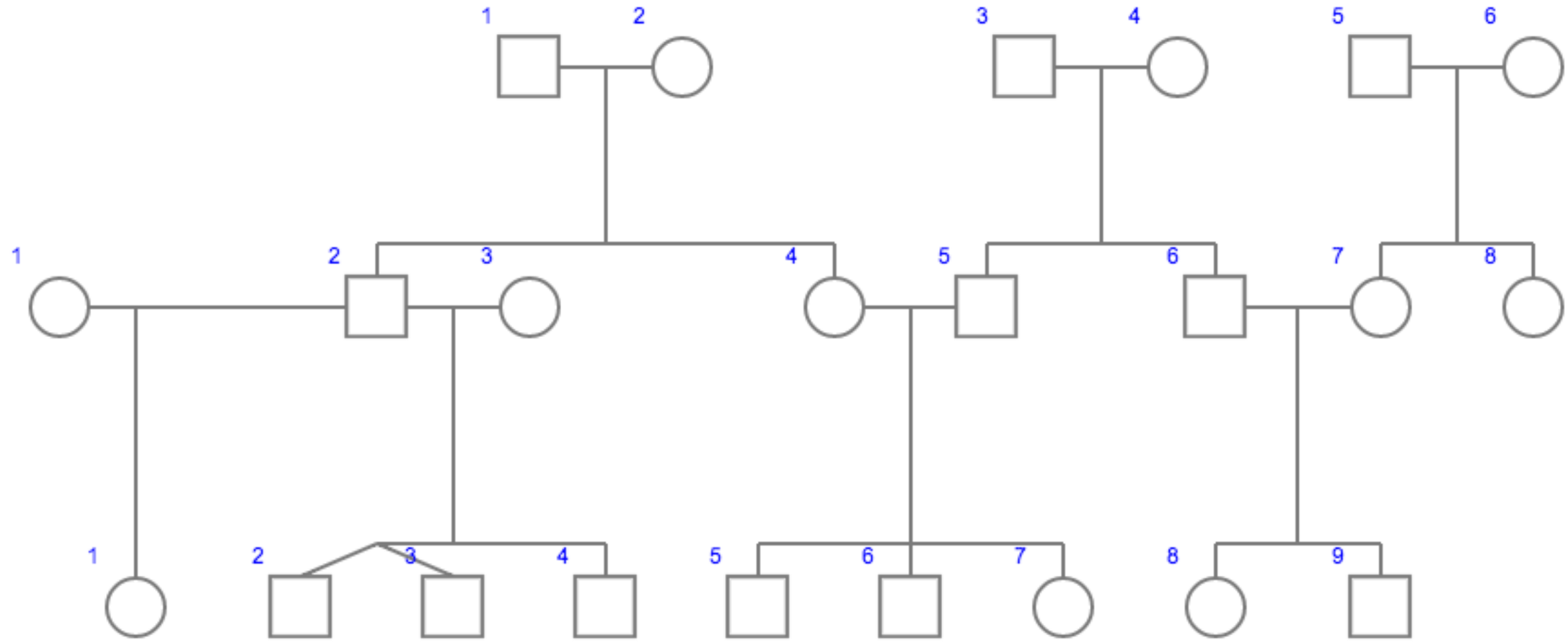
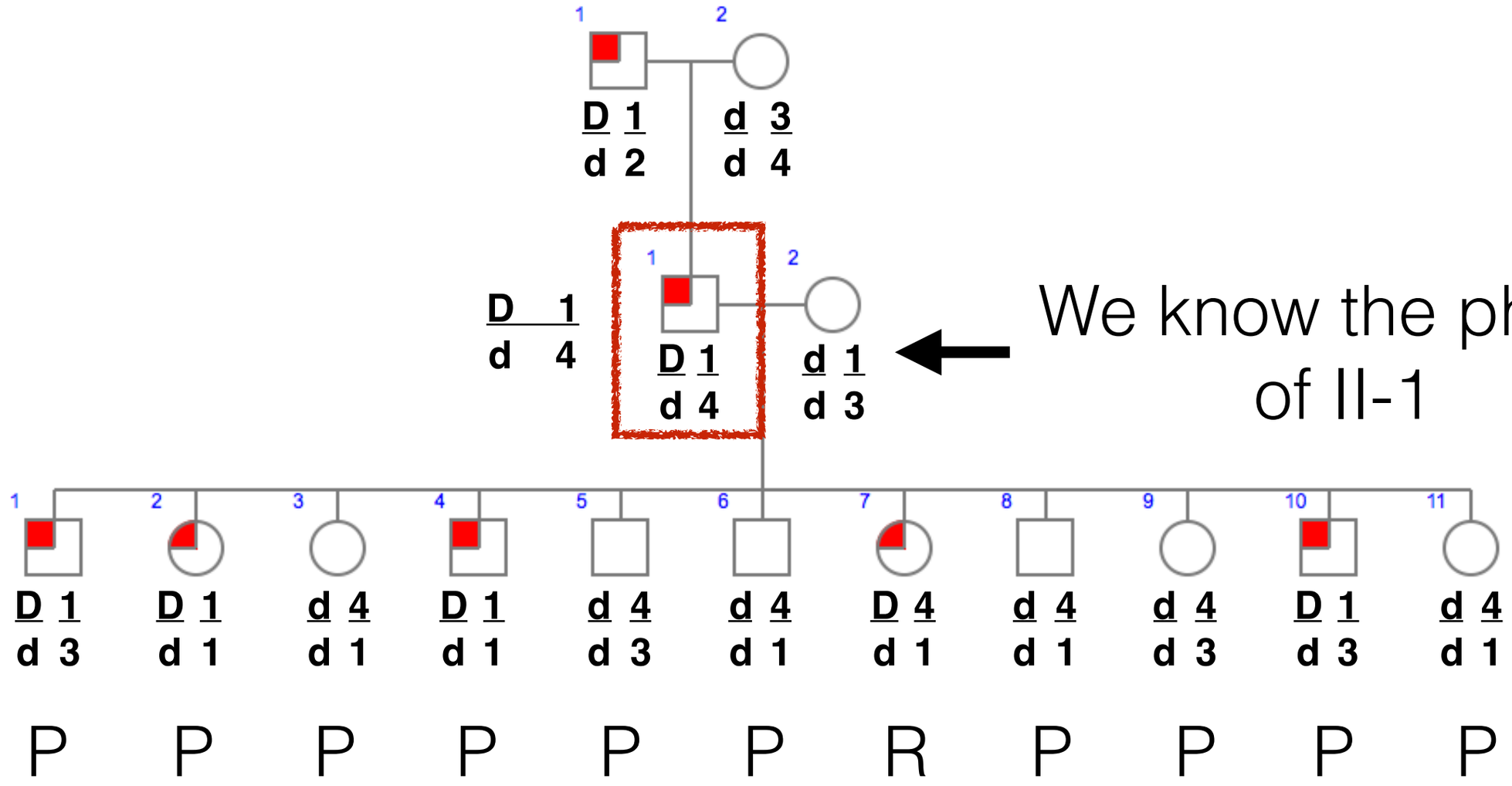
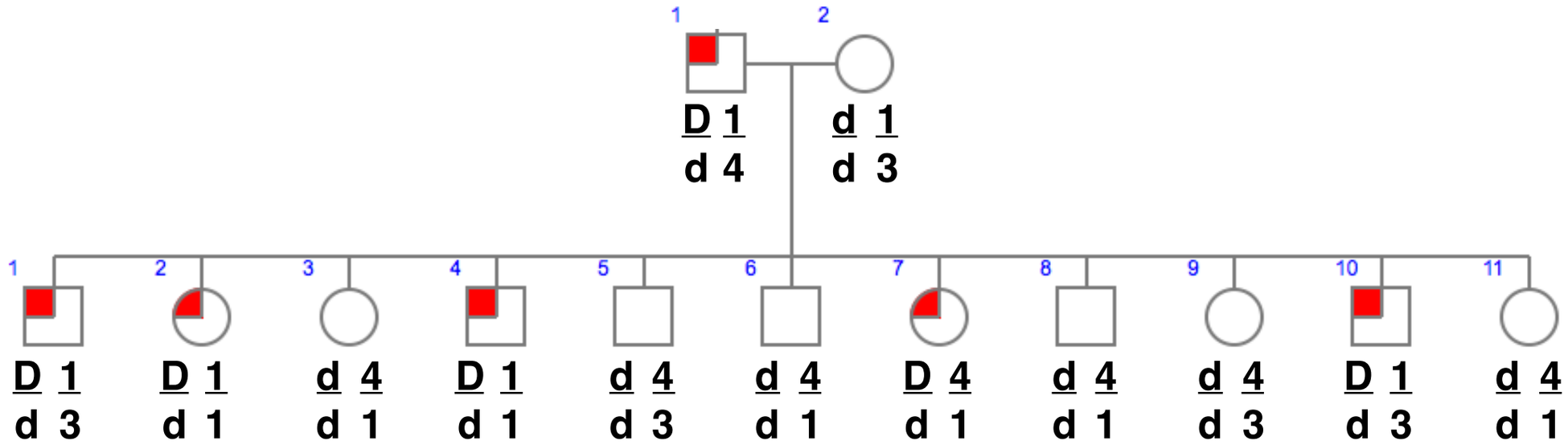
Linkage mapping in families



