

# **Hardy-Weinberg Equilibrium (HWE)**

## **Exercise part-II**

**Core course: ZOOL 3014**

**B.Sc. (Hons'): VIth Semester**

**Prof. Pranveer Singh**

Recurrent mutation will maintain alleles in the population, even if there is strong selection against them

In population genetics, the **Wahlund effect** refers to reduction of heterozygosity in a population caused by subpopulation structure. Namely, if two or more subpopulations have different allele frequencies then the overall heterozygosity is reduced, even if the subpopulations themselves are in a Hardy-Weinberg equilibrium. The underlying causes of this population subdivision could be geographic barriers to gene flow followed by genetic drift in the subpopulations.

**The simple derivation above can be generalized for more than two alleles and polyploidy.**

**Generalization for more than two alleles**

$$(p + q + r)^2 = p^2 + q^2 + r^2 + 2pq + 2pr + 2qr$$

More generally, consider the alleles  $A_1, \dots, A_n$  given by the allele frequencies  $p_1$  to  $p_n$ ;

$$(p_1 + \dots + p_n)^2$$

giving for all homozygotes:

$$f(A_i A_i) = p_i^2$$

and for all heterozygotes

$$f(A_i A_j) = 2p_i p_j$$

**Generalization for polyploidy**

The Hardy-Weinberg principle may also be generalized to polyploid systems, that is, for organisms that have more than two copies of each chromosome. Consider again only two alleles. The diploid case is the binomial expansion of:

$$(p+q)^2$$

and therefore the polyploid case is the polynomial expansion of

$$(p+q)^c$$

where  $c$  is the ploidy, for example with tetraploid ( $c = 4$ ):

Genotype	Frequency
AAAA	$p^4$

AAAa	$4p^3q$
AAaa	$6p^2q^2$
Aaaa	$4pq^3$
aaaa	$q^4$

Whether the organism is a 'true' tetraploid or an amphidiploid will determine how long it will take for the population to reach Hardy–Weinberg equilibrium.

### **Complete generalization**

For  $n$  distinct alleles in  $c$ -ploids, the genotype frequencies in the Hardy–Weinberg equilibrium are given by individual terms in the [multinomial expansion](#) of  $(p_1 + \dots + p_n)^c$

$$(p_1 + \dots + p_n)^c = \sum_k \binom{c}{k} \dots$$

Where the A gene is sex linked, the heterogametic sex (*e.g.*, mammalian males; avian females) have only one copy of the gene (and are termed hemizygous), while the homogametic sex (*e.g.*, human females) have two copies. The genotype frequencies at equilibrium are  $p$  and  $q$  for the heterogametic sex but  $p^2$ ,  $2pq$  and  $q^2$  for the homogametic sex.

## Solved Examples

1.

You have sampled a population in which you know that the percentage of the homozygous recessive genotype (aa) is 36%. Using that 36%, calculate the following:

- A. The frequency of the "aa" genotype.
- B. The frequency of the "a" allele.
- C. The frequency of the "A" allele.
- D. The frequencies of the genotypes "AA" and "Aa."
- E. The frequencies of the two possible phenotypes if "A" is completely dominant over "a."

You have sampled a population in which you know that the percentage of the homozygous recessive genotype (aa) is 36%. Using that 36%, calculate the following:

- A. The frequency of the "aa" genotype. Answer: 36%, as given in the problem itself.
- B. The frequency of the "a" allele. Answer: The frequency of aa is 36%, which means that  $q^2 = 0.36$ , by definition. If  $q^2 = 0.36$ , then  $q = 0.6$ , again by definition. Since  $q$  equals the frequency of the a allele, then the frequency is 60%.
- C. The frequency of the "A" allele. Answer: Since  $q = 0.6$ , and  $p + q = 1$ , then  $p = 0.4$ ; the frequency of A is by definition equal to  $p$ , so the answer is 40%.
- D. The frequencies of the genotypes "AA" and "Aa." Answer: The frequency of AA is equal to  $p^2$ , and the frequency of Aa is equal to  $2pq$ . So, using the information above, the frequency of AA is 16% (i.e.  $p^2$  is  $0.4 \times 0.4 = 0.16$ ) and Aa is 48% ( $2pq = 2 \times 0.4 \times 0.6 = 0.48$ ).
- E. The frequencies of the two possible phenotypes if "A" is completely dominant over "a." Answers: Because "A" is totally dominate over "a", the dominant phenotype will show if either the homozygous "AA" or heterozygous "Aa" genotypes occur. The recessive phenotype is controlled by the homozygous aa genotype. Therefore, the frequency of the dominant phenotype equals the sum of the frequencies of AA and Aa, and the recessive phenotype is simply the frequency of aa. Therefore, the dominant frequency is 64% and, in the first part of this question above, you have already shown that the recessive frequency is 36%.

