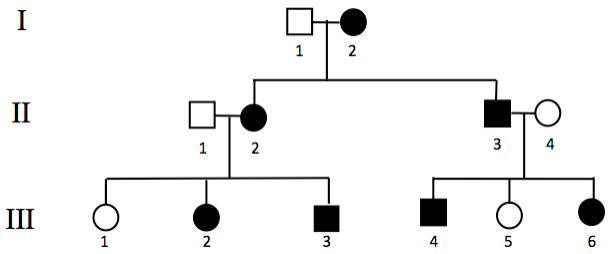
# Genetics and Genomics in Medicine Chapter 8 Questions

## Linkage Analysis Question

### Question 8.1




Affected members of the pedigree above have an autosomal dominant disorder, and cytogenetic analyses using conventional chromosome banding did not identify a disease-associated chromosomal abnormality. However, recent previous studies in other families mapped the disease locus to a region on the short arm of chromosome 17 and showed that four polymorphic microsatellite markers, *P, Q, R* and *S* are closely linked to the unidentified disease locus. (The order of the markers from most distal to most proximal is given by the following sequence: *P-Q-R-S*). When the same four markers were used to genotype each member of the family above, the alleles obtained were as listed below.

| | I-1 | I-2 | II-1 | II-2 | II-3 | II-4 | III-1 | III-2 | III-3 | III-4 | III-5 | III-6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***P*** | 1,4 | 1,2 | 1,4 | 1,2 | 2,4 | 1,4 | 1,4 | 1,2 | 2,4 | 2,4 | 4 | 1,2 |
| ***Q*** | 2,3 | 1,4 | 2,4 | 1,3 | 1,2 | 3 | 2,3 | 1,4 | 1,2 | 1,3 | 2,3 | 1,3 |
| ***R*** | 1 | 1,3 | 2 | 1,3 | 1,3 | 2,4 | 1,2 | 2,3 | 1,2 | 3,4 | 1,4 | 2,3 |
| ***S*** | 1,3 | 2,4 | 1,2 | 3,4 | 1,4 | 2 | 1,3 | 2,4 | 1,3 | 2,4 | 1,2 | 2,4 |

a) If we were to assume that the disease locus is also at the same 17p region in this family, what is the disease haplotype originating from I-2 (that is, the sequence of alleles at the four marker loci on the chromosome that has the disease allele)?
b) Has the disease haplotype been conserved in all family members?
c) Is it likely that the disease in this family is linked to 17p?
d) If the disease locus were to be at 17 p in this family, what do you deduce about the position of the disease locus with respect to the marker loci?

# Multiple Choice Questions

## Question 8.2

The same genomic DNA samples recovered from family members in Question 8.1 have been genotyped with an additional polymorphic marker T that has been identified in the same region of 17q. The alleles recorded were as follows:

| | I-1 | I-2 | II-1 | II-2 | II-3 | II-4 | III-1 | III-2 | III-3 | III-4 | III-5 | III-6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***T*** | 2,3 | 2 | 2 | 3 | 2 | 3 | 2,3 | 2 | 2 | 3 | 2,3 | 3 |

Which of the following statements is most likely to apply?

a) The genotypes have been recorded wrongly, possibly because samples were switched, and there is a need to repeat the genotyping.
b) There is something inherently unstable about marker *T*; there is no point in repeating the genotyping, and a different marker on 17q should be used instead.
c) The assumed biological relationships shown in the pedigree are wrong.
d) The disease-causing mutation has affected the marker locus.

## Question 8.3

Which, if any, of the following statements is true?

a) The heritability of a disease can vary from one human population to another, but remains fairly constant within any individual population.
b) The risk ratio in complex diseases is high compared to in monogenic disorders.
c) In complex diseases the concordance in phenotype between monozygotic twins is higher than between dizygotic twins
d) Disorders in which genetic factors have a large role always have a significantly higher concordance in phenotype between monozygotic twins and between dizygotic twins than do disorders where genetic factors play a less significant role.

## Question 8.4

With regard to affected sib pair (ASP) analysis which, if any, of the following statements is true.

a) ASP analysis is a type of parametric linkage analysis that is popularly used to study complex diseases.
b) It is highly convenient because samples are required from the affected sibs only and samples from just 100 affected sib pairs are usually enough to obtain decent results.
c) It is less suited to studying complex disease where the relative risk of disease is low.
d) A lod score of 3 is highly significant evidence for linkage in ASP analysis.

