BIS101 F2013 Lecture 6: Population Genetics

Reading

Ch. 20, skip 20.3,20.4,20.5

Ch. 14, skip 14.1 and 14.6

Notes

Seminar on Wed. 12pm. 3001 PES

Population genetics

What is a population?

a group of individuals of the same species, usually within some geographically delimited area, usually with the possibility of intermating

Evolution:

change in frequency of an allele over time (not the same as natural selection)

Frequency

- A1A1 20 A1A2 13 A2A2 17
- what's the genotype frequency of A1A1 ? 40% (20/50)
- what is the allele frequency of A2 ? (p=47% or 0.47 because 34+13/100)
- what is the frequency of A1 ? (1-p)=0.53 b/c has to add to 1

I come back in 10 years, pop is now A1A1 80 A1A2 52 and A2A2 68

Has evolution occurred ? (not at this locus)

Population genetics is the study of allele frequency change in populations

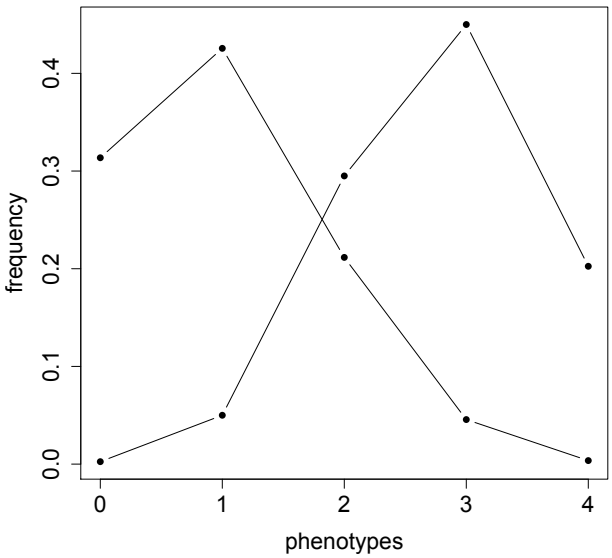
Could argue it is synonymous with evolution

We also care about it because it has effect on phenotypes!

In talking quant gen, we showed phonos for a cross where at each het we have 1:2:1 segregation in F2, but in a population phenotype distribution determined by allele freqs.

So distribution would be freq. of genotype * value, e.g.

Using our model of two loci with each big letter adding 1 to phenotype, @ locus A p_A=0.2 and at locus B p_B 0.3



Mean phenotype in the population = 1

Now change allele frequencies to pA=0.5 and pB=0.9 and redraw. New mean phenotype = 2.8. So allele frequencies will affect mean phenotype in pop without changing any of the genetics!

Hardy-Weinberg Equilibrium

- We talked about Mendel. HWE **NOT** 1:2:1
- Hardy & Weinberg 1908 (Hardy palyed crickey w/ Punnet)
    - How to solve provlem of blending & loss of diversity
- Model (what's a model ❓) (what's an equilibrium?)
    - model is a simplified description (mathematical) of a system
    - most of the time oversimplified -- "all models are wrong, some models are useful"
    - focus on the important parts of a system (noise in biology)
    - allow predictions of expected outcome & comparison to real data
    - if obs. data do not fit model -- biological interesting
    - assumptions wrong, try new model
    - what if data fit model? (my model is that aliens came down from outerspace and put chalk in the room)
    - data consistent w/ model doesn't PROVE model (that's how science works)

HW Model (write on board assumptions)

Assumptions:

- autosomal (which is?)
- locus, 2 alleles, diploid
- mendelian segregation
- random mating
- no other evol. forces (which?)

- no selection
- no migration
- no mutation
- no drift (large. pop size -> inf.)
- equal freq. in both sexes (or all hermaphrodites)
- generations discrete and nonoverlapping (annual plant) (explain)

Define variables (observed outcome of a system), parameters (things that define the model or system)

variables - X, Y, Z obs. freqs of 3 genotypes (AA, Aa, aa) in our sample

parameter p = freq. A1 allele, q=freq. A2 allele = 1-p (why?)

if we have sample (not whole pop) of 18 A1A1 and 24 A1A2 and 8 A2A2

X=0.36, Y=0.48, Z=0.16

p=X+(1/2)Y=0.6 (60 copies of A/100 total copies)

Equations

In next generation:

