Understanding the Hardy-Weinberg Equation

By:
Laura Martin

Understanding the Hardy-Weinberg Equation

By:
Laura Martin

Online:
< http://cnx.org/content/col10472/1.1/ >

C O N N E X I O N S

Rice University, Houston, Texas

Collection structure revised: October 22, 2007
PDF generated: February 4, 2011
For copyright and attribution information for the modules contained in this collection, see p. 35.

Table of Contents

Chapter 1

A Population Geneticist's Defintion of Evolution[1]

1.1 Are Humans Evolving?

In 2005, Stefasson *et al.* reported the fascinating discovery of an allele in humans whose presence is associated with increased fertility in Icelandic and European populations. Females with at least one copy of the allele have approximately 3.5%, and males 2.9%, more children on average than non-carriers. The exact mechanism by which the allele affects fertility is unknown.

1. **Predict:** Consider what you know about evolution. Do you think these populations are likely to be evolving with respect to this allele? Why or why not? Please make sure your response illustrates your understanding of evolution.

2. **Test:** Describe one way you could test your prediction **quantitatively**. What data would enable you to conclude that these human populations are evolving? What results would support the contention they are not?

1.2 Differential Survival and Reproduction Underpin Evolution

The fascinating discovery above should have called to mind the key causes, a.k.a. 'agents', of evolution

- natural selection
- sexual selection
- genetic drift (including bottle necks and founder effects)
- immigration/emigration
- mutation

and their consequences. All cause a population to evolve by altering the frequency (Definitions, p. 2) with which particular phenotypes, their underlying genotypes and most importantly the responsible alleles, occur. This quantitative description of the genetic consequences of these evolutionary mechanisms is encapsulated by the population geneticists definition of evolution:

- **Evolution** is a **change** in the **allele frequencies (Definitions, p. 2)** observed in a population over time (i.e. over generations).

Agents of evolution cause allele frequencies to change because they result in differential survival and reproduction. That is, not every individual has an equal chance of surviving, reproducing and contributing

[1] This content is available online at <http://cnx.org/content/m15262/1.1/>.

surviving offspring to the next generation. (The survival of offspring is key; if you reproduce but your kids don't survive to reproduce then you are evolutionarily inconsequential in terms of your allelic contribution to future generations.) Instead, for reasons that vary with the agent, some phenotypes, and their responsible genotypes, are more likely to survive or to reproduce and thus, to leave behind offspring than other phenotypes (genotypes). As a result, alleles of reproductively successful individuals become more common, and those associated with relatively unsuccessful individuals become less common, in subsequent generations. This change in allele frequencies is, of course, evolution.

Conversely, a population will not evolve if every phenotype (genotype) has an equal probability of surviving and producing surviving offspring. To imagine this, conjure a population in which all of the following conditions are simultaneously met:

- all phenotypes are equally likely to survive and to reproduce surviving offspring; there is no natural selection.
- all phenotypes are equally attractive or have equal access to potential mates; there is no sexual selection.
- no phenotypes leave behind more offspring than others just by chance; the population must be very large as there is no genetic drift.
- breeding individuals (and their genotypes) are not leaving or entering the population; there is no emigration or immigration.

The genetic consequence of all this equality is that the **same allele frequencies** are maintained from one generation to the next so the population does not evolve.

Of course, not all organisms reproduce sexually but the point is that a population will not evolve if all genotypes are equally likely to leave behind offspring with their alleles, even if reproduction occurs asexually.

3. Consider these definitions as you reflect your answers to questions 1 and 2 in the "Are Humans Evolving?" scenario above. Have you learned anything that would encourage you to modify your answers? If so, please do. If not, explain why your responses are appropriate.

Definitions

- **frequency** - the number of times an event or observation, for example a particular measurement or condition like blue eyes, is observed in a collection of events or observations like those comprising a sample, population or study. In this statistical sense, a frequency is equivalent to a proportion. For example, the frequency of a particular allele is equal to the number of times that allele is observed in a population over the total number of alleles for that locus in the population. Can be expressed as a fraction, a percentage, a decimal, or a probability.

Works Cited

- Stefansson, H., Helgason, A., Thorleifsson, G. et al. 2005. A common inversion under selection in Europeans. *Nature Genetics*. 37:129-137.

Chapter 2

Why are Allele Frequencies Maintained Across Generations When a Population is Not Evolving?[1]

In some respects, understanding how agents of evolution (p. 1) like natural selection, sexual selection and genetic drift drive changes in allele frequencies (Definitions, p. 7) is easier than understanding why in their absence allele frequencies remain unaltered from one generation to the next (p. 2). Understanding genetic equilibrium (Definitions, p. 7), however, is incredibly important as it forms the foundation of population genetics. It is the null hypothesis postulating the absence of evolutionary forces and, thus, against which the possibility of evolution is assessed. Work through the material below to increase your understanding of mechanism responsible for genetic equilibrium.

2.1 What is happening genetically when all individuals are equally likely to survive and reproduce?

Clearly for alleles to be perpetuated in a population through time, they must be passed from parent to offspring via reproduction. There is no other way (in the absence of continuous immigration)!

Thus, genetically speaking, when all individuals in a sexually reproducing population have an equal chance of surviving to reproduce and of producing surviving offspring:

- each individual has an equal chance of contributing one of the two required gametes to every fertilization event and thus, to the next generation.

A simple way to visualize this is depicted below. Each individual in this population holds two buckets of gametes. The buckets represent the two types of gametes the individual produces in equal numbers based on its genotype for a single locus. For example, Individual 1 with genotype **Aa** will produce equal numbers of **A** and **a** allele-containing gametes. In contrast, Individual 2's two buckets contain equal numbers of **A** allele-containing gametes reflecting its **AA** genotype. If you are unsure why we expect 50% of an individual's gametes to contain one of two alleles for a given locus and 50% the other allele for that locus, please review meiosis.

[1] This content is available online at <http://cnx.org/content/m15263/1.2/>.

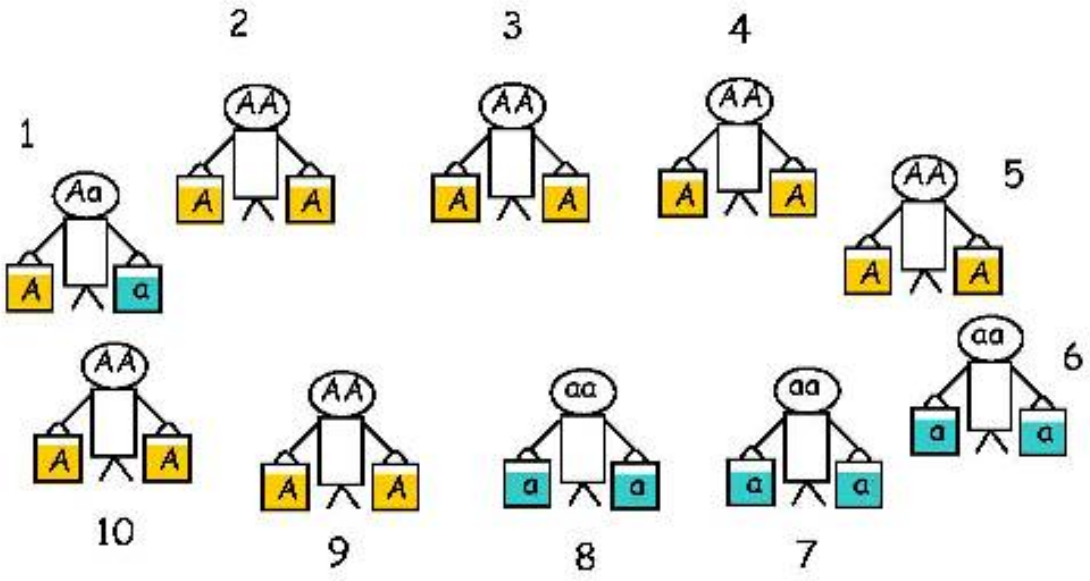

Figure 2.1: Population of 10 individuals. Each individual's genotype appears in its 'head'. Each individual's buckets represent the two types of gametes the individual produces in equal numbers based on its genotype for a single locus.

When no agents of evolution are acting on this population, each individual and therefore each bucket, because they contain equal quantities of gametes, has an equal likelihood (probability) of donating one of the two necessary gametes to a successful fertilization event and thus, to the next generation. Taken as a whole, this population of 10 individuals offers 20 buckets of gametes from which the two gametes for fertilization could possibly come. From which buckets the two gametes actually come depends upon which two individuals end up mating by chance and which of their two alleles the successful gamete contains.

To test your understanding, consider the questions below:

Exercise 2.1 ***(Solution on p. 8.)***

Imagine a situation in which a key nutrient the local bird population needs to build egg shells thick enough to withstand the weight of a parent during incubation occurs in very low levels. In this environment, the eggs of birds with **aa** genotypes crack twice as often during incubation as the eggs of **AA** and **Aa** individuals. A cracked shell always results in chick death.

Would this situation, and if so how, affect the likelihood that the alleles appearing in the next generation come from the buckets of **aa**, **AA** and **AA** individuals? Please explain.

