GGSB 2015 Prelim

Charles Czysz

September 2015

Required

   1-8: General Genetic Principles

   9-12: Mapping

   39-50: Study Design and Statistical Data Analysis

Choice between

   13-17: Genetic Architecture of Human Phenotypes

or

   18-28: Population and Evolutionary Genetics

   29-32: Molecular Mechanisms and Model Organisms in Human Genetics

or

   33-38: Gene Regulation and Human Phenotypes

A good answer would show in escalating order:

- Basic understanding via descriptions and definition of basic terms and concepts

- Knowledge of biology/literature via empirical examples of concepts in action,

- Engagement of critical thinking by highlighting well-known limitations or novel critiques of a concept or its common application

- Recognition of open problems and novel research opportunities.

Contents

General Genetic Principles

1. Explain the distinction between allelic heterogeneity, genetic (locus) heterogeneity, and clinical heterogeneity. Give examples of each.

Great review of this topic:
Genetic Heterogeneity in Human Diseases

Overview
Each of the following topics have implications in the type of studies which can or cannot be used. Overall, heterogeneity ensures that large-scale association tests or case-control studies will be poorly powered to detect causal variants or genes. See the review linked above for more detail.

Definition Allelic Heterogeneity - *In a given population, different mutations in the same gene result in a similar phenotype.*

Example
Cystic Fibrosis is caused by defective cystic fibrosis transmembrane conductance regulator proteins (CFTR). Many mutations in the CFTR gene can give rise to non-functioning proteins, which all lead to the same CF phenotype.

Two-thirds of all CF mutations are a 3bp deletion at position 508, resulting in a loss of phenylalanine. 1,500 other mutations also exist which lead to CF. However, this disease is haplosufficient..

Unknown alleleic heterogeneity can affect GWA results[1] when LD methods are used.

[1] http://hmg.oxfordjournals.org/content/11/20/2417.short

Sources
http://hmg.oxfordjournals.org/content/20/20/4082.short

Definition Genetic (locus) heterogeneity - *Mutations in different genes result in a similar phenotype.*

Example - *The BRCA1 and BRCA2 genes are a good example of how mutations in different genes lead to the same phenotype.*

Definition Clinical Heterogeneity - *Variability in clinical manifestations, or phenotypes, with the same underlying mutation/genetic disorder.*

Mutations in several genes can lead to familial hypercholesterolemia, high level of LDL cholesterol. Mutations in LDLR (LDL receptor), Apolipoprotein B, proprotein convertase subtilisin/kexin type 9 (PCSK9), and the ARH/LDLRAP1 genes can all lead to familial hypercholesterolemia.

2. What is the relationship between the inbreeding coefficient, kinship coefficient, and coefficient of relatedness? How are they calculated in pedigrees? Can they be estimated in the absence of pedigree information?

Definition Identity by descent - *Two alleles at the same locus that are descended from the same ancestral allele within the recent past. Can be 0, 1, or 2 depending on how many ancestral alleles shared between individuals.*

Definition Coefficient of Kinship - f_{xy}: *The probability that two alleles, one from X and the other from Y, are IBD.*
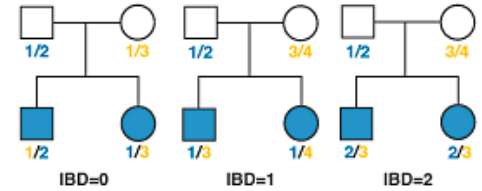
Figure 1: IBD Pedigree Example

Calculating F

♦ $F = \Sigma \, (\frac{1}{2})^n$

   ▪ n = number of connections between inbred offspring and each common ancestor

   ▪ Summed over all "loops" (or common ancestors)

First cousins
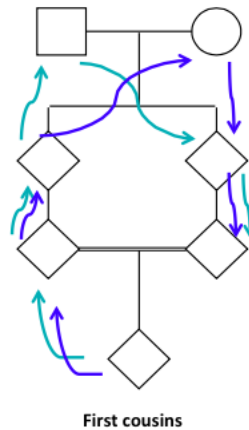
$$F = (\tfrac{1}{2})^5 + (\tfrac{1}{2})^5 = (^1/_{16})$$

Definition Coefficient of Relatedness - R_n, *the probability of sharing* $n \in \{0, 1, 2\}$ *IBD alleles. Mean relatedness,* $\bar{r} = 0 \times r_0 + 1 \times r_1 + 2 \times r_2$.

