- **complementation**
    - **complement** - 2 genes/mutations complement when one provides something that the other doesn't have
    - **complementation group** - things that don't complement each other
    - dominant mutations don't go into a complementation group
    - complementation tests can see if mutation is in the same gene or not - if it isn't in the same gene, then mutations complement each other and produce WT. if in the same gene, then they don't complement.
- **auxotroph** - something that requires something on the outside to survive
- **codominance** - when two parents hybrid together

- insert bayes' theorem in here for pedigree qs

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

- **test cross** - cross to a homozygous recessive
- **homologues** - 2 copies of a chromosome (one from mom, one from dad)
- **meiosis** -
    - **chromosome nondisjunction** - aberrant segregation of chromosomes
        - doesn't split evenly (either stage of meiosis)
    - **linkage**
        - unlinked - parentals = recombinants
        - weakly linked - parentals > recombinants
        - tightly linked - parentals > > > recombinants, maybe 0 recombs
        - map distance = 100 * #recomb/#gametes
- Experiments relating to gene linkage
    - cross heterozygotes w/ recessive - track # recombinants (3 factor cross)
        - use double recombinant #s
    - **tetrads**
        - 4 spores (4 haplotypes) = tetrad
            - parental ditypes spores (only parental combos) AB/ab
            - tetratype spores (all 4 combos) AB/ab/aB/Ab
            - NPD - 2, 2 of aB / Ab
        - tightly = distance = 100 * T/2 * #total tetrads
        - weakly = 100 * (T + 6NPD/2 * #total tetrads)
        - experiment - cross two strains where each has one mutation -
            - expect - PD - none can grow
            - NPD - half can grow
            - T - 1/4 can grow
- LOD - good luck lol

**recombinant inbred mice**

- want a strain, but w a specific mutation?
  - continuously breed w/ desired strain, and select ones who have the mutation.
- **QTL** - identify minor contributors

- probability with regards to ORF's
  - $p(stop) = 3/64$ or $p(stop) = p(UAA) + p(UAG) + p(UGA)$
  - prob of exactly n length ORF: $p(ORF) = (1 - p(stop))^n * p(stop)$
  - prob of at least n length ORF: $p(ORF) = (1 - p(stop))^n$
  - # fortuitous ORFs = #codons in genome * $\frac{3}{64} * 6 * p(above)$

**mutagens**

- base analog - incorporated into DNA
- base modifying agent - repair of DNA damage is error prone
- dna intercalator - polycyclic compounds fit between bases. mis-copying by polymerase

**GENE MUTATIONS**

- missense mutation: base change that converts codon into another
- nonsense mutation: converts a codon into a stop codon
- frameshift: addition or deletion of a base ST sequence is shifted T.T

- screen vs selection: selection kills stuff, screen is look at all, search for a desired phenotype
- **suppressor mutation**
  - back mutation - back to WT
  - intragenic suppressor - mutation in same gene that compensates
  - extragenic suppressor - compensating mutation in diff gene
    - nonsense suppressor, like in a mutation in the tRNA portion of a ribosome
- **transposons (like Tn5)**
  - phage reqs:
    - $P_{am}$ - P protein w/ nonsense mutation, can't replicate
    - $int^-$ - can't integrate
    - going into $Su^+$ bacteria means that amber mutation is resolved, which is why you need to go into a $Su^-$ strain
    - need Tn5 w/ KanR resistance, which is the only way you're going to integrate something from virus into

**TRANSDUCTION - moving mutations from one strain to another.**

- experiment
  - infect Tn5+ (Kan-resis) strain w/ phage
  - collect virus, infect w/ Tn5- (Kan-sens) strain

        - select w/ kanamycin
        - count recombinants
    - 3-factor cross -
        - get one that's like phage: Tn5, A-, B+ and infect bacteria w/ /A+, B-
        - and then the opposite - Tn5, A+, B- and infect bacteria w/ A-, B+
        - kan resis screen
        - then count recombs, WT in each cross
    - cotransduction freq = # w/ phenotype / # resistant
        - higher, the closer

**TRANSFORMATION** - clone a gene or find a specific protein

- R plasmid, with lots of resistances
- bacteria have M and R
- generating a library
    - E coli, restriction enzyme, mix ligate w/ plasmids, select
    - library generate from mutant into plasmids
    - transform a WT strain w/ plasmids, select
    - lyse open the ones selected for to identify gene

**PLASMID CONJUGATION**

- strains
    - F- - something w/o a F plasmid
    - F+ - something w/ an empty F plasmid
    - F' - something w/ an F plasmid w/ something on it
    - Hfr - something w/ an F plasmid that was integrated into it
- when x transfers long after everything else: its where the ORI is pointing (transferred when the F plasmid breaks back out)
    - this can help us determine gene location
- Insertion sequences flip things

- reads
    - make a config
    - paired ends indicate they are from the same piece, when repeated element
- sequence alignment
- the graph thing

**LINKAGE VS ASSOCIATION**

- linkage is the pedigrees
    - bad w/ polygenic traits (not det. by single gene)
    - non-allelic heterogeneity
- association - common mutations w/ mild effects
    - complex inheritance
    - large #s

- HW eq
  - $p = f(A/A) + f(A/a)/2$
  - $q = f(a/a) + f(A/a)/2$
  - $H_0$ - locus is at HWE
  - $H_{alt}$ - locus is not at HWE
  - assumptions (not happening)
    - assortative mating
    - population is small/shrinks
    - new mutation
    - selection (one allele is more favorable)
- mutation
  - once that happens, yikes
  - auto rec, auto dom, het cases
  - at steady state: $\Delta q_{sel} + \Delta q_{mut} = 0$
- **INBREEDING**
  - F = inbreeding coefficient, if parental generation is A1/A2 and A3/A4, what is the chance of target being A1/A1 or something, and then mul 4
  - 1st cousins = 1/16
  - brother-sister mating = 1/4
  - p(affected because of inbreeding) = $f(a/a) = F * q$
  - p(affected bc of random mating) = $q^2$

**LD/LE - LINKAGE DISEQUILIBRIUM**

- you have some variables
  - $p(AB), p(aB), p(Ab), p(ab)$
  - $f(A), f(a), f(B), f(b)$ - sometimes derived from adding above
    - make sure this is like #gametes/#total gametes
  - chi squared by deriving expected by multiplying
    - $p(AB) = f(A) * f(B)$ and so on
    - make SURE CHI SQUARED VALUES ARE IN TERMS OF EXPECTED POPULATION, NOT FREQ
  - $D, D', D_{max}, D_{min}, r, r^2$
  - $D_n = (1 - r)^n D_0$