Linkage to genetic markers tells us where disease genes are



Linkage to genetic markers tells us where disease genes are



P P P P P P R P P P P

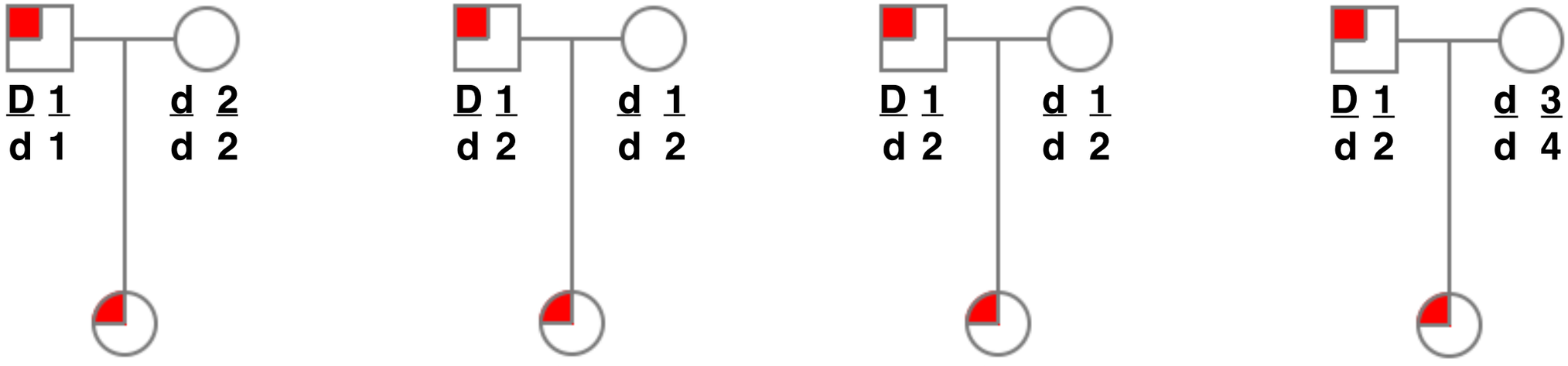
OR

R R R R R R P R R R R

Sometimes, we don't know the phase of the parent, and both possibilities of phase are equally likely

Some allelic combinations are non-informative and can not be included in mappings

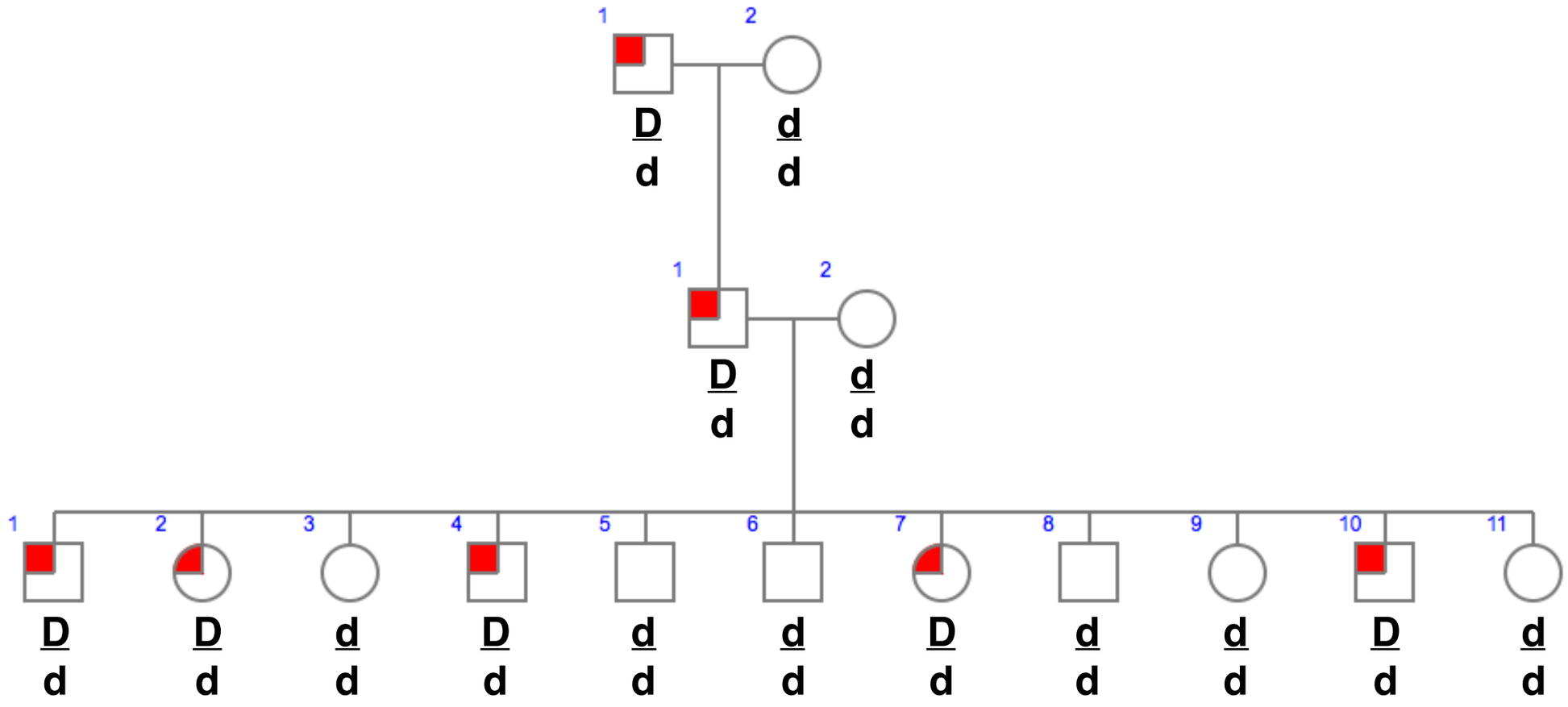
Consider a dominant trait and a variant marker:



D 1 / d 2	D 1 / d 2	D 1 / d 1	D 1 / d 3
Uninformative	Uninformative	Informative	Informative
Which marker allele is linked to disease allele?	Which allele 1 is inherited from Dad?	Each allele 1 must have come from Mom and Dad	Different marker alleles

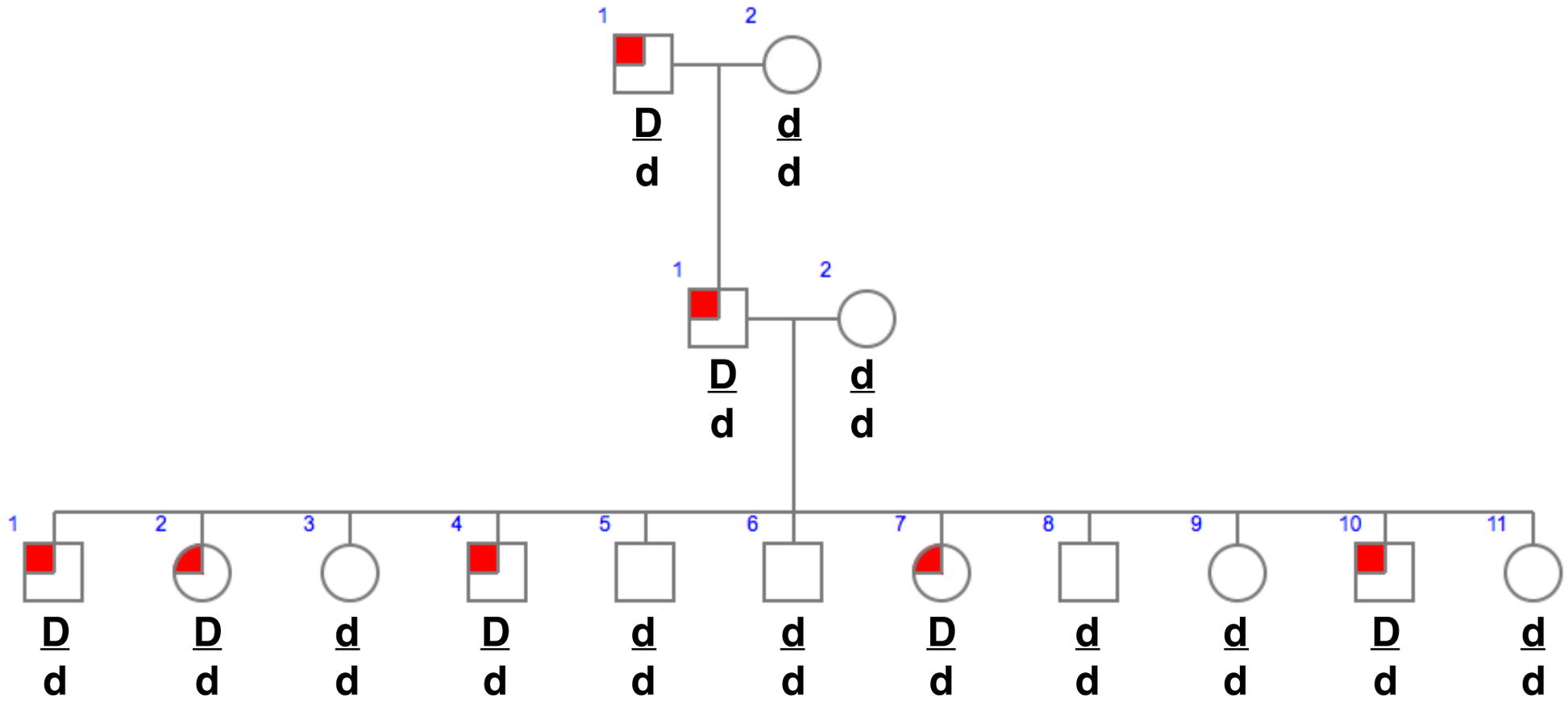
We want to determine if the daughter inherited a recombinant or parental chromosome

Imagine you could genotype millions of markers in each individual



The goal is to measure linkage of many markers to the disease-causing allele to map the gene