Exercise 2.2 ***(Solution on p. 8.)***

In the situation called 'meiotic drive', a particular allele ends up in gametes more frequently than others for the same locus. That is, the usual expectation that on average 50% of an individual's gametes contain one allele for a given locus and 50% the other allele for that locus, is violated resulting in the one of the two alleles being overrepresented in gametes of heterozygotes. (The gametes of homozygotes are not affected.)

a. Review Figure 1. What aspect of this diagram would be altered? Why? Please explain.

b. Even if all individuals have an equal probability of mating in this population (i.e. mating occurs randomly), would all alleles have an equal probability of ending up in a fertilization event

and thus the next generation? Why or why not? Please explain.

Exercise 2.3 ***(Solution on p. 8.)***

In 2005, Stefasson *et al.* reported the fascinating discovery of an allele, H2, in humans whose presence is associated with increased fertility in Icelandic and European populations. Females with at least one copy of the allele have approximately 3.5%, and males 2.9%, more children on average than non-carriers. The exact mechanism by which the allele affects fertility is unknown.

Do all people in Icelandic and European populations have an equal probability of contributing one of the two gametes to each fertilization event that successfully produces an offspring? Please explain your conclusion.

Exercise 2.4 ***(Solution on p. 8.)***

Researchers investigating the H2 allele discussed in problem 3 hypothesized that this allele could be spreading through the population because of 'transmission disequilibrium' a situation analagous to meiotic drive in that offspring are more likely to inherit the H2 allele over the alternative H1 allele from a heterozygotic parent.

To investigate this, researchers genotyped 3,286 offspring of parents, in which one parent was heterozygous for H2 and the other parent homozygous for the alternative H1 allele, and found that 1,614 of these offspring carried the H2 allele (Stefasson *et al.*, 2005). Do the data suggest that this allele is spreading through the population as a result of transmission disequilibrium? Yes or no? How do you know? Please explain.

2.2 So why do equal probabilities of survival and reproduction reproduce parental allele frequencies in the offspring generation?

Hopefully the answer to this question is more obvious now that you understand that an equal likelihood of surviving and producing surviving offspring is really about each individual having an equal chance of providing one of the two gametes to every fertilization event that occurs in a population and produces a surviving offspring. Nothing is operating to bias the chances of an individual contributing a gamete to each and every fertilization event and every individual's two alleles have an equal (50:50) chance of ending up in the gamete involved in fertilization. But in case its not quite clear, let's explore the issue a bit further.

To do this, re-examine the pool of potential gametes depicted as buckets in the cartoon above (which is repeated below for your benefit) and answer the questions that follow.

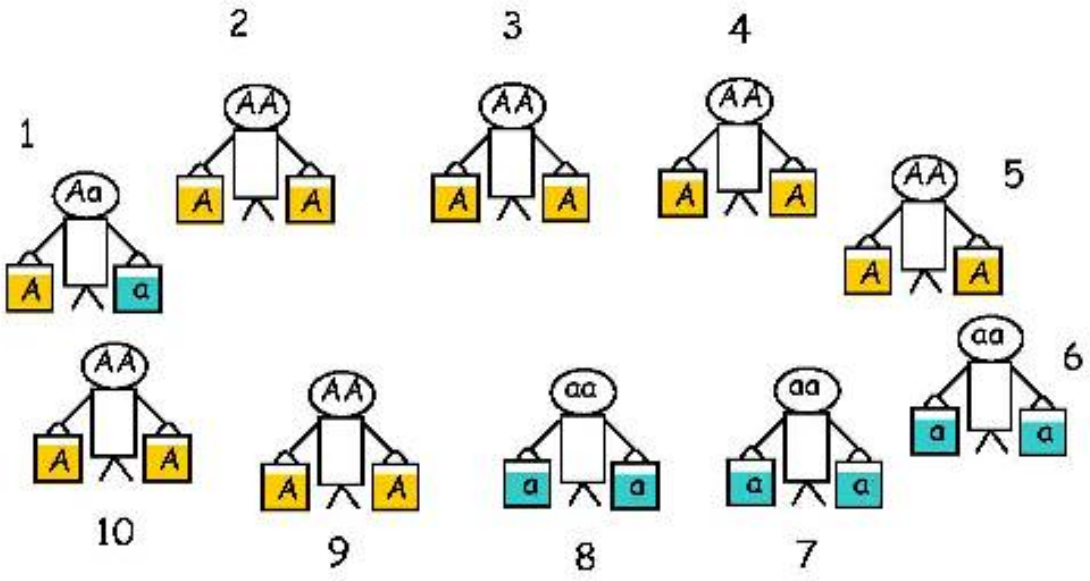

Figure 2.2: Population of 10 individuals. Each individual's genotype appears in its 'head'. Each individual's buckets represent the two types of gametes the individual produces in equal numbers based on its genotype for a single locus.

Exercise 2.5 *(Solution on p. 8.)*

a. If a gamete is randomly selected from this population, as happens when each individual has an equal chance of surviving and reproducing, what is the likelihood (probability) that it contains an **A** allele?

b. Explain the reasoning behind your response to question a.

Now consider the following questions.

Exercise 2.6 *(Solution on p. 8.)*

a. If 15 random matings took place in this population, requiring a total of 30 gametes to be randomly selected, what is the likelihood that each randomly selected gamete contains an **A** allele? How about an **a** allele?

b. Explain the reasoning supporting your response to question a.

c. Imagine all 30 of the gametes resulting from 15 random matings displayed on a piece of paper in front of you, how frequently would you expect the **A** allele to occur in this collection? The **a** allele? Why? Please explain.

d. Review your responses to questions a and c. What do they tell you about how likely (frequently) the **A** allele is to appear in the offspring population? The **a** allele? Why? Please explain.

Now that we know the frequency with which we expect **A** and **a** alleles to appear in the offspring generation when all individuals in a population have an equal probability of surviving and producing surviving offspring, let's explicitly and consciously compare them to the frequency with which these same alleles appear in the parent population and, finally, reflect on the question titling this section.

Exercise 2.7 *(Solution on p. 8.)*

a. How does the probability (likelihood) that an allele will make it into the offspring generation compare to the frequency with which that allele occurs in the parental generation?

b. Compare your responses to Problems 1, 2a, 2c, and 3a. What do they suggest about the relationship between the frequency with which an allele appears in the parental generation, its probability of appearing in parental generation gametes, and ultimately its frequency in the offspring generation when mating is random? Why? Please explain.

Exercise 2.8 *(Solution on p. 8.)*

In reality, would you expect offspring generation allele frequencies to always be perfectly identical to those of the parental generation when mating is random? Why or why not? Please explain.

Exercise 2.9 *(Solution on p. 9.)*

Test your understanding by returning to the scenario depicted in Problem 3 of the previous section. It turns out that the allele associated with increased fertility, referred to as H2, is found in 21% of the loci of people of European descent.

a. If this population is not evolving with respect to this allele, how frequently should this allele occur in this population 200 years from now? Why? Please explain.

b. Draw a figure (graph) illustrating your prediction from part a. Please make sure to label your axes and include a figure legend (Definitions, p. 7).

Definitions

- **frequency** - the number of times an event or observation, for example a particular measurement or condition like blue eyes, is observed in a collection of events or observations like those comprising a sample, population or study. In this statistical sense, a frequency is equivalent to a proportion. For example, the frequency of a particular allele is equal to the number of times that allele is observed in a population over the total number of alleles for that locus in the population. Can be expressed as a fraction, a percentage, a decimal, or a probability.
- **genetic equilibrium** - state of a population in which allele frequencies remain unchanged from one generation to the next.
- **legend** - a one or two sentence description of the variables depicted in a figure (graph).

Works Cited

- Stefansson, H., Helgason, A., Thorleifsson, G. et al. 2005. A common inversion under selection in Europeans. *Nature Genetics*. 37:129-137.

Solutions to Exercises in Chapter 2

Solution to Exercise 2.1 (p. 4)
When incubation success varies with genotype as described above, alleles from the 'buckets' of **AA** and **Aa** would be twice as likely as alleles from the 'buckets' of **aa** individuals to make it into the next generation.

Solution to Exercise 2.2 (p. 4)
a. The relative quantities of gametes in the 'buckets' of heterozygotes would no longer be 50:50. 'Buckets' corresponding to the allele that ends up in gametes more frequently would contain a larger quantity of gametes than 'buckets' representing the alternative allele.

b. No. All alleles would not have an equal likelihood of ending up in fertilization events even when all individuals are all equally likely to mate. This is true because, when a mating event involves a heterozygote, they will be more likely to contribute the over-represented allele than the alternative allele to fertilization. This is true because more than half their gametes contain the over-represented allele.

Solution to Exercise 2.3 (p. 5)
No, all individuals in these populations do not have an equal probability of contributing one of the two gametes to each fertilization event that produces a surviving offspring. This is true because individuals carrying at least one copy of the allele are more likely to successfully conceive, i.e. have more successful fertilization events, than those that do not carry it.

