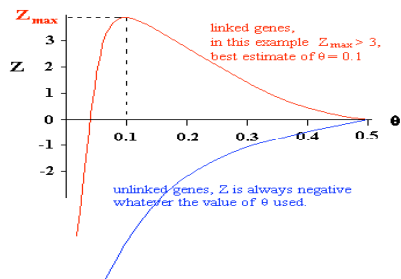


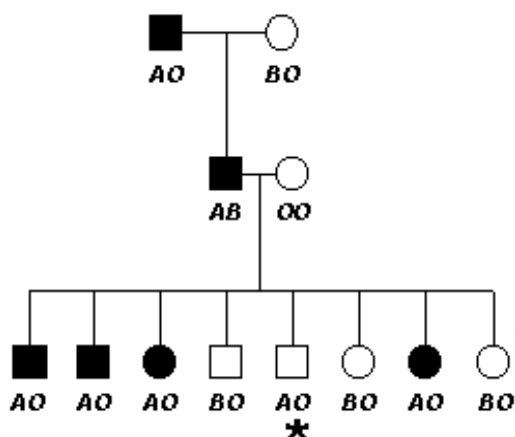
MAPPING MENDELIAN TRAITS IN HUMANS



UNDERSTANDING THE LOD SCORE METHOD

ONCE WE HAVE OBTAINED A MARKER THAT APPEARS TO BE CO-SEGREGATING WITH OUR DISEASE LOCUS WE APPLY THE LOD SCORE METHOD TO DETERMINE THE LIKELIHOOD THAT THE TWO LOCI ARE LINKED RATHER THAN JUST INDEPENDENTLY ASSORTING. WE HAVE TO TAKE A PROBABILITY APPROACH BECAUSE OUR SAMPLE SIZE IN MOST FAMILIES IS TOO SMALL TO JUST CALCULATE A RECOMBINATION FREQUENCY. LET'S GO THROUGH OUR NAIL-PATELLA SYNDROME EXAMPLE.

NAIL-PATELLA SYNDROME AND LINKAGE TO THE BLOOD TYPE LOCUS



IN THE PEDIGREE ON THE LEFT, IT APPEARS THAT AN ALLELE, A, OF THE BLOOD TYPE LOCUS IS CO-SEGREGATING WITH THE NAIL-PATELLA SYNDROME. WHAT WE NEED TO DETERMINE IS IF THE BLOOD TYPE LOCUS IS LINKED OR IF IT IS INDEPENDENTLY ASSORTING. TO DO THIS WE TAKE THE LOGARITHM OF ODDS (LOD) SCORE FOR A RANGE OF RECOMBINATION FREQUENCIES UP TO 0.50 (50%). WE DON'T GO HIGHER THAN 0.50 BECAUSE THIS IS THE MAXIMUM RECOMBINATION FREQUENCY THAT CAN BE ACHIEVED BETWEEN TWO LOCI. AT 0.50, THE GENES ARE EITHER REALLY FAR APART ON THE SAME CHROMOSOME AND APPEAR TO BE INDEPENDENTLY ASSORTING OR THEY ARE ON DIFFERENT CHROMOSOMES. NOW WE DETERMINE THE PROBABILITIES FOR EACH

Nail-patella syndrome = ● or ■

Blood Types = OO, AB, BO, AO

INDIVIDUAL DISTANCE USING THE FOLLOWING EQUATION:

$$\text{LOD} = Z(\theta) = (n - r) \log [2(1 - \theta)] + r \log (2\theta)$$

θ = Recombination Frequency we are testing

n = Total number of informative progeny

r = Number of recombinants

IN THE PEDIGREE ABOVE WE HAVE DETERMINED THAT A IS TRAVELING WITH THE NAIL-PATELLA TRAIT. IN OTHER WORDS, THE PHASE IS KNOWN. PEDIGREES IN WHICH THE PHASE IS UNKNOWN CAN STILL BE USED BUT IT REQUIRES MORE SOPHISTICATED STATISTICAL ANALYSIS THAT WE WILL NOT DISCUSS HERE. THE