## Question 8.5

Regarding two-point linkage analysis, which, if any, of the following statements is true?

a) A lod score of 3 means that the likelihood of the data, given that the two loci are linked, is 1000 times greater than is the likelihood of the data, if the two loci are unlinked.
b) A lod score of 3 is highly significant evidence for linkage.
c) A lod score of -2 is highly significant evidence against linkage.
d) A lod score of 3 is 100,000 times more convincing evidence of linkage than a lod score of -2.

## Question 8.6

With respect to genomewide association (GWA) studies, which, if any, of the following statements, is false?

a) The vast majority of genetic variants identified in GWA studies of complex association studies have been of weak effect, with odds ratios of 1.2 or less.
b) Even where GWA studies have identified disease susceptibility factors with odds ratios greater than 1.2, getting from disease variant to identifying a disease susceptibility locus is essentially impossible.
c) GWA studies have generally been of very limited use in predicting disease risk.
d) GWA studies have been very valuable.

## Question 8.7

With respect to genomewide association (GWA) studies, which, if any, of the following statements, is true?

a) GWA studies to determine genetic susceptibility to human complex disease usually use 1000 polymorphic SNPs markers that are distributed across the genome
b) Unlike linkage analyses, GWA studies are ideally suited to identifying genetic susceptibility in genetically heterogeneous diseases.
c) GWA studies have been very important in elucidating the pathogenesis of complex diseases and in providing new disease biomarkers.
d) GWA studies have transformed our ability to predict disease risk.

## Question 8.8

Genomewide association (GWA) studies have identified common single nucleotide DNA variants that are very rarely of even moderate effect and that are collectively insufficient to explain the genetic variance of complex diseases. Which of the following can at least partly explain the genetic variance?

a) Rare single nucleotide variants.

b) *de novo* variants.
c) Copy number variants.
d) Gene-gene interactions

## Question 8.9

The human *APOE* gene has three common alleles, *APOE*e2, APOE*e3*, and *APOE*e4*, that gives rise, respectively, to the common alleles apoE2, apoE3 and apoE4 at the protein level. Which, if any, of the following statements is true?

a) The **e4* allele confers a high risk of Alzheimer disease and people with two **e4* alleles have twice the disease risk of people with one **e4* allele.
b) The **e3* allele is a protective factor, conferring reduced risk of Alzheimer disease while the **e2* allele has a disease risk that is intermediate between those of **e3* and **e4*.
c) The **e4* allele is considered the ancestral *APOE* allele because although the chimp and gorilla also have three apoE proteins the apoE4 protein is the most frequent allele.
d) The **e4* allele has reached a high frequency because Alzheimer disease has such a late age at onset that reproductive success rates are not diminished.

## Question 8.10

Which, if any of the descriptions below is false. Our personal microbiomes

a) are microorganisms that are mostly located within our guts.
b) contain 10 times more cells than we have.
c) are tolerated, despite being foreign microorganisms, by down-regulating innate immune responses, mostly those that require Toll-like receptors.
d) include microbes that are highly beneficial to us.

# Fill in the Blanks Questions

## Question 8.11

Fill in the blanks

Before we had a human genetic map, disease gene identification was difficult. Sometimes, however, knowledge of the gene product provided a path to a gene underlying a Mendelian disorder. For example, haemophilia A was long known to be a deficiency of a specific blood clotting protein, ___1____ ___2___, and using that information it was possible to purify large amounts of ___1____ ___2___ from pig blood and then to design ____3_____ oligonucleotides that corresponded to all possible codon interpretations of an optimal part of the amino acid sequence of the pig ___1____ ___2___ protein. The resulting oligonucleotides were then used as ____4_____ probes to screen first a human ____5____ library and then a ____6____ ___ 7____ library to identify the underlying human gene.

## Question 8.12

Fill in the blanks

Once we had a human genetic map, a general method, ____1____ ____2____, could be applied to identifying genes underling single gene disorders. In order to carry this through, there was first the need to identify families with multiple affected individuals and to obtain samples of ____3____ ____4____ from both affected and unaffected family members. The individual samples could then be tested by assaying each of several hundred DNA ____5_____ . The DNA ____5____ were selected because they were known to be ____6____ and because they were each known to map to a ____7____ subchromosomal location. The object was to identify ____5____ that must map close to the disease gene because alleles from these markers tended to ____8____ with disease through generations in families.