Solution to Exercise 2.4 (p. 5)
No, the data suggest that this allele is not spreading through the population as a result of transmission disequilibrium because 49% of the offspring of heterozygotes [(1,614/3,286)*100] carry the H2 allele. This compares favorably to the expectation that, if transmission rates are not biased, approximately 50% of the offspring of heterozygotes will carry the H2 allele (and approximately 50% the H1 allele). The two 'buckets' of heterozygotes appear to contain equal quantities of H2 and H1 alleles.

Solution to Exercise 2.5 (p. 6)
Much as there is a 13 in 20 chance of blindly pulling a gold chip out of a bag of 20 chips of which 13 are gold and 7 blue (and they differ in no other way), there is a 13 in 20 or 65% chance that the gamete would contain an **A** allele, if a a gamete were randomly selected from the population above (13/20 = 65/100 = 0.65*100 = 65%). Buckets of **A** alleles are nearly twice as common as buckets of **a** alleles. Consequently when reproduction is random, as happens when every individual has an equal probability of surviving and reproducing, any randomly selected parent is nearly twice as likely to be carrying an **A** allele as opposed to an **a** allele.

Solution to Exercise 2.6 (p. 6)
Ideally, your responses to questions a and c were identical. For each of the 15 random mating events, the probability that a contributed gamete contains an **A** allele is 13/20 (65%) and an **a** allele 7/20 (35%). This is true because each gamete selection event is independent; the allele one gamete contains does not in anyway influence what allele the second gamete of a fertilization event will contain if mating is random. (This is not true is mating is not random. Can you provide an example?)

Since the above is true for each individual mating event, the overall allelic composition of the resulting offspring generation will simply reflect the probability associated with picking each allele during each random gamete selection event summed for all 30 events. Therefore, on average 65% of the 30 alleles will be **A** and 35% **a** in the offspring generation if mating is truly random.

Solution to Exercise 2.7 (p. 6)
As is hopefully now clear, when all individuals in a population have an equal probability of surviving and reproducing successfully the probability that an allele ends up in a fertilization event and thus, in the offspring generation is equal to the frequency with which that allele appears in the parent generation. That is, when no agents of evolution are acting on a population, the allele frequencies observed in the offspring generation will be the same as those observed in the population producing them.

Solution to Exercise 2.8 (p. 7)
Of course in practice, all populations are subject to genetic drift (an agent of evolution) where, just by chance, some individuals will reproduce more frequently than others making a disproportionate contribution

to the next generation. This effect will be most pronounced in small populations, like that in the example above, and least in very large ones.

Solution to Exercise 2.9 (p. 7)

a. In the absence of any evolutionary processes including genetic drift operating on this population, this allele should still occur in 21% of the population 200 years from now. This is expected because the H2 allele currently occurs in 21% of the population's loci, consequently 21% of all 'buckets' in the population contain this allele, leading 21% of the gametes involved in fertilization events to contain the H2 allele, causing the H2 allele to occur in 21% of the loci in the next generation. This will be repeated generation after generation for 200 years in the absence of an evolutionary force.

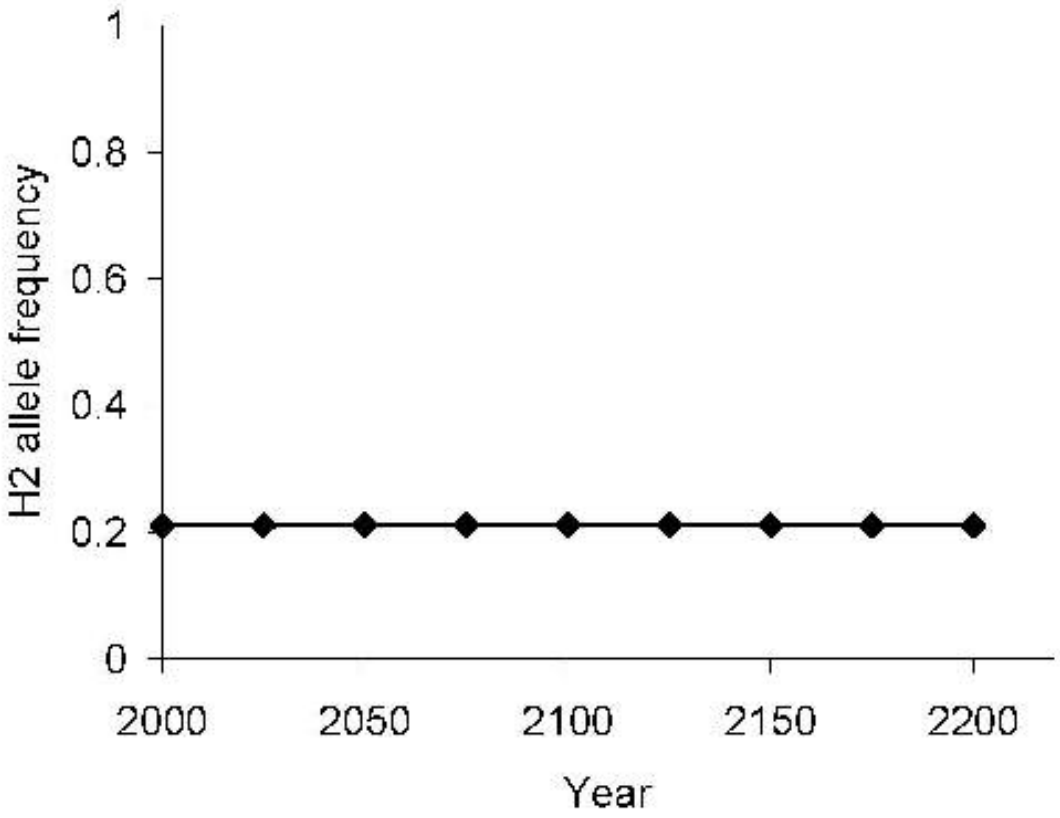

Figure 2.3: Expected frequency of the H2 allele over a 200 year period if the European population is not evolving with respect to this allele.

Chapter 3

Do Parental Generation Genotype Frequencies Affect the Probability that an Allele Will Appear in the Offspring Generation?[1]

When all individuals in a population have a truly equal likelihood of surviving and producing surviving offspring, i.e. no agents of evolution (p. 1) are operating on a population, alleles end up in gametes, and consequently in the offspring generation, in direct proportion to their frequencies (Definitions, p. 12) in the parental generation. This causes a population's allele frequencies to be perpetuated unchanged from one generation to the next (p. 3), a condition known as genetic equilibrium (Definitions, p. 12).

When reflecting on this process, it might be tempting to think that the way alleles are distributed among parents, i.e. parental genotype frequencies (Definitions, p. 12), could somehow affect the likelihood that an allele will be passed to the next generation. But does it? Let's explore this below.

Examine the two populations below.

[1] This content is available online at <http://cnx.org/content/m15267/1.1/>.

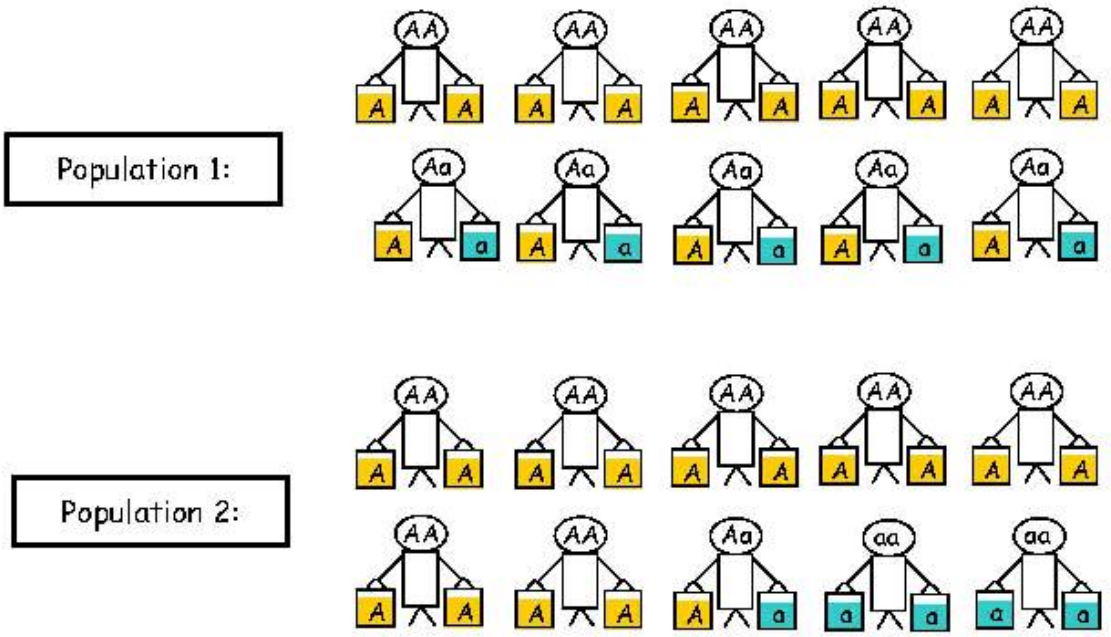

Figure 3.1

1. Confirm for yourself that although the two populations have identical allele frequencies their genotype frequencies differ. (That is, how often a particular genotype appears in a population differs between the two populations even though the two alleles are equally common in both populations.) Please be sure to calculate allele and genotype frequencies for both populations.