OTHER IMPORTANT POINT ABOUT THE PEDIGREE IS THAT IT IS **INFORMATIVE**. THIS IS BECAUSE WE KNOW THE PHASE AND THERE ARE ENOUGH ALLELES OF THE MARKER IN THE FAMILY THAT WE CAN UNAMBIGUOUSLY IDENTIFY NON-RECOMBINANT AND RECOMBINANT PROGENY. THIS IS EXTREMELY IMPORTANT. IF A FAMILY IS PHASE UNKNOWN OR DOES NOT CARRY ENOUGH DIFFERENT ALLELES OF THE MARKER THIS CAN PREVENT IDENTIFICATION OF NON-RECOMBINANTS AND RECOMBINANTS AND WE CANNOT USE THEM FOR CALCULATING A LOD SCORE AT THIS MARKER.

WE HAVE 8 INFORMATIVE PROGENY AND 1 RECOMBINANT IN OUR PEDIGREE. NOW WE CAN DETERMINE THE LOD SCORE FOR EACH RECOMBINATION DISTANCE AND THIS WILL GIVE US THE MOST LIKELY DISTANCE BETWEEN THE BLOOD TYPE LOCUS AND THE NAIL-PATELLA LOCUS

IF THE TWO LOCI ARE VERY CLOSE TOGETHER, θ COULD BE 0.05 (5% RECOMBINATION FREQUENCY). TO DETERMINE THE PROBABILITY OF THIS POSSIBILITY WE ENTER THE APPROPRIATE NUMBERS IN OUR LOD SCORE EQUATION:

$$Z(\theta) = (n - r) \log [2 (1 - \theta)] + r \log (2 \theta)$$

$$Z(0.05) = (8 - 1) \log [2 (1 - .05)] + 1 \log (2 * 0.05)$$

$$Z(0.05) = 0.95$$

WHAT IF THE RECOMBINATION FREQUENCY IS 0.1 BETWEEN THE TWO LOCI?