Relationship between Coefficients

$$f_{xy} = \frac{1}{2}\bar{r}$$

Figure 2: Relationship between Kinship and Relatedness coefficients

Relationship	Kinship coefficient	Coeffcient of relatedness
Self	0.5000	1.000
Monozygotic twins	0.5000	1.000
Parent-child	0.2500	0.500
Full siblings	0.2500	0.500
Half siblings	0.1250	0.250
First cousins	0.0625	0.1250
Unrelated	0.0000	0.0000

3. What are the key distinguishing characteristics of pedigrees segregating autosomal dominant, autosomal recessive, X-linked, Y-linked, and mitochondrial diseases?

Autosomal Dominant

On average, half of offspring of affected-nonaffected parents will be affected.

Autosomal Recessive

If both parents are carriers, then $\frac{1}{4}$th of offspring will show a disease phenotype.

X-linked Dominant

Affected men always pass disease to daughters. Affected women have $\frac{3}{4}$ths chance of passing disease to daughters and same chance to sons.

X-linked Recessive

Women can be either affected or carriers. Men always affected.

4. Explain the "non-Mendelian" concepts of uniparental disomy and imprinting. How would these be manifested in pedigrees and how are they demonstrated at the cellular or molecular levels?

Definition Mendelian Genetics - *Involves the "Laws" of random segregation, independent assortment, and dominance. Non-Mendelian genetics violate at least one of these assumptions.*

Definition Uniparental Disomy - *Offspring inherits a chromosome, or part of one chromosome, from one parent only. These offspring are euploid, but with an unequal contribution of genetic material from the parents.*

Uniparental disomy can manifest in two ways: inheriting both homologs from one parent (heterodisomy), or duplication of a single inherited chromosome (isodisomy).

Heterodisomy

Offspring inherits homologs from one parent only. Can occur during a nondisjunction error in Meiosis I, where one daughter cell inherits both sister chromatids while the other inherits none. This will lead to trisomy after fertilization, and heterodisomy if trisomy rescue occurs and results in the loss of the allele from the other parent.

Heterodisomy can also result from crossover events. In the parental generation, a cross between a balanced translocation carrier and a normal carrier can result in partial heterodisomy if the offspring inherits a crossover product and non-crossover product. However, this is an unbalanced gene load and is less prone to survival (10-15% incidence versus theoretical 50%).

In a Robertsonian translocation, acrocentric chromosomes fuse to create a single fusion chromosome. If this translocation occurs between homologs and is passed down to offspring, trisomy occurs (eg Trisomy 21). If, however, the non-fusion chromosome is lost by trisomy rescue, then heterodisomy occurs.

In a pedigree analysis, heterodisomies result in a parent passing their recessive condition to an offspring. Additionally, the parents of the affected parent are carriers.

Isodisomy

Isodisomy occurs when a single parental homolog is inherited and duplicated. This can occur after nondisjunction events in both Meiosis I and II. While the Meiosis I error gives 2 gametes which can result in heterodisomy, the other two gametes result in isodisomy if chromosomal duplication occurs. In this case, the other parental chromosome is duplicated.

During a meiosis II error, the sister chromatids fail to separate, leaving one product with two copies of the same chromosome in one gamete and none in the other. For the former gamete, trisomy rescue and loss of the other parent's chromosome leads to isodisomy. For the latter gamete, duplication of the other parental chromosome leads to isodisomy.

In a pedigree analysis, isodisomy should be suspected when an offspring shows a recessive disease when only one parent is a carrier.

This suggests that the recessive allele was inherited and duplicated.

Discovering Uniparental Disomy

Discrepancies between karyotypes of placenta and fetus.

Mosaicism of normal and trisomic cells (rescue occurs after first cell division).

Abnormal chromosome structure giving evidence of inheritance of Robertsonian product or unbalanced products. Usually seen by FISH.

Other DNA tests:

Test for STR (short tandem repeats) of known high heterozygosity, genotyping parents and offspring.

Definition Imprinting - *Imprinting results from the inheritance of epigenetic marks silencing an allele in a parent-of-origin specific manner.*

With imprinting, only one allele is expressed, and which allele depends on its parent-of-origin. When uniparental disomy of an imprinted allele occurs, the resulting expression can either go to 0x or 2x depending on if the silenced allele is the one doubled or not.