2. **Make predictions:** Imagine that individuals within each population mate randomly to produce an offspring generation, one per population.

a. Do you think that, for each population individually, allele frequencies in the offspring generation will be identical to those of the parental generation? Why or why not? Please explain the reasoning supporting your prediction.

b. Do you think that the allele frequencies of these two independent offspring generations will be identical to each other? Why or why not? Please explain the reasoning supporting your prediction.

3. **Test your predictions:** Use the SIMULATION to test your predictions described in questions 2a and b. Be sure to record your data. You may run the simulation multiple times if you like.

4. **Interpret your data:** Compare your results from question 3 to the predictions you made as part of questions 2a and b. Do the data support or refute your predictions? Please explain making sure to support your conclusions with your data.

5. **Draw a conclusion:** Do differences in parental genotype frequencies affect the probability that an allele will appear in the offspring generation when parental allele frequencies are identical and mating is random? Why or why not? Please explain using data to support your conclusion.

Definitions

- **frequency** - the number of times an event or observation, for example a particular measurement or condition like blue eyes, is observed in a collection of events or observations like those comprising a sample, population or study. In this statistical sense, a frequency is equivalent to a proportion. For example, the frequency of a particular genotype is equal to the number of times that genotype is

observed in a population over the total number of genotypes for that locus in the population (equals the number of individuals in the population). Can be expressed as a fraction, a percentage, a decimal, or a probability.

- **genetic equilibrium** - state of a population in which allele frequencies remain unchanged from one generation to the next.

Chapter 4

Are Genotype Fequencies Necessarily Maintained Across Generations When A Population is Not Evolving?[1]

When all individuals in a population have an equal likelihood of surviving and producing surviving offspring, i.e. no agents of evolution (p. 1) operate on a population, alleles end up in gametes and, in turn the offspring generation in direct proportion to their frequencies (Definitions, p. 18) in the parental generation. This causes a population's **allele frequencies** to be perpetuated unchanged from one generation to the next, a condition known as genetic equilibrium (Definitions, p. 18).

Does the same thing hold true for **genotype** frequencies? That is, when every individual has an equal probability of surviving and producing surviving offspring **are the genotype frequencies of the parental generation necessarily replicated in the offspring generation?**

4.1 How do genotypes form when every individual is equally likely to survive and reproduce?

To investigate the relationship between parental and offspring generation genotype frequencies when a population is not being influenced by an agent of evolution, let's review how offspring genotypes are formed when every potential parent is equally likely to survive and reproduce.

Under these conditions, the allele contributed by one parent does not and is not influenced in anyway by the allele contributed by the second parent; there is no sexual selection. Nor is a parent's likelihood of contributing alleles to the next generation influenced by environmental circumstances that favor survival of one parental phenotype over another or by chance events as occur during genetic drift.

Thus, if survival and reproduction are truly random events, each parent's allelic contribution to fertilization is a statistically independent event. Put differently, the genotype of the resulting offspring is the result of two **independent**, **chance** events - one chance event per allele. Imagine drawing two parents at random from a population and randomly collecting one allele from each.

Test your understanding by answering the question below.

Exercise 4.1 ***(Solution on p. 19.)***

If female[2] peacocks prefer to mate with males who have a larger number of eye-spots[3] on their elaborate tail feathers over males with fewer eye-spots (and the number of eye-spots is heritable),

[1] This content is available online at <http://cnx.org/content/m15266/1.1/>.
[2] http://www1.istockphoto.com/file_thumbview_approve/1637902/2/istockphoto_1637902_peacock_and_female.jpg
[3] http://www.birding.in/images/Birds/indian_peacock.jpg

is the likelihood that a male will contribute alleles to the next generation independent of, i.e. unaffected by, the female? Why or why not? Please explain.

4.2 How do we calculate the probability that a particular genotype will form when every individual is equally likely to survive and reproduce?

When the independent events condition described above (p. 15) is met, a given **genotype** is expected to appear in the offspring generation with a likelihood equal to:

- the mathematical product (multiplication) of the **frequency** with which each **allele** forming the genotype occurs in the parental generation.

This is just a simple rule of probability. The likelihood that an event, composed of two or more independent events, will occur is equal to the mathematical product of the likelihood that each event will occur independently.

Review the bulleted definition above to see that it also makes biological sense. The likelihood that a given genotype will form is equal to the probability of picking first one then the second of the two alleles necessary to form it. And, the probability of picking an allele directly reflects its relative commonness (frequency) in the parental population (p. 15).

Finally, note that this likelihood also represents the **frequency** with which we expect to observe the genotype in the offspring generation. This makes sense because how frequently we expect to see the genotype in the offspring generation is the direct result of how likely it is to form. Genotypes that have a low probability of forming, because the alleles that comprise them are relatively rare in the parental generation, will appear infrequently in the offspring generation.

4.3 Is it exactly that simple?

There is one more thing to consider when determining the probability that given genotype will appear in the next generation; that is the number of ways a particular genotype can form.

Remember that the likelihood a particular genotype will form is equal to the probability of picking first one then the second of the two alleles necessary to form it. There is only one way to form a homozygote; both parents must donate the same allele. For example, the **AA** genotype can only form if, when one parent donates an **A**, the other does as well. Heterozygotes, in contrast, can form in two ways; 'parent one' can donate an **A** and 'parent two' an **a** or 'parent one' can donate an **a** and 'parent two' an **A**.

The fact that there are two routes to heterozygote formation must be taken into consideration when calculating the likelihood that the heterozygous genotype will occur in the offspring generation. To account for this, **you must multiply the probability that a heterozygote will form by two, i.e. multiply the value described in the bulleted point above by two.**

You should now have the tools to answer the question posed at the start of this module:

4.4 Are parental genotype frequencies necessarily reproduced in the offspring generation when all individuals are equally likely to survive and reproduce?

To answer the question titling this section and the module, let's start by asking what you need to know. Then let's figure out how you can get that information. To do this, review the question, the information provided in previous sections and develop a table similar to that below.

What do I need to know to answer this question?	How do I get this information?

Table 4.1

Check your outline above by working through the example below.

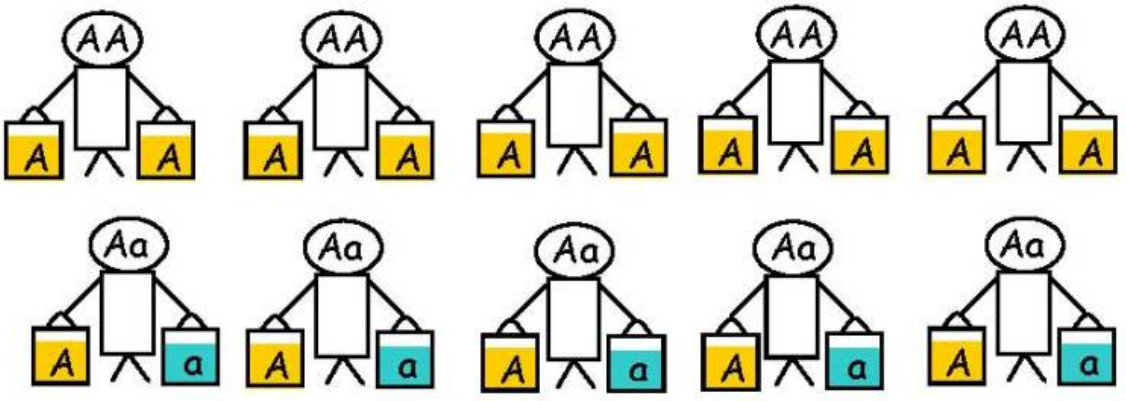

Figure 4.1

Exercise 4.2 ***(Solution on p. 19.)***

Review Figure 1. Imagine that these parents mate randomly. What genotypes could their offspring exhibit? Put another way, what allele combinations could we see in the offspring generation if every parent has an equal chance of reproducing and thus contributing an allele to the next generation?

Exercise 4.3 ***(Solution on p. 19.)***

Calculate how frequently we expect to observe each possible genotype identified in problem 1. Please show your work. Refer to the outline you developed or reread the preceding sections for guidance.

Now check your work: have you identified every possible genotype that could appear in the offspring generation when each parent has an equal chance of reproducing in the population above?

To do this, consider the following: if we have identified every possible genotype, then the frequencies calculated in question 2 will sum to 1 (i.e. 100%) because we will have accounted for 100% of the genotypes that could possibly occur in the offspring if all parents have an equal chance of reproducing and contributing alleles to the next generation.

Exercise 4.4 ***(Solution on p. 19.)***

Check your work. Have you omitted any possible offspring genotypes? How do you know? Please explain.

Exercise 4.5 ***(Solution on p. 19.)***

If you have missed a genotype, what one do you think you have missed? Why?