$$Z(\theta) = (n - r) \log [2 (1 - \theta)] + r \log (2 \theta)$$

$$Z(0.1) = (8 - 1) \log [2 (1 - .1)] + 1 \log (2 * 0.1)$$

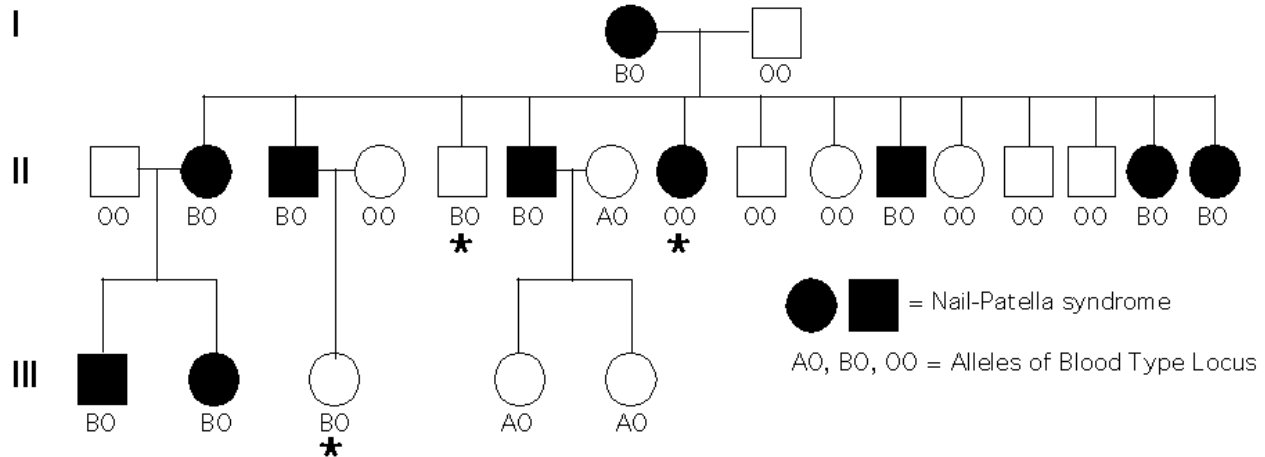
$$Z(0.1) = 1.09$$

WE FINISH THE ANALYSIS BY DETERMINING THE LOD SCORE FOR SEVERAL MORE PROBABLE DISTANCES OUT TO A MAXIMUM OF 0.5. IN OUR EXAMPLE IN CLASS WE DETERMINED SEVERAL MORE SCORES OUT TO 0.25 AND FOUND A MAXIMUM LOD SCORE AT 0.125.

RECOMBINATION DISTANCE (θ)	LOD SCORE
0.05	0.95
0.10	1.09
0.125	1.12
0.15	1.09
0.20	1.03
0.25	0.92

WHAT DOES THESE DATA MEAN? THESE SCORES MEAN THAT OUR MOST LIKELY DISTANCE BETWEEN THE NAIL-PATELLA LOCUS AND THE BLOOD TYPE LOCUS WE ARE USING AS A MARKER IS 0.125 OR 12.5 MAP UNITS. **THE LOD SCORE MEANS THAT THIS POSSIBILITY IS ABOUT 10 TIMES MORE LIKELY THAN THE 2 LOCI BEING UNLINKED. HUMAN GENETICISTS ACCEPT A LOD SCORE OF 3 (OR HIGHER) AS BEING AN INDICATION OF SIGNIFICANT LINKAGE. THIS SCORE REPRESENTS THE PROBABILITY OF BEING**

LINKED AT A PARTICULAR DISTANCE AS 1 000 TIMES MORE LIKELY THAN BEING UNLINKED. HOW DO WE ACHIEVE A HIGHER NUMBER FOR OUR ANALYSIS OF NAIL-PATELLA SYNDROME? WE NEED MORE FAMILIES THAT ARE INFORMATIVE FOR LINKAGE TO THE BLOOD TYPE LOCUS THAT WE CAN USE TO CALCULATE ANOTHER SET OF LOD SCORES. THEN WE CAN ADD THE LOD SCORES FOR BOTH FAMILIES TO ACHIEVE A TOTAL LOD SCORE FOR EACH PROBABLE DISTANCE BETWEEN THE BLOOD TYPE LOCUS AND THE NAIL-PATELLA LOCUS.



IN FAMILY 2 THERE ARE A COUPLE POINTS TO MAKE BEFORE WE DETERMINE OUR LOD SCORES. FIRST, IT IS INFORMATIVE AND THE PHASE IS KNOWN IF WE ASSUME THAT THE FEMALE IN GENERATION I RECEIVED B AND THE DISEASE FROM A PARENT (NOT SHOWN). IN THIS FAMILY THE B ALLELE APPEARS TO BE SEGREGATING WITH THE NAIL-PATELLA LOCUS. THE SECOND IMPORTANT POINT IS THAT IN THIS FAMILY IT IS THE B ALLELE AND NOT THE A ALLELE LIKE IN THE FIRST FAMILY. THIS IS PERFECTLY FINE. FOR LINKAGE STUDIES, WE ARE ONLY LOOKING TO SEE IF ANY ALLELE IS SEGREGATING WITH THE PARTICULAR SYNDROME WE ARE STUDYING. IT WILL PROBABLY BE DIFFERENT ALLELES OF THE MARKER IN UNRELATED FAMILIES.

IN THIS FAMILY THERE ARE 18 INFORMATIVE INDIVIDUALS AND 3 RECOMBINANTS (LABELED WITH AN ASTERISK). NOW WE DETERMINE THE LOD SCORES AT THE SAME RECOMBINATION DISTANCES AS FAMILY ONE.