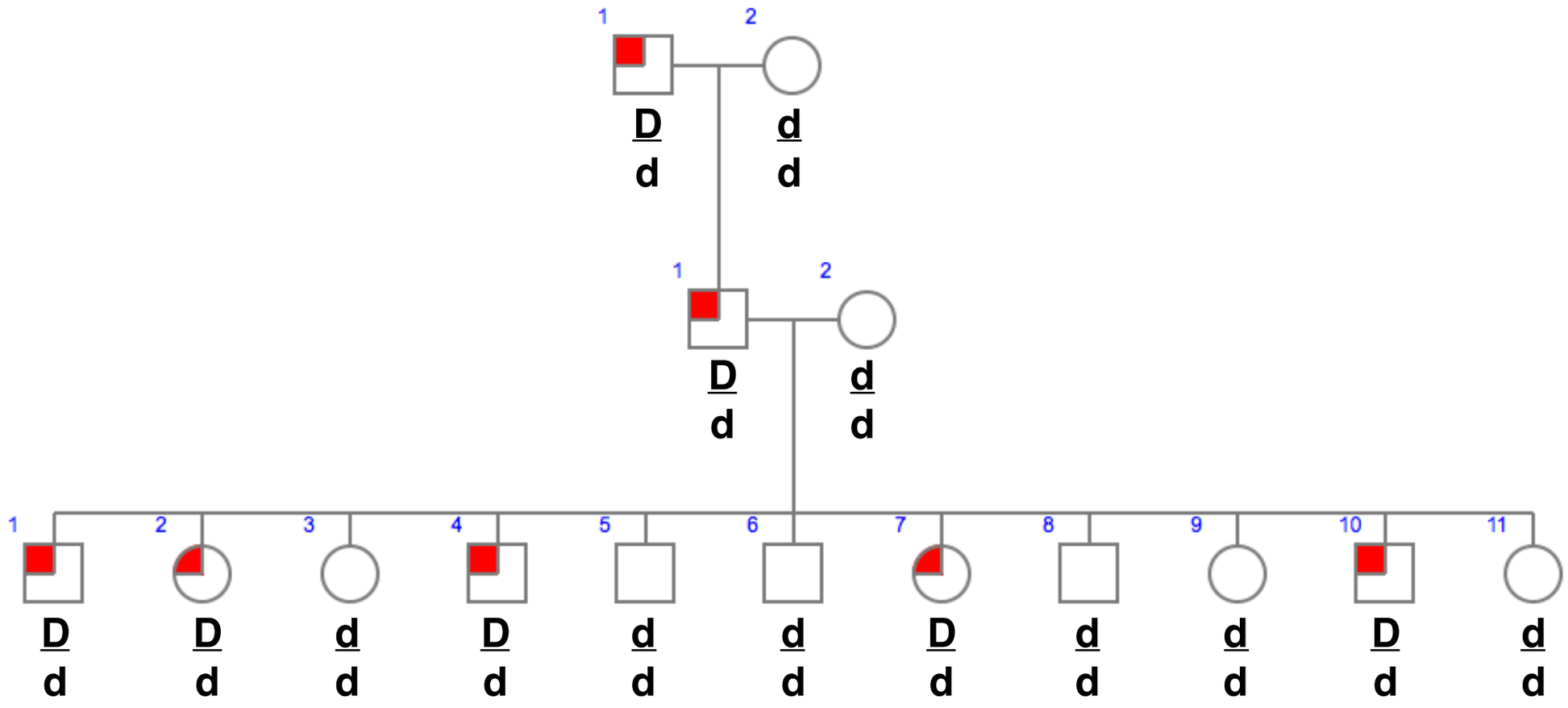
We want to measure how close each marker is to the disease-causing allele



$$\frac{\textbf{Number of recombinants}}{\textbf{Total progeny}} \times \textbf{100} = \textbf{Recombination frequency}$$

But we don't know who is a recombinant because we don't have true-breeding strains

We want to measure how close each marker is to the disease-causing allele



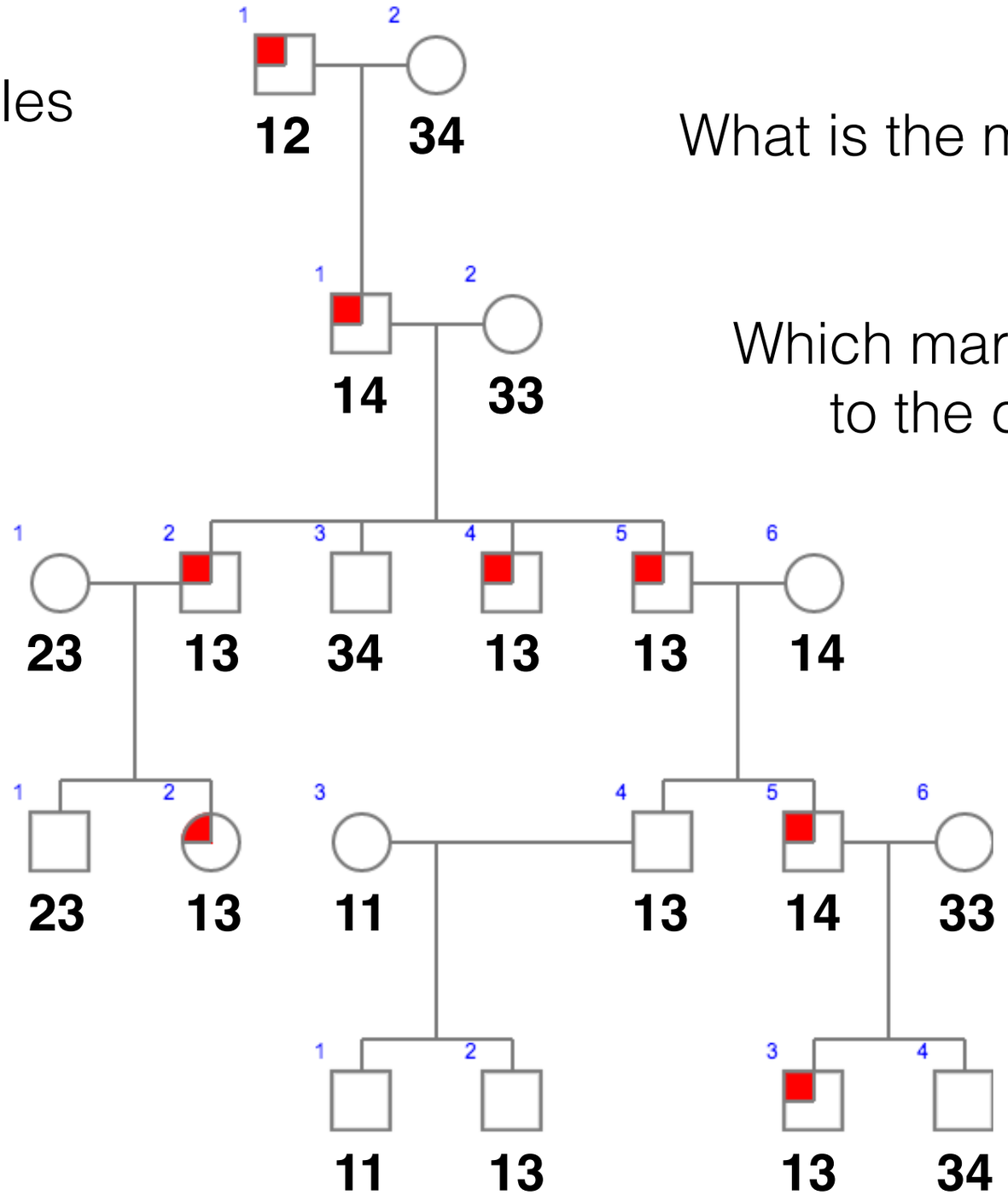
Likelihood of linkage between marker and disease-causing allele

---

Likelihood of NO linkage between marker and disease-causing allele

The goal of linkage analysis is to identify a marker nearby the disease-causing allele

Marker with alleles
1, 2, 3, or 4



What is the mode of inheritance?

Which marker allele is linked
to the disease allele?

Linkage is reported as a log of the odds ratio or LOD

The odds ratio is a statistical method to measure association of some attribute with the presence or absence of another attribute.

Probability of pedigree under linkage versus no linkage

$$\text{LOD} = \log_{10} \frac{\text{P(pedigree with linkage)}}{\text{P(pedigree with no linkage)}}$$

LOD > 3 is good evidence of linkage
1 in 1,000 chance of data by random assortment

LOD < 0 means no linkage
and observed data are more likely by chance

How linked is our marker to the disease-causing allele? Recombination frequency

Recombination frequency is written as θ

Percentage of recombinant gametes passed down

θ = 0 is perfect linkage

θ = 0.5 is no linkage

Linkage is reported as a log of the odds ratio or LOD

The odds ratio is a statistical method to measure association of some attribute with the presence or absence of another attribute.

Probability of pedigree under linkage versus no linkage

$$\text{LOD} = \log_{10} \frac{\text{P(pedigree with linkage)}}{\text{P(pedigree with no linkage)}}$$

$$\text{LOD} = \log_{10} \frac{\text{P(data} \mid \theta)}{\text{P(data} \mid \theta = 0.5)}$$

P(data) = probability that a particular gamete was inherited

P(data $\mid \theta = 0.5$) is independent assortment

Linkage is reported as a log of the odds ratio or LOD

$$LOD = \log_{10} \frac{P(data \mid \theta)}{P(data \mid \theta = 0.5)}$$

P(data) = probability that a particular gamete was inherited

$P(data \mid \theta)$ =

Probability that recombination did not occur for parental $1 - \theta$

Probability that recombination did occur for recombinant θ

Probability of phase of parent

Each individual is independent.
Use product rule to multiply probabilities.

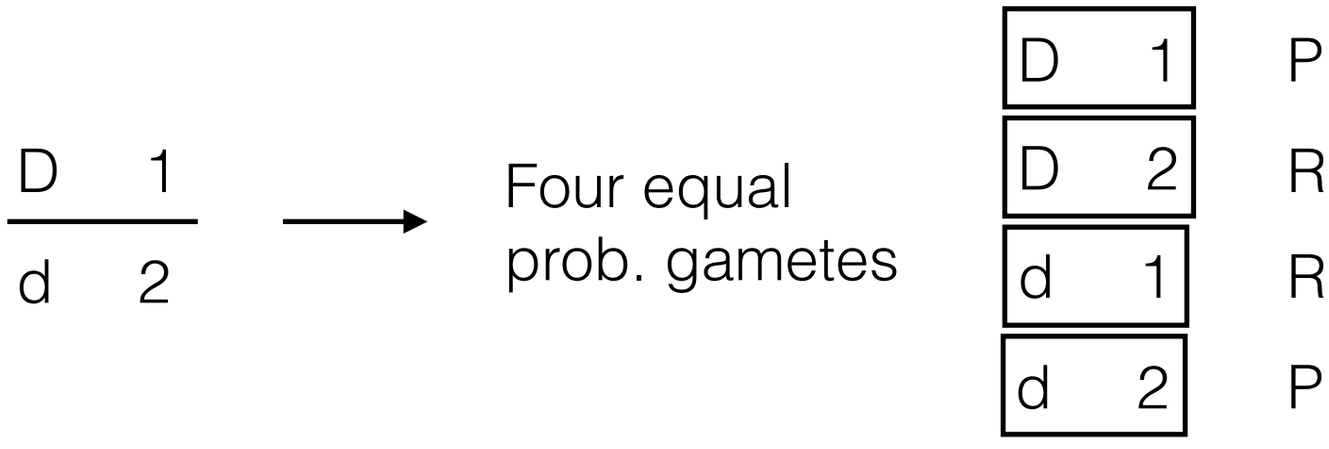
Linkage is reported as a log of the odds ratio or LOD

$$LOD = \log_{10} \frac{P(data \mid \theta)}{P(data \mid \theta = 0.5)}$$

P(data) = probability that a particular gamete was inherited

$$P(data \mid \theta = 0.5)$$

Equal probability of two loci independently assorting and one gamete being passed down



Each individual is independent.
Use product rule to multiply probabilities.