If you have missed one, correct your response to question 2 above. Now you should have an accurate description of the offspring genotype frequencies that will materialize when every parent has an equal chance of reproducing in this population.

Exercise 4.6 *(Solution on p. 19.)*

Finally, we are ready to address the question heading this module, are offspring generation genotype frequencies necessarily expected to be the same as parental generation genotype frequencies when every parent has an equal chance of surviving and reproducing? Please explain using your results to support your conclusion.

Exercise 4.7 *(Solution on p. 19.)*

How confident are you that parental genotypes are not necessarily recreated in the offspring generation when every parent has an equal probability of surviving and reproducing? Please explain.

Now let's confirm that random reproduction produces the offspring genotypes we predicted based on the parental allele frequencies above. To do so, we will run a simulation in which the computer will randomly pick 50 pairs of alleles to represent 50 random fertilization events.

Exercise 4.8 *(Solution on p. 19.)*

Are the genotype frequencies in the offspring generation produced by the simulation the same as those you predicted? Yes or no? Please explain.

Exercise 4.9 *(Solution on p. 19.)*

Do the data support the prediction you made about the similarity between parental and offspring genotype frequencies in problem 6 above? Yes or no? Please explain using evidence to support your conclusion.

Exercise 4.10 *(Solution on p. 19.)*

Imagine randomly mating the offspring generation just produced. Would you expect the genotype frequencies of their offspring to differ from their own? Yes or no? Please explain your conclusion and support it with evidence either logical or actual.

Exercise 4.11 *(Solution on p. 20.)*

What do the conclusions to problems 9 and 6 suggest about the potential for similarity between parental and offspring generation allele frequencies when a population is not subject to any agents of evolution? Please explain.

Exercise 4.12 *(Solution on p. 20.)*

a. Finally, return to the simulation and click the box marked X, run it and record the results. Notice that the parent generation is identical to that in Figure 1. Review the results. Do you think that the offspring generation is the product of random mating? Why or why not? Please explain using data to support your conclusion.

b. What do these results indicate about the similarity between the allele frequencies of the parent and offspring generations? Why? Please explain.

Definitions

- **frequency** - the number of times an event or observation, for example a particular measurement or condition like blue eyes, is observed in a collection of events or observations like those comprising a sample, population or study. In this statistical sense, a frequency is equivalent to a proportion. For example, the frequency of a particular allele is equal to the number of times that allele is observed in a population over the total number of alleles for that locus in the population. Can be expressed as a fraction, a percentage, a decimal, or a probability.
- **genetic equilibrium** - state of a population in which allele frequencies remain unchanged from one generation to the next.

Solutions to Exercises in Chapter 4

Solution to Exercise 4.1 (p. 15)
No, male and female contributions to a fertilization event are not independent. Females preferentially mate with males with a larger number of tail feather eye-spots. This favors the likelihood that the alleles associated with this trait will end up in the next generation over alleles associated with fewer eye-spots. This is sexual selection, i.e. non-random mating.

Solution to Exercise 4.2 (p. 17)
Since two alleles exist for this locus in this population, **A** and **a**, three possible genotypes can occur in the offspring of these parents. These three genotypes are:

- **AA**
- **aa**
- **Aa**

Recall that the last can form in two ways: **Aa** or **aA**.

Solution to Exercise 4.3 (p. 17)
The frequency with which a genotype will occur in the offspring generation is equal to (p. 16) the multiplicative product of the frequency with which each allele comprising the genotype appears in the parent generation.

Counting the buckets depicted in Figure 1 above reveals that the **A** allele occurs with a frequency equal to 15 out of 20 or 0.75 and the **a** with a frequency equal to 5 out of 20 or 0.25 in the parental generation. Consequently,

- the **AA** genotype will occur with a frequency equal to 0.75 x 0.75 = 0.5625 in the offspring generation.
- the **aa** genotype will occur with a frequency equal to 0.25 x 0.25 = 0.0625 in the offspring generation.
- the **Aa** genotype will occur with a frequency equal to (0.75 x 0.25) x 2 = 0.375 in the offspring generation.

Solution to Exercise 4.4 (p. 17)
We have not omitted any possible genotypes because the expected genotype frequencies (0.5625 + 0.0625 + 0.375) sum to 1 as expected if we have accounted for every possible genotype that could occur in the offspring generation of a randomly mating, parental generation with these allele frequencies.

Solution to Exercise 4.5 (p. 17)
If you missed one genotype, it was probably a heterozygote, either **Aa** or **aA**. Review the information found here (p. 16) to understand the likely source of your error.

Solution to Exercise 4.6 (p. 18)
No, parental and offspring generation genotype frequencies will not necessarily be the same in a population that is not evolving. Our data support this conclusion because in the parental generation above, the **AA**, **aa**, and **Aa** genotypes occur with frequencies equal to 5 out of 10 or 0.5, 0 out of 10 or 0.0 and 5 out of 10 or 0.5, respectively. Whereas, if these parents mate randomly, the **AA**, **aa**, and **Aa** genotypes will occur in their offspring with frequencies equal to 0.5625, 0.0625 and 0.375 respectively.

Solution to Exercise 4.7 (p. 18)
Any response in which you discuss elements about which you are confused or describe your understanding is acceptable.

Solution to Exercise 4.8 (p. 18)
Yes, they are. The **AA**, **aa**, and **Aa** genotypes occur with frequencies equal to 0.5625, 0.0625 and 0.375 respectively in the offspring generation of the simulation exactly as predicted.

Solution to Exercise 4.9 (p. 18)
Best answers describe the prediction you made in problem 6 and how the data from the simulation do or do not support that expectation.

Solution to Exercise 4.10 (p. 18)

The genotypes frequencies should not differ between these two generations. You can support this using data collected from the simulation or by calculating the genotype frequencies one expects to see in their offspring as a result of random mating by these parents. This is illustrated below. F1 refers to the offspring generation just produced and F2 refers to their offspring.

To calculate F2 genotype frequencies, the first pieces of information we need are the allele frequencies of their parent's generation, F1.

We know that the allele frequencies of the F1 generation are identical to those of their parent's generation (p. 19) because every parent had an equal likelihood of surviving and mating (i.e. they were not subject to an agent of evolution). Consequently, in the F1 generation

- **A** occurs with a frequency of 0.75
- **a** occurs with a frequency of 0.25

Since the same allele frequencies and the same method are used to calculate the genotype frequencies for both the F1 and F2 generations, the genotype frequencies of these two generations must be identical.

To confirm this, review below:

- the **AA** genotype will occur in the F2 generation with a frequency equal to 0.75 x 0.75 = 0.5625.
- the **aa** genotype will occur in the F2 generation with a frequency equal to 0.25 x 0.25 = 0.0625.
- the **Aa** genotype will occur in the F2 generation with a frequency equal to (0.75 x 0.25) x 2 = 0.375.

and these calculations are identical to those we made to calculate the genotype frequencies for the F1 generation in problem 3 (p. 19).

Solution to Exercise 4.11 (p. 18)

In some cases, parental and offspring generation genotype frequencies may be the same and in others they may differ. The results demonstrate that if a parental generation, like F1, is the product of random mating and is itself not subject to any evolutionary pressures then its own genotype frequencies will be replicated in its offspring as observed in F2.

Solution to Exercise 4.12 (p. 18)

a. The data suggest that the offspring were not the product of random mating because the genotype frequencies observed in the offspring generation differ greatly from those predicted Hardy-Weinberg and thus when the population is not subject to any agents of evolution.

b. The parent and offspring generations must have different allele frequencies because if they were the same, then the genotype frequencies of the offspring generation would be equal to those calculated by Hardy-Weinberg.

Chapter 5

Generalizing from the specific case: Introducing the Hardy-Weinberg Equation[1]

5.1 Formulating the Components of the Hardy-Weinberg Equation

The probability rule we used in another module (p. 16) to predict **genotype** frequencies (Definitions, p. 23) in the offspring generation of a specific population:

> NOTE: If every individual in a population has an equal probability of surviving and producing surviving offspring, then a **genotype** is expected to appear in the offspring generation with a frequency (probability) (Definitions, p. 23) equal to the mathematical product of the frequency with which each **allele** that forms the genotype occurs in the population producing it.

can be use to generate a **general formula** for doing the same thing. That is, we can create a formula that describes how frequently particular **genotypes** will appear in the offspring generation when a population is not subject to an agent of evolution. This formula is known as the Hardy-Weinberg equation.

To do this, we will first generate the individual elements of the Hardy-Weinberg equation.

As the boxed rule above says, offspring **genotype** frequencies are calculated using parental **allele** frequencies. So imagine a population that has only two alleles for a given locus, **A** and **a**, and that in this population

- the **A** allele occurs with a frequency equal to p
- the **a** allele occurs with a frequency equal to q

To make sure you understand these phrases, substitute a number for p or for q. For example, p might be 0.4 meaning that the **A** allele occurs in 40% of the population's loci for this gene.