Gamete Table	prob. A1 from parent1	prob A2 from parent1
prob A1 from parent2	p*p	p*(1-p)
prob A2 from parent2	(1-p)*p	(1-p)*(1-p)

use prime to denote next generation

$X'=p^2$ $Y'=p(1-p)$ $Z'=(1-p)^2$

and p'=X'+Y'/2

and substituting: $p'=p^{2+2*p(1-p)/2}$ = p <- equilibrium

** In HWE, genetic composition (genotype and allele freqs) predictable w/ one parameter -> p **

Once in HWE, allele freqs. do not change (equilibrium) w/o disturbance

After a single generation of random mating -> HWE

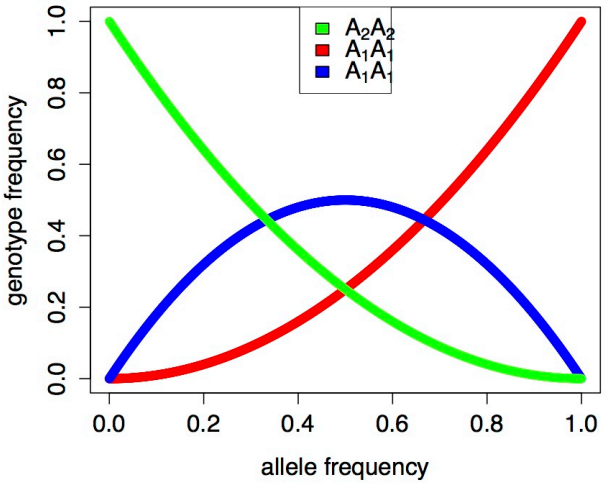
- start with 0.2 A1A1 and 0.8 A2A2 -> figure out what's p, and what's p' -> HWE after one generation

Conclusions

Single generation of random mating will almost always -> HWE

- start with 0.2 AA and 0.8 aa -> figure out p and p' -> HWE in one generation

Rare alleles more common in hets ($p^2 < 2pq$ for whenever p < q)



Does dominance change HWE?

- no because we haven't said anything about phenotype -- just genotype
- so will recessive alleles go extinct ???

Surprising # of loci in diff. organisms cannot reject HWE

- does this mean no selection, drift, mutation ??? why not?

Can use Chi-square to test if loci are in HWE!

- e.g.sample AA= 72 (68.1) Aa = 21 (28.9) aa = 7 (3.1)
    - X^2 >7.28 so not in HWE (tell them to show for selves)
- e.g. AA= 82 Aa = 38 aa = 5
    - is in nigh-perfect HWE (test it)

Violate an assumption

Can show effects of violating an assumption (nonequal # of sexes etc.)

Some assumptions have sm. FX

- nonequal sexes, takes longer than 1 gen of random mating to reach HWE but still reach it
- drift or selection -> never reach

nonrandom mating -- inbreeding

decrease in (2p(1-p)) beyond expectations

can use inbreeding coefficient F to detect

- F = 1-Ho/He = 1 - (obs # heterozygotes)/(expected number) = 1-X/2p(1-p)
- F is probability of IBD (identity by descent) of two alleles

- probability that allele picked at random from each individual is identical

Decrease in population size -> inbreeding.

Imagine only 10 unrelated individuals. Eventually everyone will be mating with a relative!

LD

We've looked at single loci, and multiple loci with recombination.

What is LD

- LD is nonrandom association of alleles at two loci -- if nonrandom assoc., loci are "in LD"
- LD ≠ linkage -- can have LD even if far apart or on different chromosomes
- what is linkage ???
- show phase A_B / a_b and ask:
    - will I ever see a_B ? yes
    - will A_B be more common in population or A_b ? (first)

Why care?

- mapping (usually don't sample REAL cause, but sample marker nearby)
- for mutations in complex traits, most important feature of genome
- history: demography, structure, selection

Unlinked loci

haplotype (define?): combination of alleles at multiple loci along a stretch of chromosomes

for two biallelic loci unlinked 4 gametes (phase!):

- Ab AB ab AB
- locus 1: A and a w/ freqs p_A and (1-p_A)
- locus 2: B and b w/ freqs p_B and (1-p_B)

Table: freqs of gametic haps if unlinked ?

(make sure class doesn't say 1:1:1:1)!