Linkage is reported as a log of the odds ratio or LOD

Probability of pedigree under linkage versus no linkage

$$LOD = \log_{10} \frac{P(\text{pedigree with linkage})}{P(\text{pedigree with no linkage})}$$

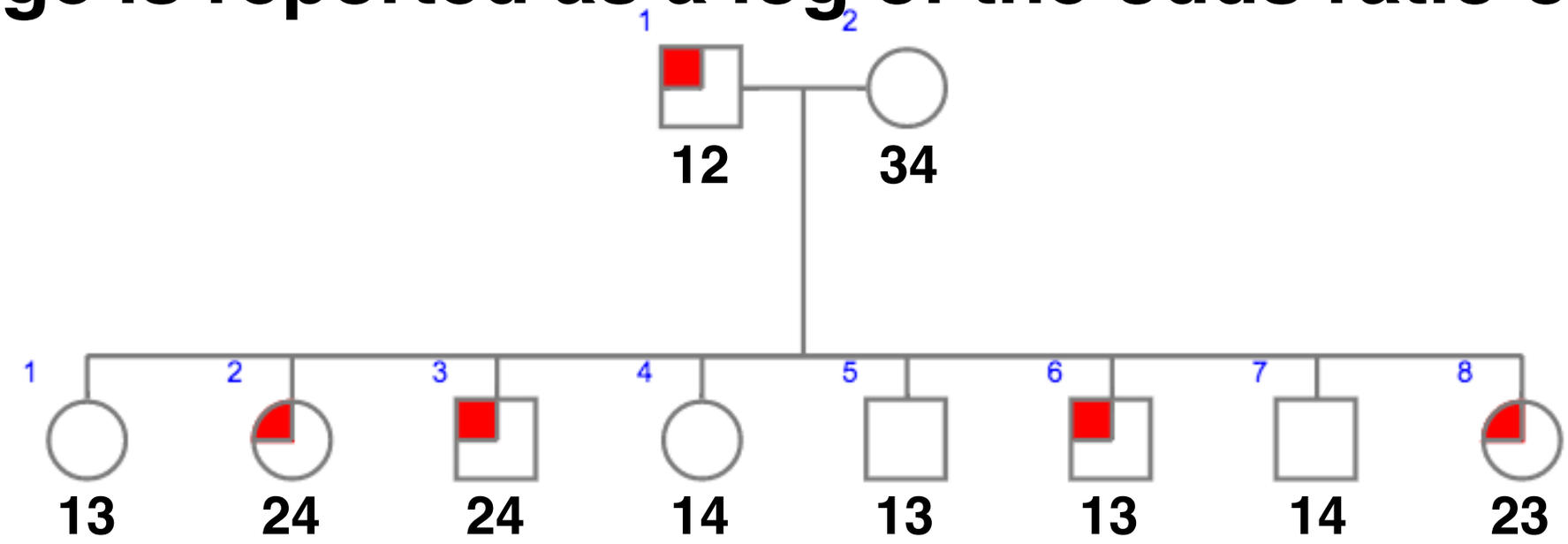
$$LOD = \log_{10} \frac{P(data \mid \theta)}{P(data \mid \theta = 0.5)}$$

P(data) = probability that a particular gamete was inherited

$$LOD = \log_{10} \frac{(1 - \theta)^P \times \theta^R}{0.5^{(P + R)}}$$

Use this equation when we know phase.

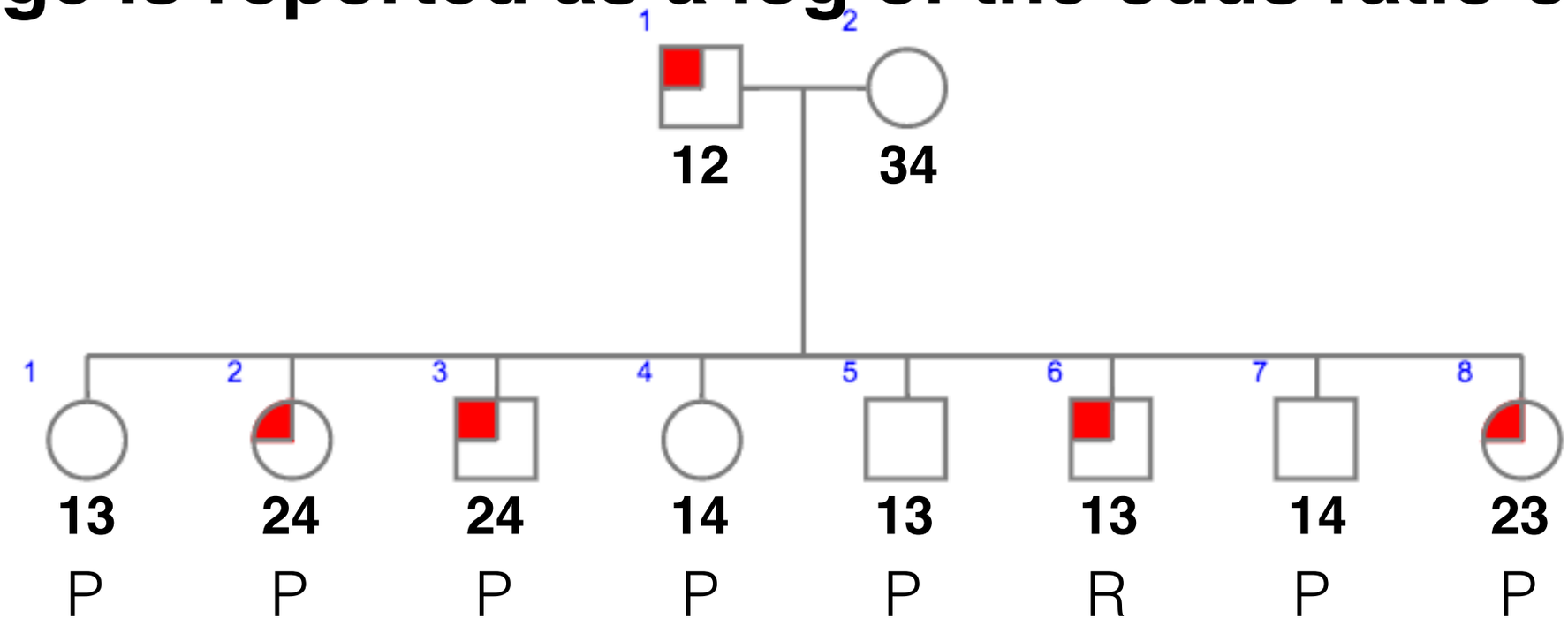
Linkage is reported as a log of the odds ratio or LOD



The paternal grandmother had the disease and her genotype was 2,4.
Her husband's genotype was 11 and he was not affected

$$LOD = \log_{10} \frac{(1 - \theta)^P \times \theta^R}{0.5^{(P + R)}}$$

Linkage is reported as a log of the odds ratio or LOD

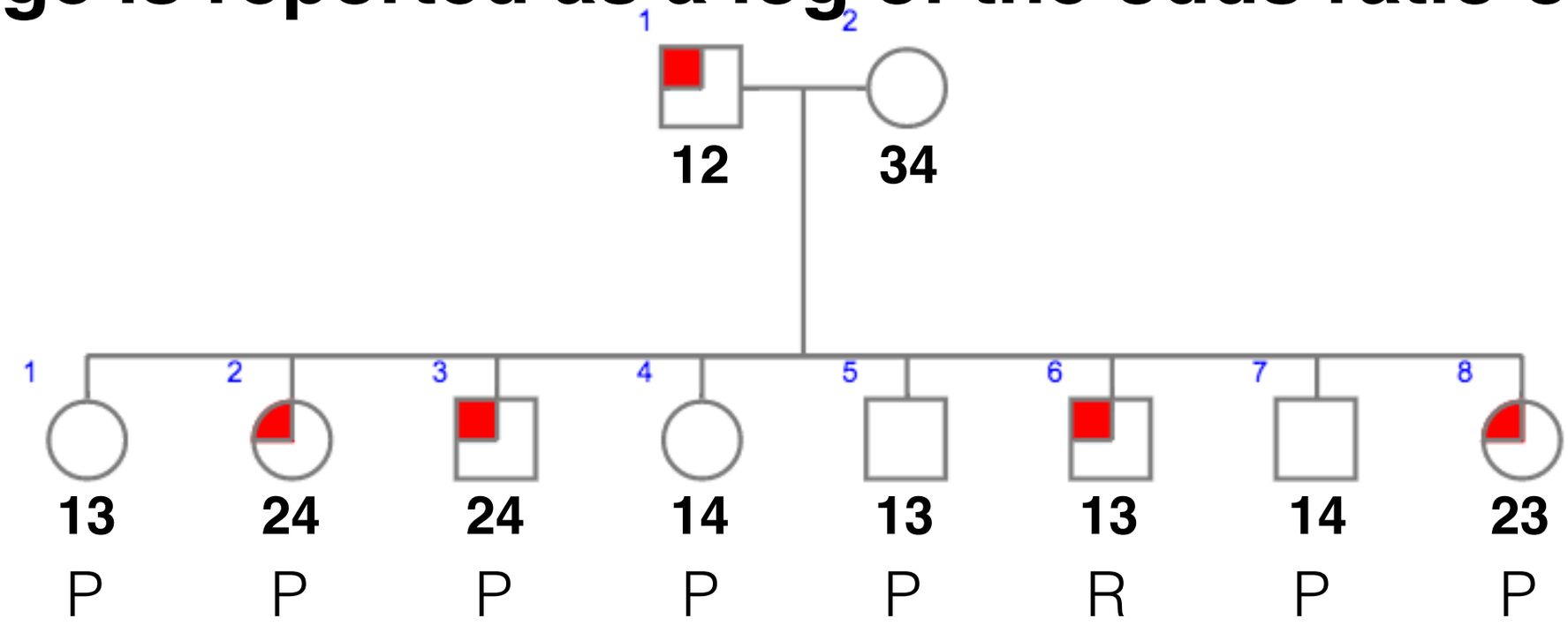


The paternal grandmother had the disease and her genotype was 24.
Her husband's genotype was 11 and he was not affected