## Question 8.13

Fill in the blanks

Poor family structure means that it may not be easy, or even feasible, to use ____1____ ___2____ in order to map the underlying gene for certain types of single gene disorders. The disorder may be extremely rare so that insufficient samples can be obtained, or the vast majority of affected individuals are sporadic cases that arise through a ____3____ ___4____ mutation, such as in severe congenital ____5____ disorders. In the past, gene identification in these cases

was sometimes possible using cytogenetic analyses to identify a disease-associated chromosome abnormality. Two or more affected individuals might be identified to have a translocation or ____6____ with a common ____7_____ or a recurring interstitial or terminal ____8_____ might be identified. Once the subchromosomal location of the gene had been identified it was possible to use ____9_____ ____10_____ or ____9_____ ___11____ gene approaches to identify the underlying gene. More recently, a more direct approach, ___12____ ____13____ makes it possible to quickly get to the underlying gene without knowing its chromosomal location.

## Question 8.14

Fill in the blanks

In a _____1____-control study affected individuals ( ___1___ ) are compared with control subjects that are suitably ___2___ to the ____1___ with regard to age, sex and so on. In genetic studies of complex diseases the ___1___ and controls are genotyped for genetic ___3____ in the hope of identifying ___3____ that may confer significantly increased disease risk ( ____4____ ___5____ ) or significantly reduced disease risk ( ____6____ ____5____ ). A popular way of assessing the data is to calculate an ___7___ ___8_____, that is the ___7____ of being affected if the genetic ____3____ is present divided by the odds of being affected if the genetic ____3____ is absent.

## Question 8.15

Fill in the blanks

While linkage is a ____1_____ phenomenon, association is a _____2_____ property. Linkage is concerned with seeking to identify the map relationship between two or more ___3___, such as disease and marker ___3____, by analysing samples from individuals within ____4____. Association, on the other hand, is a relationship between ____5____and is studied by analysing samples from individuals within ____6____. Linkage works over ____7____ distances whereas, in practice, association works over very ____8_____ distances. If allele *A*1* is shown to be positively associated with a complex disease it would be described as a disease ____9____ factor, and if allele *A*2* is negatively associated with the disease it would be described as a ____10____ factor

## Question 8.16

Fill in the blanks

In an autoimmune disorder, cells in the body come under attack from certain antibodies known as ____1_____ , and also from certain types of ____2_____ T cell that may attack specific host cells (such as insulin-producing ____3_____ ____4_____ cells in the case of type I diabetes). In ____2_____ T cell responses the host cell peptides serve as ____5_____ and they are presented to T cells after they have been bound by ____6_____ proteins. ____6_____ proteins differ in their ability to bind individual peptide ____5____ and that is the primary basis for ____6____-disease associations.

## Question 8.17

Fill in the blanks

Imagine a genetic variant that is tightly linked to a disease susceptibility allele. On a single chromosome, the ____1____ containing the genetic variant and the linked disease susceptibility allele will have a higher ____2____ than would be expected (that is, it would be higher than the ____2____ of the genetic variant multiplied by the ____2_____ of the disease susceptibility allele. This is an example of _____3_____ ____4_____ , the non-random _____5_____ of alleles at two or more loci. Although _____3_____ ____4_____ describes *any* non-random _____5_____ of alleles at different loci, in practice, the alleles are at very closely _____6_____ loci. Although _____3_____ _____4_____ can occur if a particular combination of alleles offer some advantage and is _____7_____ selected, it may often simply reflect reduced _____8_____ between loci (certain regions of the genome, such at the HLA complex, show significantly reduced _____8_____ ). When a new DNA variant is created by mutation it will show very tight ____3_____ with alleles at neighboring loci. However, the _____3_____ ____4_____ will gradually be eroded by _____8_____ but that will take a very _____9_____ time for any locus that is physically very close to the locus with the new mutation.

## Question 8.18

Fill in the blanks with single words or single letters.