Sources

http://www.nature.com/gim/journal/v3/n3/full/gim200144a.html

5. What evidence is there for the presence of modifier loci? How is this related to the concept of epistasis and how is it distinct (or not) from polygenic and other models of inheritance?

6. What are distinctions among the concepts linkage, linkage disequilibrium, and association? Under what circumstances would each be preferable for genetic mapping? Consider both sample composition and types of diseases.

Definition Linkage - *Two loci tend to be inherited together due to lack of recombination.*

Linkage Mapping

Useful when samples are from large and informative pedigrees. Informative indicates that each allele can be assigned to

7. Define epistasis. Describe approaches that allow epistasis to be detected or quantified. Describe some biological mechanisms that can produce epistasis. Discuss the implications of epistasis for efforts to map the genetic causes of phenotypes. Discuss the potential implications of epistasis for the evolutionary process.

8. Define heritability. Describe methods used to quantify the heritability of a phenotype. Discuss the value and limitations of heritability as a descriptor of the extent to which a phenotype has genetic causes. Describe the "missing heritability problem" and its potential explanations.

Mapping

9. What is the difference between an odds ratio and relative risk? When would you use each and how might these relate to the concept of heritability in a genetic study?

Definition Relative Risk - *λ_R is the risk that an affected person acquires a trait divided by the risk for the general population.*

Consider a 2x2 contingency table testing a response (yes or no) based on group (placebo versus medicine). π_1 is defined as the probability of a "yes" response for the placebo group. π_2 is the probability of a "yes" response for the medicine group.

Group	Diseased	Not Diseased	Total
Placebo	π_1	$1 - \pi_1$	1
Medicine	π_2	$1 - \pi_2$	1

The relative risk for this table is given by

$$\text{relative risk} = \frac{\pi_1}{\pi_2}$$

Example Data:

Group	Diseased	Not Diseased	Total
Placebo	189	10,845	11,034
Medicine	104	10,933	11,037

Table 1: Adapted from Table 2.3 from *Introduction to Categorical Data Analysis* by Alan Agresti

In this example, $\hat{\pi}_1 = \frac{189}{11,034}$ and $\hat{\pi}_2 = \frac{104}{11,037}$. A small difference, but a relative risk of 1.83, meaning the placebo group is 83% more at risk than the medicine group.

Definition Odds - *Expected ratio of successes. Define p as probability of success, $odds_1 = \frac{p}{1-p}$. Thus, an odds of 3 indicates a probability of success of 0.75 and of failure 0.25.*

For the previous table, we can define the odds ratios for each group, giving $odds_1 = \frac{\pi_1}{1-\pi_1}$ and $odds_2 = \frac{\pi_2}{1-\pi_2}$, respectively.

Definition Odds Ratio - *A ratio of odds. For a 2x2 contingency table,* $OR = \theta = \frac{odds_1}{odds_2}$.

Odds ratios of 1 describe the null hypothesis that the group and response are independent. When $\theta > 1$, the odds of success for group 1 are θ times higher than for group 2. When $\theta < 1$, the odds of success for group 2 are $\frac{1}{\theta}$ times higher than for group 1.

Definition Log Odds - *As a fraction of positive values, the odds ratio is only lower-bounded by 0. Thus, the distribution is very skewed. Taking the log odds removes this skewedness since it is symmetric around its null value of 0.*

Relative Risk versus Odds Ratio

Recall that

$$RR = \frac{\pi_1}{\pi_2}$$

and

$$OR = \frac{\pi_1/(1 - \pi_1)}{\pi_2/(1 - \pi_2)}$$

Thus,

$$OR = RR \times \frac{1 - \pi_2}{1 - \pi_1}$$

and when the probabilities of success are both small, the RR and the OR are very similar.

Sources

Lander and Schork. "Genetic Dissection of Complex Traits" 1994. *Science*

10. What are the differences between quantitative and qualitative trait mapping, correlated phenotypes, multi-trait mapping?

Genetic Architecture of Human Phenotypes

Population and Evolutionary Genetics

Molecular Mechanisms and Model Organisms in Human Genetics

Gene Regulation and Human Phenotypes

Study Design and Statistical Data Analysis