$$LOD = \log_{10} \frac{(1 - \theta)^P \times \theta^R}{0.5^{(P + R)}}$$

1. Determine who is recombinant or parental

Linkage is reported as a log of the odds ratio or LOD

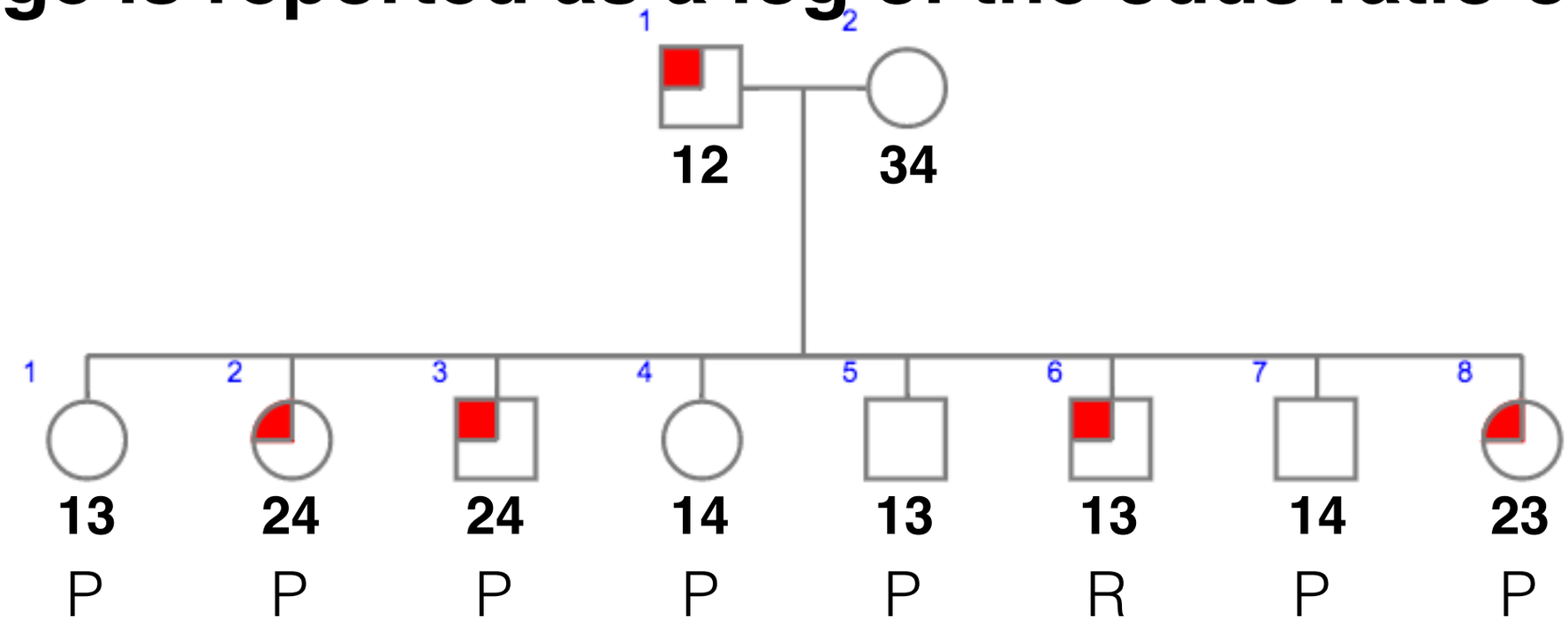


The paternal grandmother had the disease and her genotype was 24.
Her husband's genotype was 11 and he was not affected

$$LOD = \log_{10} \frac{(1 - \theta)^7 \times \theta^1}{0.5^{(7+1)}}$$

2. Put the P and R numbers in the equation

Linkage is reported as a log of the odds ratio or LOD



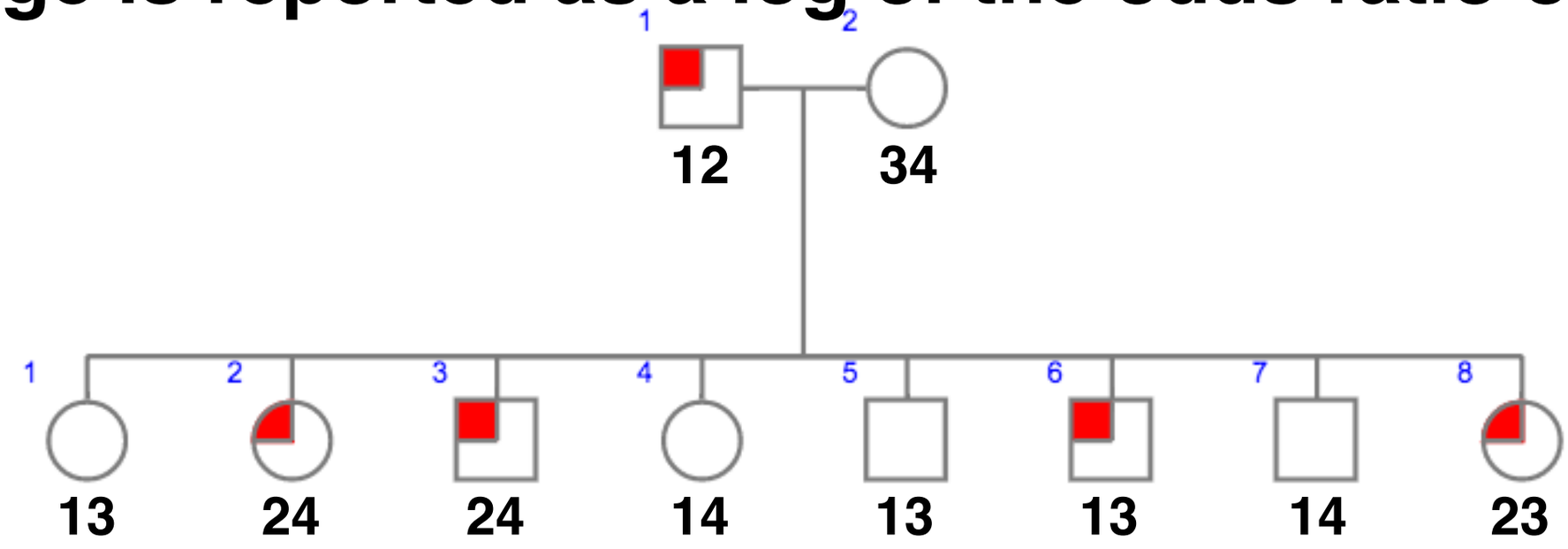
The paternal grandmother had the disease and her genotype was 24.
Her husband's genotype was 11 and he was not affected

$$LOD = \log_{10} \frac{(1 - 0.125)^7 \times 0.125^1}{0.5^{(7+1)}}$$

3. Put the recombination fraction in the equation

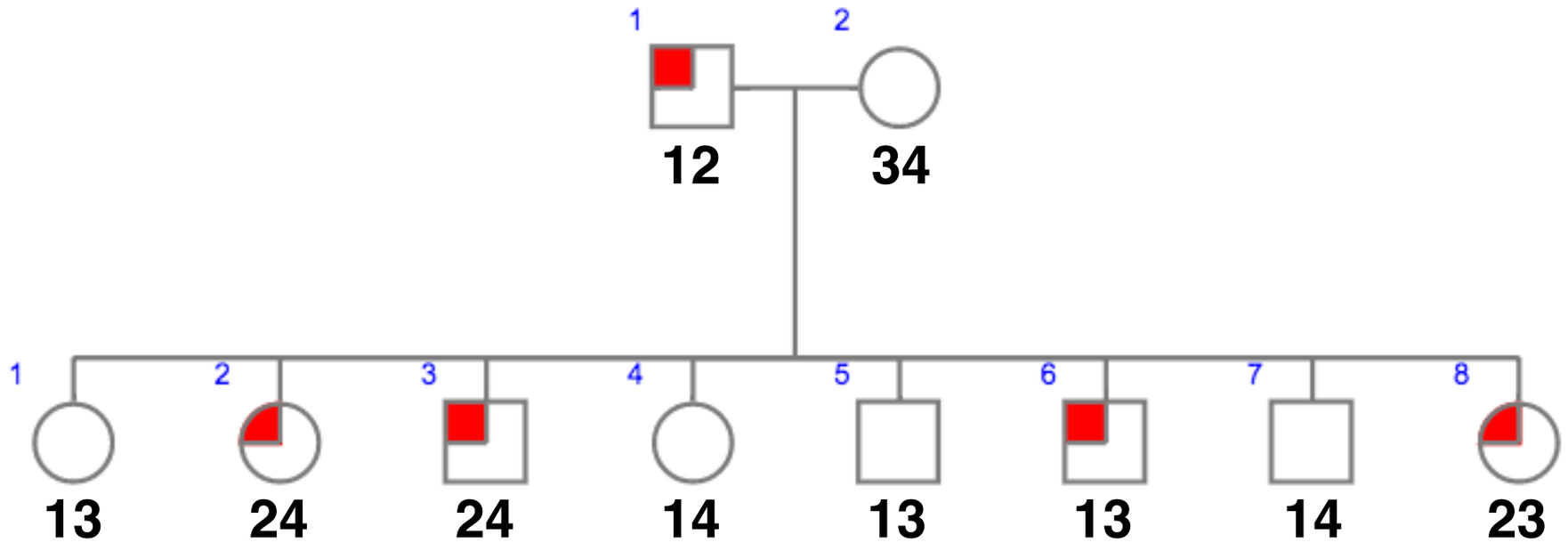
$\theta = 0.125$

Linkage is reported as a log of the odds ratio or LOD



LOD = 1.1 $\theta = 0.125$

What happens when don’t know the phase?