- FOR RECOMBINATION DISTANCE OF 0.05:
 $Z(\theta) = (n - r) \log [2 (1 - \theta)] + r \log (2 \theta)$
 $Z(0.05) = (18 - 3) \log [2 (1 - .05)] + 3 \log (2 * 0.05)$
 $Z(0.05) = 1.18$

- FOR RECOMBINATION DISTANCE OF 0.10:
 $Z(\theta) = (n - r) \log [2 (1 - \theta)] + r \log (2 \theta)$
 $Z(0.10) = (18 - 3) \log [2 (1 - .10)] + 3 \log (2 * 0.10)$
 $Z(0.10) = 1.73$

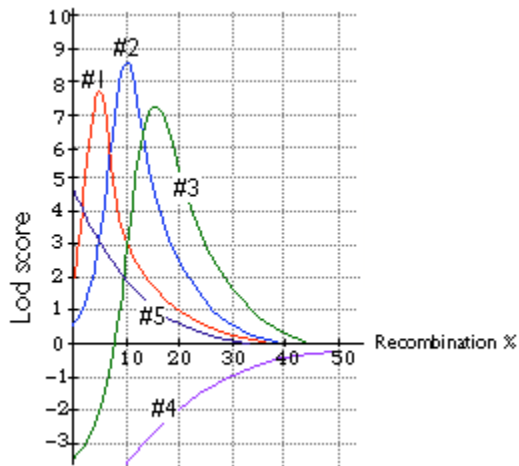
WE DO THIS FOR ALL THE RECOMBINATION FREQUENCIES THAT WE CALCULATED FOR THE FIRST FAMILY AND THEN ADD THEM TOGETHER:

RECOMBINATION DISTANCE (θ)	LOD SCORE FAMILY 1	LOD SCORE FAMILY 2	TOTAL LOD SCORE
0.05	0.95	1.18	2.13
0.10	1.09	1.73	2.82
0.125	1.12	1.84	2.96
0.15	1.09	1.89	2.98
0.20	1.03	1.87	2.9
0.25	0.92	1.74	2.66

NOW WE CAN SEE THAT WITH TWO INFORMATIVE FAMILIES WE ARE MUCH CLOSER TO SHOWING SIGNIFICANT LINKAGE ($LOD > 3$) FOR THE BLOOD TYPE LOCUS AND NAIL-PATELLA SYNDROME. NOW LETS EMPHASIZE SEVERAL POINTS ABOUT OUR DATA SO FAR:

- OUR HIGHEST LOD SCORE IS ALMOST 3 IN THE RANGE OF 0.125 TO 0.15 RECOMBINATION FREQUENCY. IF WE DETERMINED THE LOD SCORES FOR 0.30, 0.40. AND 0.50 THEY WOULD PROBABLY CONTINUE TO DECREASE. IF WE OBTAIN ONE MORE INFORMATIVE FAMILY THE TOTAL LOD SCORES WOULD DEFINITELY PASS 3, GIVING US SIGNIFICANT LINKAGE. THUS OUR ANALYSIS TELLS US THAT THE MOST LIKELY SCENARIO IS THAT THE NAIL-PATELLA LOCUS AND THE BLOOD TYPE LOCUS ARE BETWEEN 12.5 AND 15.0 MAP UNITS AWAY.
- FAMILIES WITH MORE INFORMATIVE INDIVIDUALS GIVE YOU HIGHER LOD SCORE NUMBERS AT EACH POSITION BECAUSE THE SAMPLE SIZE IS LARGER
- IF WE WERE NEVER ABLE TO FIND ANOTHER FAMILY WITH THIS SYNDROME AND WE WERE LEFT WITH ONLY THE FIRST FAMILY, LINKAGE STUDIES WOULD NOT BE ABLE TO IDENTIFY A SIGNIFICANT LINKAGE FOR THIS DISORDER AT THIS MARKER LOCUS. WE WOULD HAVE TO TAKE ANOTHER APPROACH TO IDENTIFYING THE NAIL PATELLA GENE.
- IF THE BLOOD TYPE LOCUS WAS COMPLETELY UNLINKED, THEN ALL OF THE LOD SCORES WOULD HAVE BEEN NEGATIVE AT ALL OF THE DISTANCES

