	CT	One working copy of alpha-actinin-3 in fast-twitch muscle fiber. **Many** world-class sprinters and some endurance athletes have this genotype.
John Witte, Lilly Mendel (Mom)	TT	**No working copies of alpha-actinin-3 in fast-twitch muscle fiber. Few world-class sprinters have this genotype, but many world-class endurance athletes do.**


ATHLETIGEN
YOUR FOUNDATION FOR
PRECISION PERFORMANCE.

UNLOCK YOUR BODY

- Use the kit to collect and send your saliva to our labs
- Your Athletigen profile will be generated automatically
- Refine your training based on your DNA

BUY NOW

$199 US

All prices are in US Dollars. You will only be charged when your kit is shipped.

Genetic Taste Testing & Earwax Project

- Test for variants in genes that impact bitter taste receptors and earwax (wet or dry).
- What do you think your phenotypes are?

Additional Reading

From your literature reviews

JAMA Genetics Collections:
http://jama.jamanetwork.com/collection.aspx?categoryid=5664

Nature Reviews Genetics Study Design Series:
http://www.nature.com/nrg/series/studydesigns/index.html#d201601

COMMENT

Time for one-person trials

Precision medicine requires a different type of clinical trial that focuses on individual, not average, responses to therapy, says **Nicholas J. Schork**.

Nature, 2015: 520: 609