Do we know the phase of I-1?

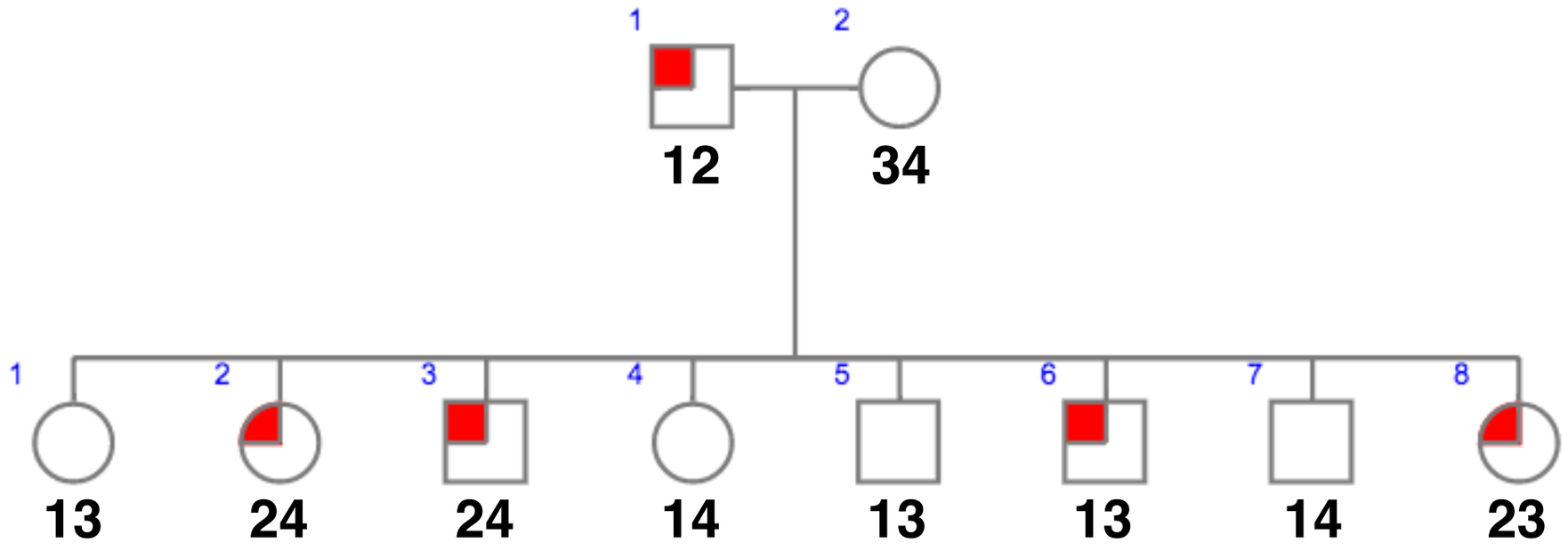
NO

Every child has an equal chance of being a recombinant or parental

Difference between being informative and knowing phase

Informative = Parent heterozygous at each of two loci

What happens when don’t know the phase?



With unknown phase of parent, the LOD equation is…

$$LOD = \log_{10} \frac{\frac{1}{2}((1 - \theta)^P \times \theta^R) + \frac{1}{2}((1 - \theta)^R \times \theta^P)}{(0.5^{(P + R)})}$$

½ chance for each phase

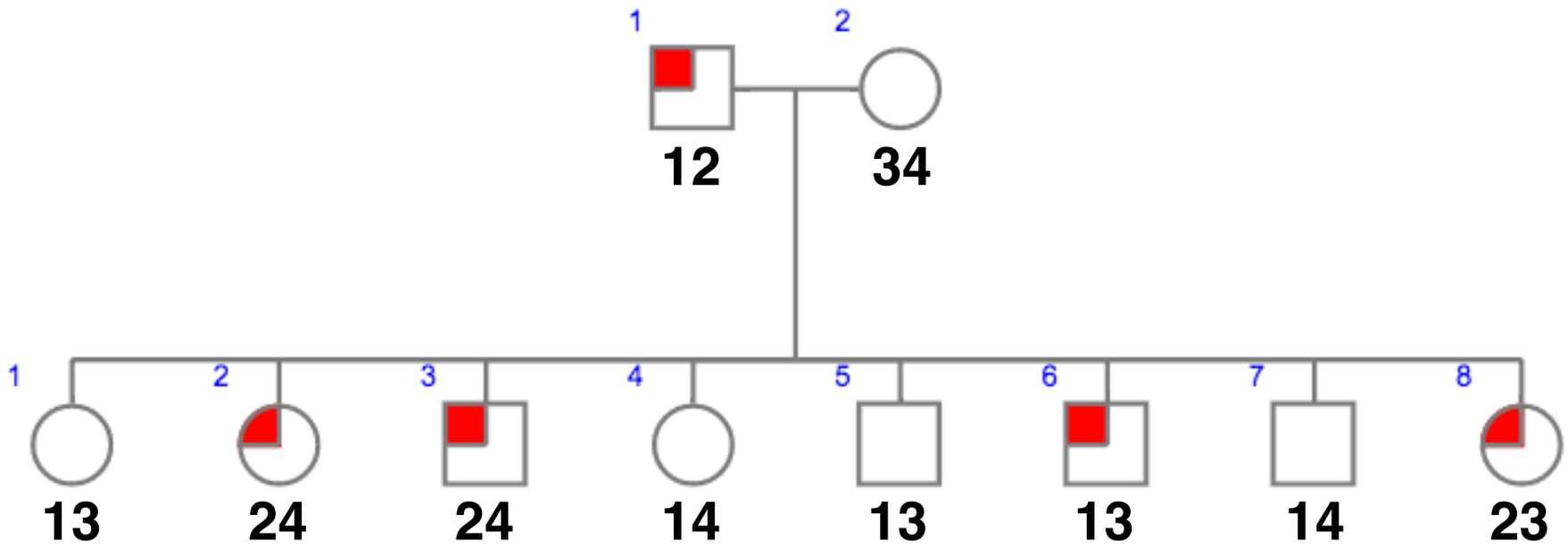
Phase 1

$$\frac{D \quad 2}{d \quad 1}$$

Phase 2

$$\frac{D \quad 1}{d \quad 2}$$

What happens when don't know the phase?



Phase 1	P	P	P	P	P	R	P	P

$$LOD = \log_{10} \frac{\frac{1}{2}((1-\theta)^{P} \times \theta^{R}) + \frac{1}{2}((1-\theta)^{R} \times \theta^{P})}{(0.5^{(P+R)})}$$

½ chance for each phase

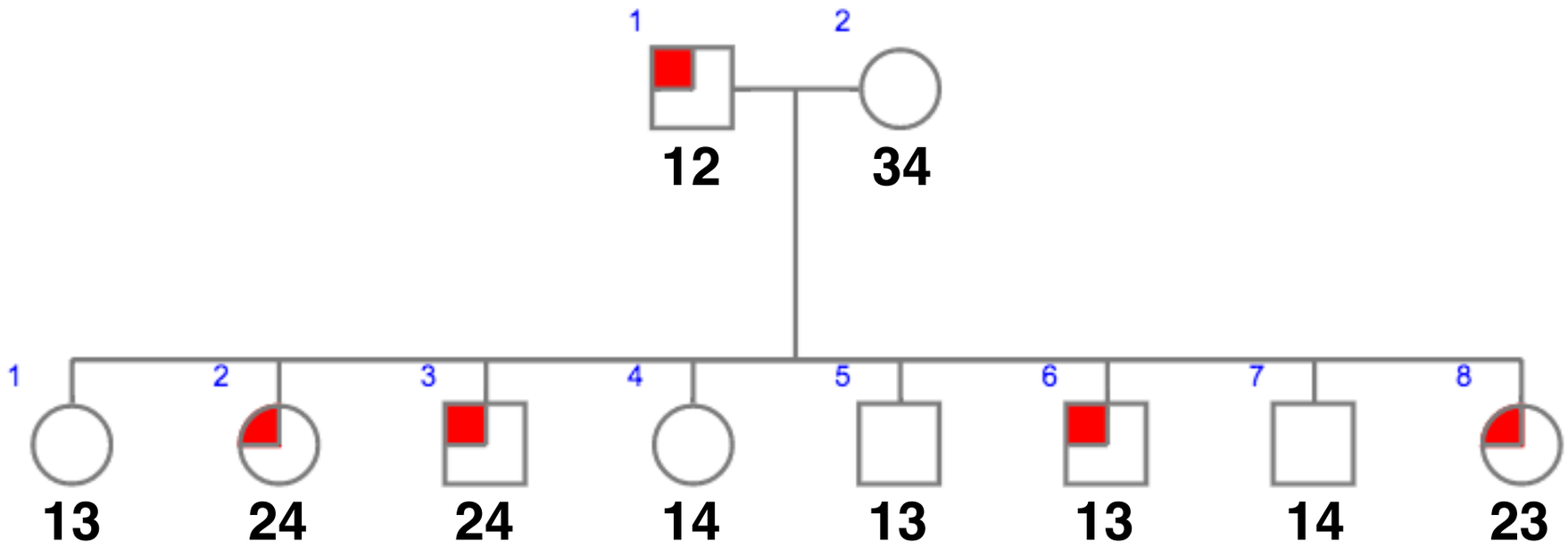
Phase 1

$$\frac{D \quad 2}{d \quad 1}$$

Phase 2

$$\frac{D \quad 1}{d \quad 2}$$

What happens when don't know the phase?