Hap	Exp. freq.
AB	p_Ap_B
Ab	p_A(1-p_B)
aB	(1-p_A)p_B
ab	(1-p_A)(1-p_B)

When observed = this, linkage equil

- if not, we say LD
- not the same as HWE -- not arrived at in one generation if out of Equil. (but eventually)
- but will eventually arrive

- we can measure these deviations (but will skip the math here)

How to get LD?

Lots of processes generate LD, including mutation, mating system, demographic change, and selection.

Example: if all ab die, then will only observe Ab aB and AB and will determine there is LD, even if A and B on diff. chromosomes!

selection can cause LD, even for loci on diff. csomes

How to lose LD

LD breaks down as things recombine. So things farther away will on average by at **?** lower LD because more recombination.

DRIFT

in HWE with N-> inf what happens to allele freqs over time (nothing)

in small pops, random chance is imp. (think about freq. heads depends on sample size)

Smaller N -> more drift; Bigger N -> less drift

Draw on board 6 individuals (1 AA 3 Aa 2 aa) (can do X^2 and show sample does not reject HWE)

- If this is whole pop.: (calculate freqs.) use die roller app to pick mates for next gen.
- Recalculate freqs. (has evolution happened?)
- do a couple more (maybe until fixation?)

Other things associated with drift (define):

- bottleneck
- founder effect

Other deviations from HWE cause drift -- separate sexes, etc, uneven offspring production. Humans have as much drift as a theoretical pop of size 10K in spite of being 7Billion of us!

Will skip the math, but:

Drift causes inbreeding: random mating in pop of sample 10, soon you're mating w/ relatives by random!

Chance of fixation = frequency. So most new mutations (at freq. 1/2N) are lost by to drift!

Differences between species: 2N*mu mutations per gen. * 1/2N chance of fixing = mu differences between species per gen. (or between genes, i.e. Ks)

Selection

Natural Selection not same as evolution

- change in frequency of a variant due to its effect on fitness
- multiple components to fitness: viability, mating success, fecundity
- Think in terms of relative fitness: some most fit genotype, and all other genotypes are competing with

it

- variant that makes you compete better and make more copies of your genes -> natural selection will increase freq.

Fitness Table

Genotype	A1A1	A1A2	A2A2
Freq.	p^2	2p(1-p)	$(1-p)^2$
Fitness	w11	w12	w22

relative fitness of AA is w_AA

- not faster than the bear, faster than the other guy

mean fitness of populatons is weighted avg.

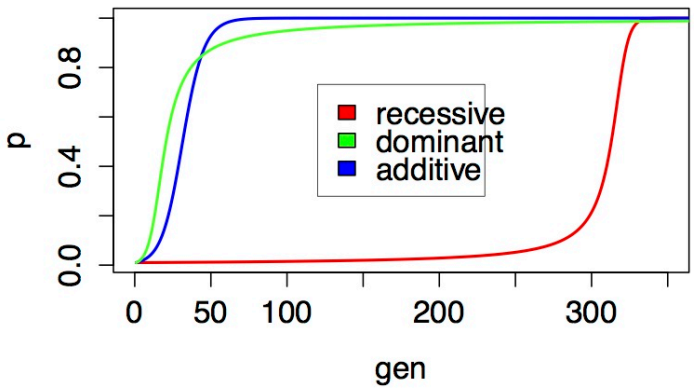
wbar = p^2w11+2p(1-p)w12+w22$(1-p)^2$

can do some math and show allele freq:

- p'= p^2 w11+p(1-p)w12)/wbar

results:

- change in allele freq. depends on difference in fitness b/t heterozygote and homozygote for the allele
- allele freq. (greater change with more middling allele freq)
- stronger s -> faster change in p,
- this general formula allows variation.
    - e.g. wAA=1 w12=1-s w22=1-2s -- when het is intermediate additive or codominant
    - w11=1 w12=1 w22=1-s (dominance of A1)
    - w11=1 w12=1-s w22=1-s (recessive A1)
- draw graph for recessive, dominant, codominant



Other forces

Gene flow: movement of genes from one population to another

- can impact allele frequencies and counteract selection and drift
- different pops should drift independently, but even one migrant/generation enough to prevent extensive divergence

Mutation

- fairly straightforward, increases freq of particular allele

Combos

Mutation-selection balance:

- assuming selection against recessive: $p=(mu/s)^{0.5}$
- for a completely recessive mutation, even under lethal selection (s=1) the freq. of A will be mu