Each of us carries our personal ____1____ that shares our body space (but is principally distributed within our ____2____) and that contains ____3____ times more cells than our body. These cells are foreign ____4_____ that are nevertheless tolerated by the body, largely by suppressing _____5_____ ____6_____ responses, notably those that depend on _____7____ receptors. Our personal _____1_____ is normally beneficial to us because some of the _____4_____ are beneficial to us in different ways. They can help us derived additional energy through the fermentation of undigested _____8_____, help us break down _____9_____ , and they synthesize vitamins _____10_____ and _____11_____ for us. In ____12____ ____13____ diseases, however, an abnormal ____6____ response is directed against antigens

carried by foreign ____4_____ within our ___2____ and that leads to accumulation of ____14____ blood cells within the linings of the _____15_____ , producing chronic _____16_____.

Essay and List Questions

## Question 8.19

The table below shows the percentage phenotype concordance in monozygotic (MZ) and dizygotic (DZ) twins in four hypothetical genetic diseases A to D. Which disease would you estimate to have the highest heritability and which one has the lowest heritability, and why?

| **Disease** | **% Concordance in MZ twins** | **% Concordance in DZ twins** |
|---|---:|---:|
| A | 25.5 | 5.6 |
| B | 15.2 | 11.3 |
| C | 7.0 | 3.0 |
| D | 41.0 | 5.2 |

## Question 8.20

Strategies to identify the genes that underlie single gene disorders have often relied on first obtaining a subchromosomal location for the disease gene. List two approaches that have been taken to identify subchromosomal locations for these disorders.

## Question 8.21

The first genome-wide human genetic map was created by taking a completely different approach to the approaches used to create genetic maps in model organisms. What was the essential difference?

## Question 8.22

To carry out exome sequencing the desired exome is first captured from a sample of genomic DNA. How is that achieved?

## Question 8.23

The risk ratio, $\lambda$, is commonly used in complex disease. What does it mean and how do risk ratios compare in complex disease and monogenic disorders?

## Question 8.24

Illustrate how the heritability of a disease can change over time using an example of a) a monogenic disorder and b) a complex disease.

## Question 8.25

What is the difference between parametric and non-parametric linkage analyses and under what circumstances are they applied to studying human genetic disease.

## Question 8.26

What is a haplotype block, and how are they organized in the human genome?

## Question 8.27

What is the basis of HLA associations with autoimmune disorders?

## Question 8.28

Before genomewide association (GWA) studies became successful, association studies used to rely on candidate gene approaches. How successful were the candidate gene approaches?

## Question 8.29

In the years when association studies were limited to candidate gene approaches what were the technological drawbacks that prevented genomewide association (GWA) studies and what developments made GWA studies possible?

## Question 8.30

How successful have genomewide association studies been in identifying genetic susceptibility to complex disease? What has been the main value of these studies?

## Question 8.31

List three possible explanations for the general failure of GWA studies to identify genetic factors that collectively might explain the heritability of complex diseases.

## Question 8.32
Certain common alleles are known to be associated with specific complex diseases. Why has purifying selection not led to these alleles being eliminated from the population?

## Question 8.33
Balancing selection might explain why certain common alleles known to be associated with specific complex diseases have not been eliminated by natural selection. List some examples that might suggest this to be the case.

## Question 8.34
Certain common alleles known to be associated with specific complex diseases may have been maintained in the human population because of beneficial advantage in the past. List some examples that might suggest this to be the case.

## Question 8.35
List three observations that support the common disease-common variant hypothesis.

## Question 8.36
List three justifications for the common disease-rare variant hypothesis

## Question 8.37
To what extent do late-onset common Alzheimer disease and rare autosomal dominant Alzheimer disease show common pathologies at the physiological level and what is the evidence for common biological pathways in the Mendelian and complex disease versions of this disease?

## Question 8.38
HLA associations with autoimmune disorders are among the very strongest of associations between genetic variants and disease, but have rarely been mapped to the amino acid level. An exception is the HLA association with rheumatoid arthritis. What is known about this association, and how did detailed knowledge of how HLA proteins function help to identify the amino acids implicated in the association?

## Question 8.39

Naturally occurring genetic variants are known to act as protective factors to confer reduced risk of infectious disease. List some examples.

## Question 8.40

Genomewide association studies have been very successful in identifying susceptibility factors in inflammatory bowel diseases, and have provided valuable insights into the pathogenesis of these diseases. What types of insights have they offered?

## Question 8.41

Outside of cancers and infectious diseases, various other complex diseases are known to be strongly influenced by environmental factors. Give three examples of environmental factors that are known to increase the risk of specific complex diseases other than cancers and infectious diseases.