Phase 1	P	P	P	P	P	R	P	P

$$LOD = \log_{10} \frac{\frac{1}{2}((1-\theta)^{7} \times \theta^{1}) + \frac{1}{2}((1-\theta)^{R} \times \theta^{P})}{(0.5^{(8)})}$$

½ chance for each phase

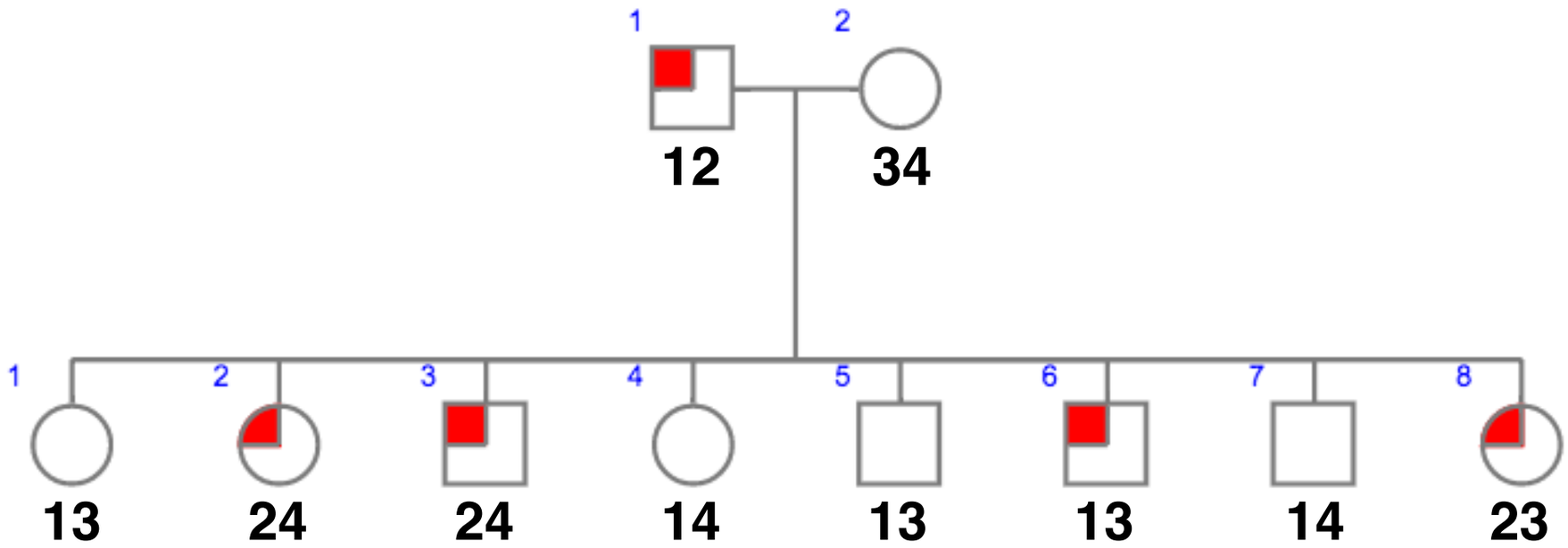
Phase 1

$$\frac{D \quad 2}{d \quad 1}$$

Phase 2

$$\frac{D \quad 1}{d \quad 2}$$

What happens when don't know the phase?



Phase 2	R	R	R	R	R	P	R	R

$$LOD = \log_{10} \frac{\frac{1}{2}((1-\theta)^7 \times \theta^1) + \frac{1}{2}((1-\theta)^1 \times \theta^7)}{(0.5^{(8)})}$$

½ chance for each phase

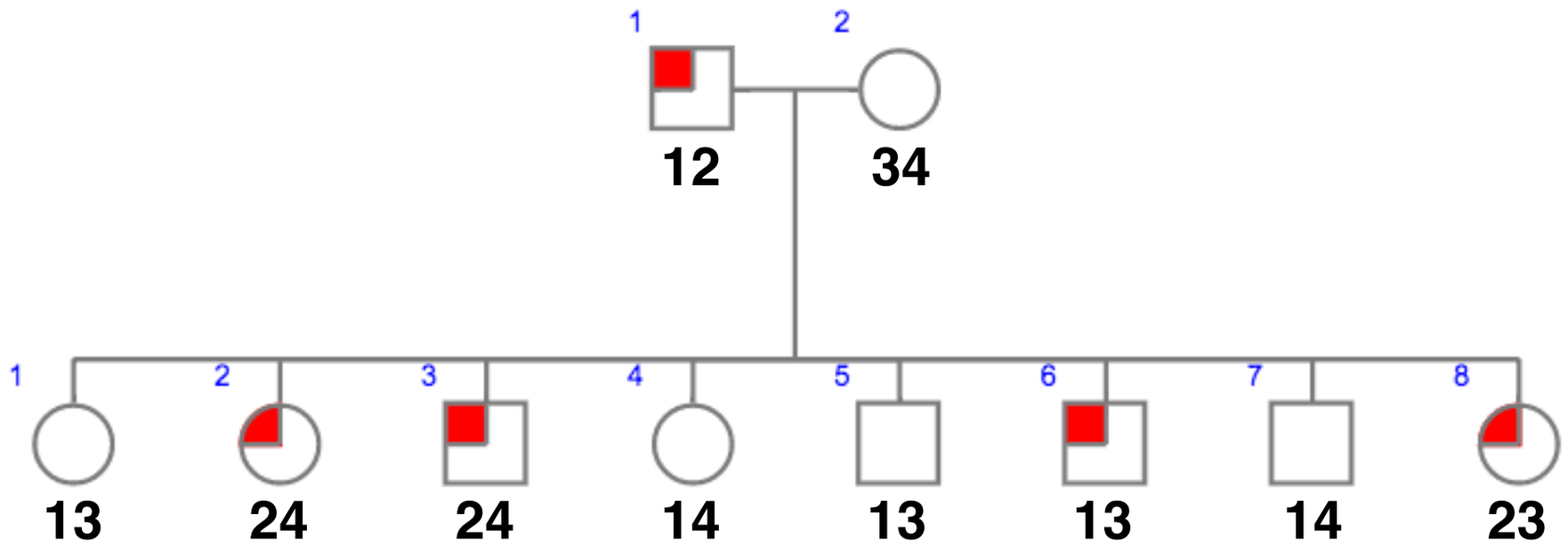
Phase 1

$$\frac{D \quad 2}{d \quad 1}$$

Phase 2

$$\frac{D \quad 1}{d \quad 2}$$

What happens when don't know the phase?



Phase 2	R	R	R	R	R	P	R	R

$$LOD = \log_{10} \frac{½((1 - 0.125)^7 \times 0.125^1) + ½((1 - 0.125)^1 \times 0.125^7)}{(0.5^{(8)})}$$

½ chance for each phase

$\theta = 0.125$

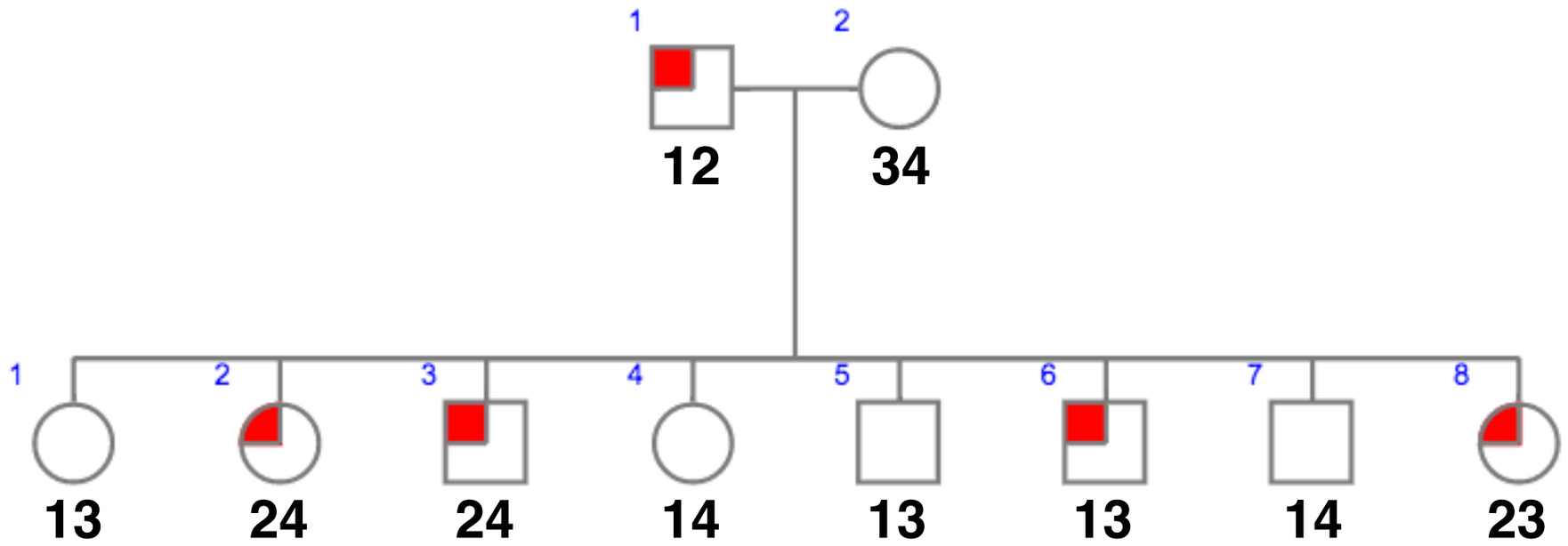
Phase 1

$$\frac{D \quad 2}{d \quad 1}$$

Phase 2

$$\frac{D \quad 1}{d \quad 2}$$

What happens when don’t know the phase?