Notice that, because only two alleles exist in the population for this locus, the frequencies of the **A** and **a** alleles, p and q respectively, must sum to 1, the equivalent of 100%. Or

- $p + q = 1$

Also notice that, if we know the frequency of only one of the two alleles in a population, we can use simple algebra to work out the frequency of the second. For example, if we know p, the frequency of the **A** allele, then

[1] This content is available online at <http://cnx.org/content/m15268/1.1/>.

- q = 1 - p

And, of course, if we know q, the frequency of the **a** allele, then

- p = 1 - q

To confirm your understanding of this relationship between the frequencies with which two alleles occur in a population when only two alleles exist for a given locus, answer the following questions.

Exercise 5.1 *(Solution on p. 24.)*

Please explain in your own words why p + q must always equal 1 when only two alleles exist in a population for a given locus.

Exercise 5.2 *(Solution on p. 24.)*

Let's consider a real example of this. In 2005, Stefasson **et al.** reported the fascinating discovery of an allele in humans whose presence is associated with increased fertility in Icelandic and European populations. Females with at least one copy of the allele have approximately 3.5%, and males 2.9%, more children on average than non-carriers. The exact mechanism by which the allele, known as H2, affects fertility is unknown.

If we know that 21% of European loci for this gene house the H2 allele, then how frequently must the single alternative allele, H1, for that locus occur in this population? Why?

Exercise 5.3 *(Solution on p. 24.)*

A colleague determines that the B1 and B2 alleles of the B locus both occur with a frequency equal 0.45. Surprised, she redoes her work and confirms her results.

a. What could be the cause of your colleague's surprise? Please explain.

b. Because your colleague confirms her results she now needs to explain them. She turns to you for assistance. What do you suggest? Please be sure to explain how your explanation accounts for her observations.

Now that we have designated p to represent the freqeuncy of **A** allele and q, the **a** allele, we are ready to move forward with our efforts to construct the elements of the Hardy-Weinberg equation. Imagine that every individual in a population, in which both copies of both the **A** and **a** allele occur, is equally likely to survive and to reproduce.

What possible genotypes could occur in the offspring of this population?

Exercise 5.4 *(Solution on p. 24.)*

To answer this question, determine all the possible genotypes that could be formed from a population of individuals whose loci collectively warehouse numerous copies of **A** and **a** alleles. Remember that, because these individuals are all equally likely to reproduce, all combinations of these two alleles have the potential to form. Visit this module (p. 15) if you have questions.

Now that we know what genotypes could form, we can use the rule highlighted at the very beginning of this module to predict how frequently each of these **genotypes** will appear in the offspring generation.

Exercise 5.5 *(Solution on p. 24.)*

What are these frequencies? Apply the highlighted (boxed) rule above to complete the phrases below using the symbols p and q.

If all individuals are equally likely to survive and to reproduce, then the

- frequency of genotype **AA** will equal _____
- frequency of genotype **aa** will equal _____
- frequency of genotype **Aa** will equal _____
- frequency of genotype **aA** will equal _____

To test your understanding of these relationships, answer the following questions.

Exercise 5.6 *(Solution on p. 24.)*

Please explain in your own words what these three formulae tell us about the relationship between allele frequencies in the population and genotype frequencies in the offspring generation when all individuals are equally likely to survive and to reproduce.

Exercise 5.7 *(Solution on p. 24.)*

Return to the scenario described in problem 2. How frequently do you expect to the H1H1, H1H2, and H2H2 genotypes to appear in Europeans if the population is not evolving with respect to this allele?

To solve this problem, review the section above and generate a list of the information you need to generate and describe how you plan to get it.

5.2 Formalizing the Hardy-Weinberg Equation

The formulae generated above - p2, 2pq and q2 - constitute the fundamental components of the Hardy-Weinberg equation. Thus, they describe the genotype frequencies you will see in a population, with respect to a single locus with only two alleles, if the population is not subject to any agent of evolution. That is, all individuals in the population are equally likely to survive and to produce offspring that survive.

Because together these formulae account for 100% of the genotypes this population could produce, they can be summarized and are often written in the following way:

- p2 (**AA**) + 2pq (**Aa**) + q2 (**aa**) = 1

In words, this equation says that the values, p2, 2pq and q2, which describe the frequency with which the **AA**, **Aa** and **aa** genotypes occur respectively, sum to 1.

Importantly, and as you applied it in the previous section, the Hardy-Weinberg equation is not necessarily used in the form in which it is written above. That is, you do not set the equation equal to 1 and solve for an unknown. Rather the individual elements p2, 2pq and q2 along with the relationship p + q = 1 are used as needed to solve problems.

Interestingly, the Hardy-Weinberg equation was actually formulated and published independently by both the British mathematician G. H. Hardy and the German physician cum geneticist W. Weinberg in 1908. Because Weinberg published in native German, however, his contribution was not recognized until 1943 at which point the principle was renamed to recognize both contributions.

Definitions

- **frequency** - the number of times an event or observation, for example a particular measurement or condition like blue eyes, is observed in a collection of events or observations like those comprising a sample, population or study. In this statistical sense, a frequency is equivalent to a proportion. For example, the frequency of a particular allele is equal to the number of times that allele is observed in a population over the total number of alleles for that locus in the population. Can be expressed as a fraction, a percentage, a decimal, or a probability.
- **genetic equilibrium** - state of a population in which allele frequencies remain unchanged from one generation to the next.

Works Cited

- Stefansson, H., Helgason, A., Thorleifsson, G. et al. 2005. A common inversion under selection in Europeans. *Nature Genetics*. 37:129-137.

Solutions to Exercises in Chapter 5

Solution to Exercise 5.1 (p. 22)
An example answer: I imagine a locus as slots for alleles. In diploid organisms, each individual will have two 'slots' or two loci for the particular gene of interest, one per chromosome. If only two alleles exist to fill every slot in this population, the slots that are not filled with one of those two alleles must be filled with the second. Consequently, if 50% of the slots are filled with **A** alleles, then the remaining 50% must be filled with **a** alleles to account for 100% of the population's loci. Because 50% is equivalent to a frequency of 0.5, then the frequency of the **A** allele or p equals 0.5 as does the frequency of the **a** allele or q so that p + q = 1.

Solution to Exercise 5.2 (p. 22)
If 21% of the loci (equivalent to a frequency of 0.21) in a population contain the H2 allele and H1 is the only other possible allele for this locus, then 79% of the remaining loci (equivalent to a frequency of 0.79) must have this allele. No other alleles exist for this locus consequently, if H2 does not occur at a locus then H1 must be there instead.

Solution to Exercise 5.3 (p. 22)
Your colleague was probably surprised because she thought that B1 and B2 were the only two alleles that occurred at this locus in this population. Consequently, the discovery that their frequencies, p and q, summed to 0.9 as opposed to 1 was startling. Your suggestion to look for at least one additional allele to account for the 10% of the alleles unaccounted for in her study is well taken. She realizes that the existence of one or more additional alleles would explain the missing 10% and enable her to bring the summed allele frequencies for the B locus to 1.

Solution to Exercise 5.4 (p. 22)
There are four possible genotypes:

- **AA**
- **aa**
- **Aa**
- **aA**

Solution to Exercise 5.5 (p. 22)
Because the **Aa** and **aA** genotypes are genetically equivalent, we can summarize the relationships you articulated above as

- frequency of genotype **AA** will equal p x p = p2
- frequency of genotype aa**aa** will equal q x q = q2
- frequency of genotype **Aa** will equal 2 x (p x q) = 2pq

And there you have it, the three fundamental elements of the Hardy-Weinberg equation that describe how frequently the three possible **genotypes** will appear in the offspring generation of a population that is not subject to an agent of evolution! Remember that only three genotypes are possible because we are only working with a gene for which only two alleles exist in a population.

Solution to Exercise 5.6 (p. 22)
In plain English, these three relationships tell us that

1. If we want to know the frequency of the homozygous genotype (**AA** or **aa**) in the offspring of a population in which all individuals are equally likely to survive and reproduce, then we simply square the frequency with which the appropriate allele (**A** or **a**) occurs population.

2. If we want to know the frequency of the heterozygous genotype (**Aa**) in the offspring of a population in which all individuals are equally likely to survive and reproduce, then we multiple the frequency with which each allele (**A** and **a**) occurs in the population and multiply this result by 2.

Solution to Exercise 5.7 (p. 23)
Check your outline by answering the questions below.

1. How frequently do the H1 and H2 alleles occur in this population? This can be found in solution to problem 2.

2. Calculate the expected frequency of the H1H1, H1H2 and H2H2 genotypes in offspring of this population. To do this, square the frequency with which the H1 allele occurs in the population (p2), multiply the frequency with which the H1 allele occurs with the frequency with which the H2 allele occurs and multiply this result by two (2pq), and finally square the frequency with which the H2 allele occurs in the population (q2).

Chapter 6

What Makes the Hardy-Weinberg Equation So Useful?[1]

The Hardy-Weinberg Equation

- p2 (**AA**) + 2pq (**Aa**) + q2 (**aa**) = 1

is incredibly useful because it describes mathematically the genetic product of a population in which all individuals are equally likely to survive and to produce surviving offspring. Specifically, it calculates the genotype frequencies that will be observed in a population that is not evolving. This information functions as a null hypothesis or standard against which we can judge if a population is evolving.