OUR NEXT STEP WOULD BE TO IDENTIFY NEW MARKERS SUCH AS MICRO-SATELLITE LOCI THAT ARE NEAR TO THE BLOOD TYPE LOCUS. WE WOULD ANALYZE OUR FAMILIES TO DETERMINE THE ALLELES THAT EACH FAMILY CARRIES, IF ANY SPECIFIC ALLELE IS SEGREGATING WITH THE DISORDER AND ARE THE FAMILIES INFORMATIVE. AS WE IDENTIFY NEW MARKERS AND DO LOD SCORES FOR THESE IN RELATION TO THE NAIL-PATELLA LOCUS WE SHOULD BE ABLE TO FIND MARKERS THAT ARE EVEN CLOSER TO THE GENE. WHAT WOULD THIS DATA LOOK LIKE. IF WE FOUND 5 INFORMATIVE MARKERS AND CALCULATED LOD SCORES FOR EACH MARKER WE COULD GENERATE THE FOLLOWING GRAPH.



WHAT DOES THIS GRAPH TELL US?

MARKERS 1,2, AND 3 ARE ALL LINKED TO THE NAIL-PATELLA LOCUS, ALTHOUGH AT DIFFERENT DISTANCES AWAY.

- MARKER 1'S HIGHEST LOD SCORE WAS CLOSE TO 8 AT ABOUT 0.05 (5%) RECOMBINATION FREQUENCY . THIS SUGGESTS THAT THE DISTANCE BETWEEN THIS MARKER AND THE NAIL-PATELLA LOCUS IS ABOUT 5 MAP UNITS.

- MARKER 2'S HIGHEST LOD SCORE WAS 8.5 AT ABOUT 0.10 (10%) RECOMBINATION FREQUENCY. THIS SUGGESTS THAT THE DISTANCE BETWEEN THIS MARKER AND THE NAIL-PATELLA LOCUS IS ABOUT 10 MAP UNITS.
- MARKERS 3'S HIGHEST LOD SCORE WAS A LITTLE MORE THAN 7 AT ABOUT 0.15 (15%) RECOMBINATION FREQUENCY. THIS SUGGESTS THAT THE DISTANCE BETWEEN THIS MARKER AND THE NAIL-PATELLA LOCUS IS ABOUT 15 MAP UNITS.

MARKER 4'S LOD SCORES WERE ALL NEGATIVE SUGGESTING THAT THIS MARKER IS NOT LINKED TO THE NAIL-PATELLA LOCUS. IT IS PROBABLY FAR AWAY FROM THE NAIL-PATELLA LOCUS ON THE SAME CHROMOSOME.

MARKER 5 APPEARS TO BE TIGHTLY LINKED TO THE NAIL-PATELLA LOCUS. HOW DO WE KNOW THIS? THIS MARKER'S HIGHEST LOD SCORE, ~4.5, WAS FOR A RECOMBINATION DISTANCE VERY CLOSE TO 0.0 (0%) RECOMBINATION FREQUENCY. SO IT IS HIGHLY PROBABLE THAT THIS MARKER IS TIGHTLY LINKED. WE WOULD THEN FIND THIS MARKER ON THE PHYSICAL MAP OF THE GENOME AND BEGIN TO LOOK FOR POSSIBLE CANDIDATE GENES FOR NAIL PATELLA SYNDROME.

WE CAN USE THIS TECHNIQUE FOR ANY HUMAN TRAIT/DISORDER FOR WHICH WE CAN FIND PEDIGREES THAT ARE SEGREGATING FOR THE DISORDER. THE FIRST STEP IS TO IDENTIFY A MARKER LOCUS THAT IS CO-SEGREGATING WITH THE TRAIT IN THE PEDIGREES. WITH THESE TWO PIECES OF INFORMATION YOU CAN BEGIN YOUR LINKAGE ANALYSIS AND THE QUEST FOR THE GENE RESPONSIBLE FOR THE TRAIT/DISORDER