Phase 2	R	R	R	R	R	P	R	R

LOD = 0.79 $\theta = 0.125$

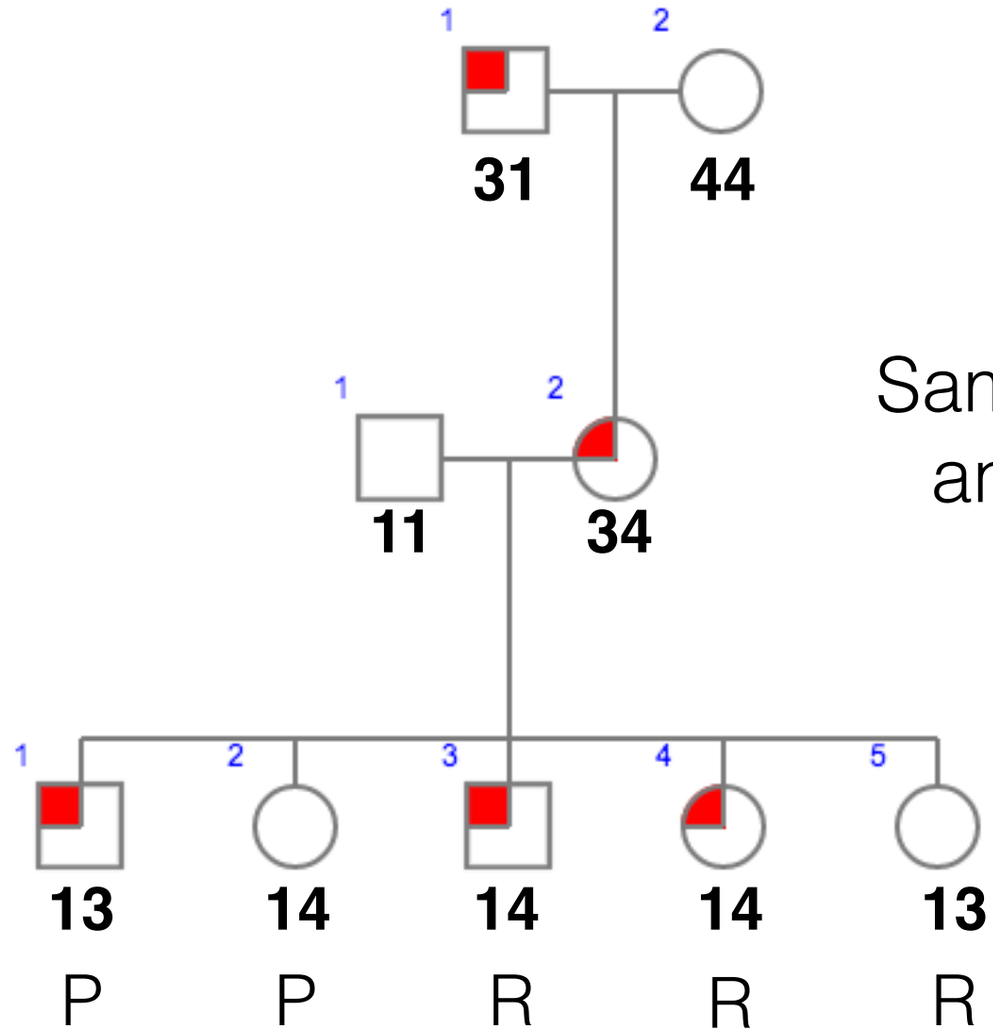
Without phase information, LOD scores decrease

Some properties of LOD scores

LOD scores from independent families can be added
(product rule with logarithms)

Determining phase increases the LOD score

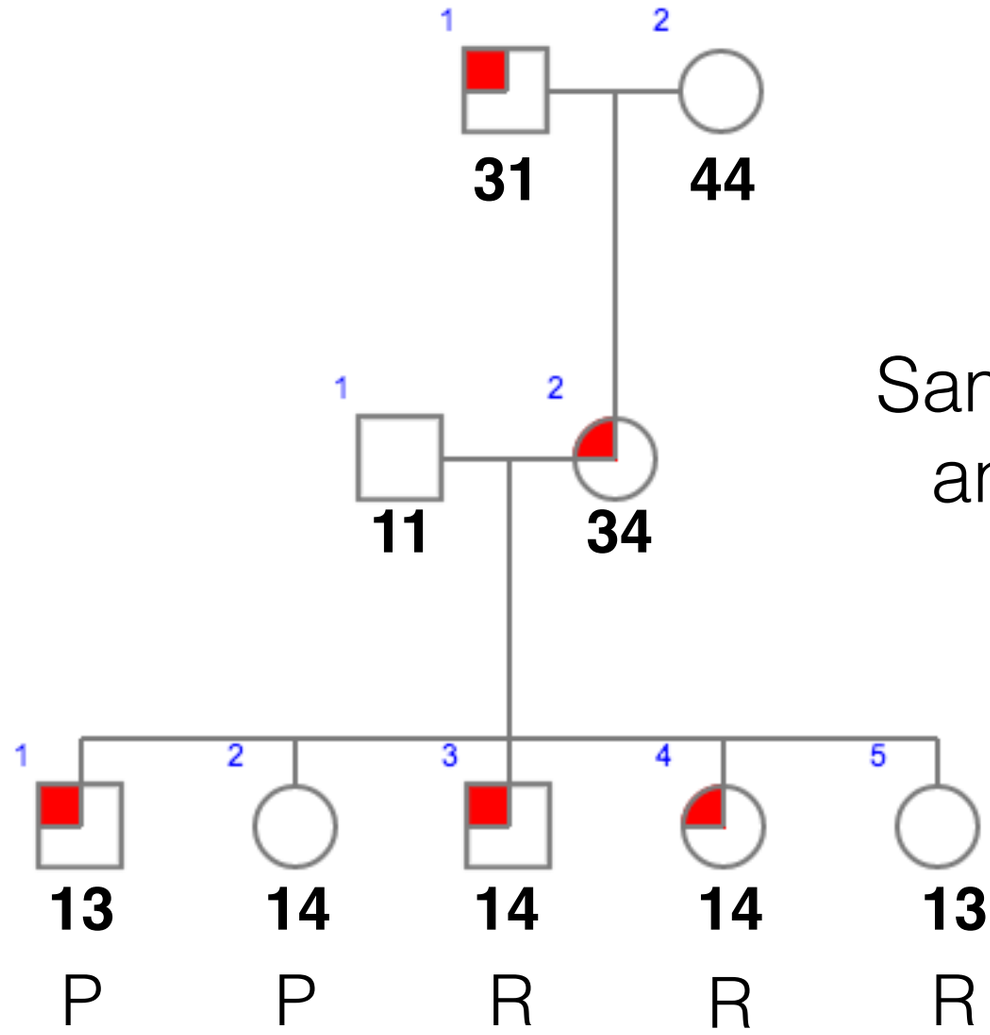
Let’s add another family



Same marker as before and we know phase

$$LOD = \log_{10} \frac{(1 - \theta)^P \times \theta^R}{0.5^{(P + R)}}$$

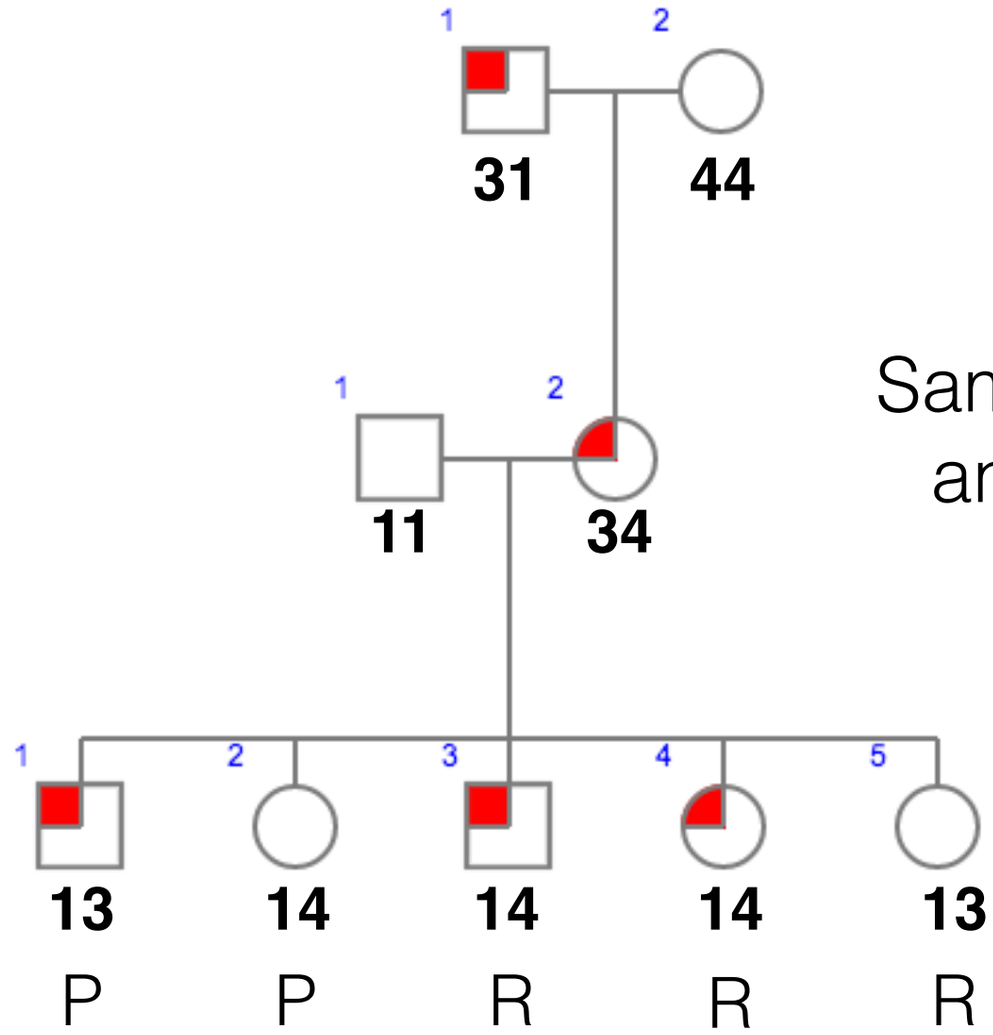
Let’s add another family



Same marker as before and we know phase

$$LOD = \log_{10} \frac{(1 - \mathbf{0.125})^2 \times \mathbf{0.125}^{\,3}}{0.5^{(2+3)}}$$

Let’s add another family



Same marker as before and we know phase

LOD = **-0.1159** $\theta = 0.125$

Change theta to 0.6? No, 0.5 is unlinked