6.1 Summarizing the Concepts on which Hardy-Weinberg is Based

To understand how Hardy-Weinberg enables this, let's review what we have learned:

- When agents of evolution favor the survival and/or reproduction of some individuals (and their alleles) over others, the population's allele frequencies change over time and the population is said to evolve.
- When all individuals are equally likely to survive and to produce offspring that survive, allele frequencies do not change from one generation to the next because alleles end up in fertilization events, and thus the offspring generation, in proportion to their relative commonness (frequency) in the parental generation.
- When alleles end up in fertilization events (and the offspring generation) in proportion to their relative commonness in the parental generation, **genotype** frequencies of the offspring generation can be determined from the **allele** frequencies of the parental generation using simple rules for calculating the probability of an event composed to two independent events.
- Specifically, these rules say that, if we let p equal the frequency of the **A** allele and q the frequency of the **a** allele in the parental generation, and thus p equals the probability of picking an **A** allele and q an **a** allele from the parental generation, the:
- frequency of the **AA** genotype in the offspring generation is equal to the probability of picking two **A** alleles from the parent generation or p x p = p2
- frequency of the **aa** genotype in the offspring generation is equal to the probability of picking two **a** alleles from the parent generation or q x q = q2
- frequency of the **Aa** genotype in the offspring generation is equal to two times the probability of picking an **A** and **a** allele from the parent generation or 2 x (p x q) = 2pq
- This relationship, which is often written as p2 (**AA**) + 2pq (**Aa**) + q2 (**aa**) = 1, is known as the Hardy-Weinberg equation. It tells us what genotype frequencies (p2, 2pq, q2) we will see in a population that is not evolving based on parental allele frequencies, p and q.

[1] This content is available online at <http://cnx.org/content/m15261/1.1/>.

Finally, please recall that the Hardy-Weinberg equation only applies to a locus for which there are two alleles in a population. And that, although some individuals in the population will die or fail to reproduce, population allele frequencies will be maintained because these losses will remove alleles in proportion to their commonness in the population.

6.2 So what is the power of these concepts?

Notice that these principles apply to the production of a **single** generation. Thus, if an offspring generation is the product of a parental generation that was not subject to an agent of evolution, then the offspring generation must have the following two properties.

1. The allele frequencies must be the same as those of the parental generation.
2. Genotypes must occur with the frequencies predicted by Hardy-Weinberg: **AA** = p2, **aa** = q2 and **Aa**=2pq

Because of this, we can use information from a **single** offspring generation to test the hypothesis that individuals in the parental generation were all equally likely to survive and to produce surviving offspring and, therefore, that the population is not evolving with respect to this locus.

6.3 How does this work?

If parents are all equally likely to survive and to produce surviving offspring, then the allele frequencies of the offspring generation can be used as proxy for the allele frequencies of the parental generation. They should be identical if the no evolution condition is met.

Now we can use these hypothetical parental generation allele frequencies, that are based on the assumption of no evolution, in the Hardy-Weinberg equation to calculate the genotype frequencies we would see in the offspring generation if the population were not evolving with respect to these alleles.

Finally, we can draw a conclusion about whether a population is subject to an agent of evolution by comparing the genotype frequencies we calculated using Hardy-Weinberg to those we actually observe in the offspring generation.

To test your understanding of these relationships, answer the questions below.

Exercise 6.1 ***(Solution on p. 31.)***

What would you conclude about a) the actual similarity of the parental and offspring generation allele frequencies and b) whether the population was evolving if

1. The observed offspring genotype frequencies equaled those predicted by Hardy-Weinberg?
2. The observed offspring genotype frequencies did not equal those predicted by Hardy-Weinberg?

Now, let's apply this logic to an actual problem.

Exercise 6.2 ***(Solution on p. 31.)***

In the mid-1990's, researchers discovered that despite repeated exposure to HIV-1, a strain of the Human immunodeficiency virus (HIV), some individuals remained uninfected (Samson *et al.*, 1996).

Subsequent investigation revealed the existence of an allele that confers immunity to HIV-1 infection in homozygotes. This allele, known as delta-*ccr-5* or CCR5-delta-32, is a mutant version of the cell-surface receptor protein CCR-5. It inhibits HIV infection because it codes for a form of the CCR-5 receptor to which HIV-1 viruses are unable to bind and, thus, to enter white blood cells and thereby establish an infection (Samson *et al.*, 1996).

For reasons unrelated to its effects on susceptibility to HIV-1 infection, this allele is found most commonly in caucasian Europeans and is absent or virtually absent from African, Asian, Middle

Eastern and American Indian populations (Galvani and Slatkin, 2003). Table 1 contains original data from Samson *et al.* (1996) documenting the genotypes of 704 caucasian Europeans. Use these data to answer the questions that follow.

Table 1: The number of individuals homozygous for either the CCR-5 or ccr-5 allele or heterozygous for these two alleles in a sample of 704 caucasian Europeans.

Genotype	**Number of Individuals**
CCR-5/CCR-5	582
CCR-5/*ccr-5*	114
ccr-5/*ccr-5*	8
Total	704

Table 6.1

Is this population of Europeans evolving with respect to this allele? How do you know? To answer this question, work your way through the steps below.

1. Review the question. Using information from earlier in this module or in related modules, complete the chart below on a piece of paper.

What do I need to know to answer this question?	How do I get this information?

Table 6.2

Exercise 6.3 ***(Solution on p. 31.)***

Samson *et al.* (1996) determined that CCR-5 allele occurred with a frequency of 0.91 and the *ccr-5* allele with a frequency 0.09. Do you agree?

Exercise 6.4 ***(Solution on p. 31.)***

Samson *et al.* (1996) also concluded that *'The genotype frequencies observed in this population were not significantly different from the expected Hardy-Weinberg distribution...'* (Samson *et al.*, 1996, p.724.) Do you agree with their conclusion? Why or why not? Please explain using data to support your conclusion.

Exercise 6.5 ***(Solution on p. 32.)***

Finally, we are ready to answer our original question. Is this population of caucasian Europeans evolving with respect to this locus? Yes or no? Please explain.

Exercise 6.6 ***(Solution on p. 32.)***

Are you confident that this result is probably true of the caucasian European population at large? Why or why not? Please explain.

If you are not confident, please describe one activity you could undertake to improve your confidence.

Exercise 6.7 ***(Solution on p. 32.)***

Imagine Europeans were not careful to practice safe sex, use clean needles, and maintain a virus-free blood supply causing exposure to the virus and infection rates to increase dramatically. Predict what would happen, if anything, to the frequency of these two alleles over time? Why? Please explain.

Exercise 6.8 ***(Solution on p. 32.)***

It turns out that European caucasian heterozygotes for the CCR-5 locus have lower infection rates and, when infected, progress more slowly to AIDS, the syndrome that ultimately kills HIV infected individuals, than CCR-5/CCR-5 homozygotes (Galvani and Slatkin, 2003).

Under a scenario in which exposure to HIV-1 and actual infection rates soar in Europe, what do you think would happen, if anything, to the frequency with which these CCR-5/CCR-5, CCR-5/*ccr-5* and *ccr-5*/*ccr-5* genotypes occur in this population? Why? Please explain.

Works Cited

- Galvani, A.P. and M. Slaktin. 2003. Evaluating plagues and small pox as historical selective pressures for the *CRR5 delta-32* HIV-resistance allele. *Proceedings of the National Academy of Sciences.* 100:15276-15279.
- Samson, M., F. Libert, B.J. Doranz, *et al.*. 1996. Resistance to HIV-1 in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature.* 382:722-725.
- Schliekelman, P., C. Garner and M. Slatkin. 2001. Natural selection and resistance to HIV. *Nature.* 411:545

Solutions to Exercises in Chapter 6

Solution to Exercise 6.1 (p. 28)

1. The parental generations allele frequencies must have been identical to those of the offspring, therefore the population is not evolving with respect to these alleles.
2. The offspring generation allele frequencies, which you used as a proxy for the parental allele frequencies, cannot be the same as those of the actual parental generation and therefore the population is evolving with respect to these alleles. This is true because actual offspring genotype frequencies did not equal those predicted by Hardy-Weinberg. The follow-up question would then be: what agent of evolution is responsible for this departure from genetic equilibrium?

Solution to Exercise 6.2 (p. 28)
To answer this question, you need to know

- the frequency with which the CCR-5 and *ccr-5* alleles occur in this population
- the genotype frequencies you would see if this population was not evolving with respect to these alleles
- the actual genotype frequencies observed in this population

To get this information, you need to

- calculate allele frequencies from the genotype frequencies provided in Table 1.
- use these frequencies as a proxy for the allele frequencies of the generation that produced this population and calculate, via the Hardy-Weinberg equation, the genotype frequencies you would see if this population is not evolving with respect to these alleles.
- calculate the actual genotype frequencies from the data provided in Table 1.
- compare the expected versus actual genotype frequencies and draw a conclusion about whether this population is evolving with respect to these alleles.

Complete these steps and move on to the following problems to confirm your answers.
Solution to Exercise 6.3 (p. 29)
To answer this question, we need to determine the total number of alleles in the population and what fraction of that total is due to the CCR-5 and *ccr-5* alleles respectively.

Since each of the 704 individuals in this sample has two alleles for this locus, one per chromosome, we are working with a total of 1408 alleles.

Of these 704 individuals, 582 have two copies of the CCR-5 allele and another 114 have one copy for a total of 1278 (2 x 582 + 114) CCR-5 alleles. Therefore, 1278 of 1408 alleles are CCR-5 alleles, a value equal to a frequency of 0.91 or 91% (1278/1408 x 100).

Now that we have calculated the frequency of the CCR-5 allele, there are two ways to confirm that the *ccr-5* allele occurs with a frequency equal to 0.09. The first is to use the relationship p + q = 1. If we set p equal to 0.91, the frequency of the CCR-5 allele, then q, the frequency of the *ccr-5*, must equal 0.09 because q = 1 - p.

We can also confirm this calculation using the genotype data from Table 1. If 114 individuals have one copy of the *ccr-5* allele and another 8 have two copies, there are 130 copies (2 x 8 +114) of the *ccr-5* allele in this population of 1408 alleles for a frequency of 130/1408 or 0.09, equivalent to 9% of the allele population.
Solution to Exercise 6.4 (p. 29)
To answer this, we must calculate the genotype frequencies we would see if this population is not evolving with respect to these alleles, the actual genotype frequencies observed in this population and finally, compare these two frequencies and draw a conclusion.

If the population is not evolving with respect to these alleles, the population of parents that produced these 704 offspring and the population of offspring themselves must have the same allele frequencies. Therefore, we can use the allele frequencies we calculated above and the Hardy-Weinberg equation to determine the genotype frequencies we would see in this offspring population under the assumption of no evolution.

According to the Hardy-Weinberg equation, the frequency of the

- CCR-5/CCR-5 genotype is equal to the square of the frequency with which this allele occurs or p2 = 0.91 x 0.91 = 0.83.
- *ccr-5*/*ccr-5* genotype is equal to the square of the frequency with which this allele occurs or q2 = 0.09 x 0.09 = 0.008.
- CCR-5/*ccr-5* genotype is equal to the two times the product of frequency with which of these two alleles occur or 2pq = 2(0.91 x 0.09) = 0.16.

Thus, if this population is not evolving, 83%, 0.8% and 16% of these 704 offspring should exhibit the CCR-5/CCR-5, *ccr-5*/*ccr-5*, and CCR-5/*ccr-5* genotypes respectively. We can confirm our mathematics and that we haven't missed any potential offspring genotypes by checking that these three frequencies sum to 1 as they do (0.83 + 0.16 = 0.008 = 1.0).

We now need to calculate the actual frequency with which these genotypes occur in this population of 704 offspring (and therefore 704 genotypes). From Table 1:

- 582 of 704 genotypes are CCR-5/CCR-5 for a frequency of 582/704=0.83.
- 8 of 704 genotypes are *ccr-5*/*ccr-5* for a frequency of 8/704=0.011
- 114 of 704 genotypes are CCR-5/*ccr-5* for a frequency of 114/704=0.16

Thus, 83%, 1.1% and 16% of this population of 704 offspring exhibit the CCR-5/CCR-5, *ccr-5*/*ccr-5*, and CCR-5/*ccr-5* genotypes respectively. We can confirm our mathematics by checking that these three frequencies sum to 1 as they do (0.83 + 0.16 + 0.011 = 1.0).

We are now prepared to compare the genotype frequencies we actually see in this population of 704 individuals to those we would see if it is not evolving with respect to these alleles. Reviewing the summarized data below, we can see that both sets of genotype frequencies are nearly identical confirming Samson *et al.'s* (1996) conclusion.

- 83% of the genotypes are CRR-5/CRR-5 in both data sets
- 16% of the genotypes are CRR-5/*crr-5* in both data sets.
- 0.08 and 1.1% , or ∼0.1%, of the genotypes are *crr-5*/*crr-5* in the observed and expected data sets respectively.

Solution to Exercise 6.5 (p. 29)
The nearly identical expected and observed genotype frequencies calculated in problem 4 suggest that this population of caucasian Europeans is not evolving with respect to this locus.

Solution to Exercise 6.6 (p. 29)
The best answers to this question will discuss the lack of information describing the source of the 704 individuals genotyped for this study. Are they representative of caucasian Europeans at large or is this sample biased in some way? If you are not confident, your confidence might be improved by gathering background information on these individuals to assess their representativeness or by looking for or conducting a complimentary study to see if its results support these reported here.

Solution to Exercise 6.7 (p. 29)
One would expect the frequency of the CCR-5 allele to decline and the frequency of the *ccr-5* allele to increase as individuals carrying the *ccr-5* allele resist infection and therefore leave behind more offspring relative to those who, without it, contract and succumb to the disease and leave behind fewer offspring as a result.

Significantly, something like this appears to be happening in Africa, where different variations in the CCR-5 receptor are conferring increased resistance to HIV infection and, in those infected, increased time to development of AIDS, the condition that ultimately kills HIV infected individuals. Individuals with these alleles are expected to produce 15 to 30% more children than individuals that do not carry these alleles (Schliekelman, Garner and Slatkin, 2001).

Solution to Exercise 6.8 (p. 30)
Because individuals carrying at least one copy of the *ccr-5* allele would be less likely to both contract HIV-1 and, if infected, to progress more slowly to AIDS, these individuals would leave behind more offspring

than CCR-5/CCR-5 homozygotes. Thus, one would expect the frequency of CCR-5/CCR-5 homozygotes to decline and the frequency of *ccr-5*/*ccr-5* homozygotes and heterozygotes to increase over time at least initially.

Index of Keywords and Terms

Keywords are listed by the section with that keyword (page numbers are in parentheses). Keywords do not necessarily appear in the text of the page. They are merely associated with that section. *Ex.* apples, § 1.1 (1) **Terms** are referenced by the page they appear on. *Ex.* apples, 1

Attributions

Collection: *Understanding the Hardy-Weinberg Equation*
Edited by: Laura Martin
URL: http://cnx.org/content/col10472/1.1/
License: http://creativecommons.org/licenses/by/2.0/

Module: "A Population Geneticist's Defintion of Evolution"
By: Laura Martin
URL: http://cnx.org/content/m15262/1.1/
Pages: 1-2
Copyright: Laura Martin
License: http://creativecommons.org/licenses/by/2.0/

Module: "Why are Allele Frequencies Maintained Across Generations When a Population is Not Evolving?"
By: Laura Martin
URL: http://cnx.org/content/m15263/1.2/
Pages: 3-9
Copyright: Laura Martin
License: http://creativecommons.org/licenses/by/2.0/

Module: "Do Parental Generation Genotype Frequencies Affect the Probability that an Allele Will Appear in the Offspring Generation?"
By: Laura Martin
URL: http://cnx.org/content/m15267/1.1/
Pages: 11-13
Copyright: Laura Martin
License: http://creativecommons.org/licenses/by/2.0/

Module: "Are Genotype Fequencies Necessarily Maintained Across Generations When A Population is Not Evolving?"
By: Laura Martin
URL: http://cnx.org/content/m15266/1.1/
Pages: 15-20
Copyright: Laura Martin
License: http://creativecommons.org/licenses/by/2.0/

Module: "Generalizing from the specific case: Introducing the Hardy-Weinberg Equation"
By: Laura Martin
URL: http://cnx.org/content/m15268/1.1/
Pages: 21-25
Copyright: Laura Martin
License: http://creativecommons.org/licenses/by/2.0/

Module: "What Makes the Hardy-Weinberg Equation So Useful?"
By: Laura Martin
URL: http://cnx.org/content/m15261/1.1/
Pages: 27-33
Copyright: Laura Martin
License: http://creativecommons.org/licenses/by/2.0/

Understanding the Hardy-Weinberg Equation
This collection develops the Hardy-Weinberg equation, the standard null model for detecting evolution in populations, from first principles. Some common student misconceptions are approached as questions to be quantitatively tested using simple computer simulations and problems taken from recently published research papers examining evolution in human populations.

About Connexions
Since 1999, Connexions has been pioneering a global system where anyone can create course materials and make them fully accessible and easily reusable free of charge. We are a Web-based authoring, teaching and learning environment open to anyone interested in education, including students, teachers, professors and lifelong learners. We connect ideas and facilitate educational communities.

Connexions's modular, interactive courses are in use worldwide by universities, community colleges, K-12 schools, distance learners, and lifelong learners. Connexions materials are in many languages, including English, Spanish, Chinese, Japanese, Italian, Vietnamese, French, Portuguese, and Thai. Connexions is part of an exciting new information distribution system that allows for **Print on Demand Books**. Connexions has partnered with innovative on-demand publisher QOOP to accelerate the delivery of printed course materials and textbooks into classrooms worldwide at lower prices than traditional academic publishers.