2. Sickle-cell anemia is an interesting genetic disease. Normal homozygous individuals (SS) have normal blood cells that are easily infected with the malarial parasite. Thus, many of these individuals become very ill from the parasite and many die. Individuals homozygous for the sickle-cell trait (ss) have red blood cells that readily collapse when deoxygenated. Although malaria cannot grow in these red blood cells, individuals often die because of the genetic defect. However, individuals with the heterozygous condition (Ss) have some sickling of red blood cells, but generally not enough to

**cause mortality. In addition, malaria cannot survive well within these "partially defective" red blood cells. Thus, heterozygotes tend to survive better than either of the homozygous conditions. If 9% of an African population is born with a severe form of sickle-cell anemia (ss), what percentage of the population will be more resistant to malaria because they are heterozygous (Ss) for the sickle-cell gene?**

Sickle-cell anemia is an interesting genetic disease. Normal homozygous individuals (SS) have normal blood cells that are easily infected with the malarial parasite. Thus, many of these individuals become very ill from the parasite and many die. Individuals homozygous for the sickle-cell trait (ss) have red blood cells that readily collapse when deoxygenated. Although malaria cannot grow in these red blood cells, individuals often die because of the genetic defect. However, individuals with the heterozygous condition (Ss) have some sickling of red blood cells, but generally not enough to cause mortality. In addition, malaria cannot survive well within these "partially defective" red blood cells. Thus, heterozygotes tend to survive better than either of the homozygous conditions. If 9% of an African population is born with a severe form of sickle-cell anemia (ss), what percentage of the population will be more resistant to malaria because they are heterozygous (Ss) for the sickle-cell gene? Answer:  $9\% = .09 = ss = q^2$ . To find q, simply take the square root of 0.09 to get 0.3. Since  $p = 1 - 0.3$ , then p must equal 0.7.  $2pq = 2 (0.7 \times 0.3) = 0.42 = 42\%$  of the population are heterozygotes (carriers).

- 3. There are 100 students in a class. Ninety-six did well in the course whereas four blew it totally and received a grade of F. Sorry. In the highly unlikely event that these traits are genetic rather than environmental, if these traits involve dominant and recessive alleles, and if the four (4%) represent the frequency of the homozygous recessive condition, please calculate the following:**
- A. The frequency of the recessive allele.**
  - B. The frequency of the dominant allele.**
  - C. The frequency of heterozygous individuals.**

There are 100 students in a class. Ninety-six did well in the course whereas four blew it totally and received a grade of F. Sorry. In the highly unlikely event that these traits are genetic rather than environmental, if these traits involve dominant and recessive alleles, and if the four (4%) represent the frequency of the homozygous recessive condition, please calculate the following:

- A. The frequency of the recessive allele. Answer: Since we believe that the homozygous recessive for this gene ( $q^2$ ) represents 4% (i.e. = 0.04), the square root (q) is 0.2 (20%).
- B. The frequency of the dominant allele. Answer: Since  $q = 0.2$ , and  $p + q = 1$ , then  $p = 0.8$  (80%).
- C. The frequency of heterozygous individuals. Answer: The frequency of heterozygous individuals is equal to  $2pq$ . In this case,  $2pq$  equals 0.32, which

means that the frequency of individuals heterozygous for this gene is equal to 32% (i.e.  $2(0.8)(0.2) = 0.32$ ).

- 4. Within a population of butterflies, the color brown (B) is dominant over the color white (b). And, 40% of all butterflies are white. Given this simple information, which is something that is very likely to be on an exam, calculate the following:**
- A. The percentage of butterflies in the population that are heterozygous.**
  - B. The frequency of homozygous dominant individuals.**

Within a population of butterflies, the color brown (B) is dominant over the color white (b). And, 40% of all butterflies are white. Given this simple information, which is something that is very likely to be on an exam, calculate the following:

- A. The percentage of butterflies in the population that are heterozygous.
- B. The frequency of homozygous dominant individuals.

Answers: The first thing you'll need to do is obtain p and q. So, since white is recessive (i.e. bb), and 40% of the butterflies are white, then  $bb = q^2 = 0.4$ . To determine q, which is the frequency of the recessive allele in the population, simply take the square root of  $q^2$  which works out to be 0.632 (i.e.  $0.632 \times 0.632 = 0.4$ ). So,  $q = 0.63$ . Since  $p + q = 1$ , then p must be  $1 - 0.63 = 0.37$ . Now then, to answer our questions. First, what is the percentage of butterflies in the population that are heterozygous? Well, that would be  $2pq$  so the answer is  $2(0.37)(0.63) = 0.47$ . Second, what is the frequency of homozygous dominant individuals? That would be  $p^2$  or  $(0.37)^2 = 0.14$ .

- 5. A rather large population of Biology instructors have 396 red-sided individuals and 557 tan-sided individuals. Assume that red is totally recessive. Please calculate the following:**
- A. The allele frequencies of each allele.**
  - B. The expected genotype frequencies.**
  - C. The number of heterozygous individuals that you would predict to be in this population.**
  - D. The expected phenotype frequencies.**
  - E. Conditions happen to be really good this year for breeding and next year there are 1,245 young "potential" Biology instructors. Assuming that all of the Hardy-Weinberg conditions are met, how many of these would you expect to be red-sided and how many tan-sided?**

A rather large population of Biology instructors have 396 red-sided individuals and 557 tan-sided individuals. Assume that red is totally recessive. Please calculate the following:

- A. The allele frequencies of each allele. Answer: Well, before you start, note that the allelic frequencies are  $p$  and  $q$ , and be sure to note that we don't have nice round numbers and the total number of individuals counted is  $396 + 557 = 953$ . So, the recessive individuals are all red ( $q^2$ ) and  $396/953 = 0.416$ . Therefore,  $q$  (the square root of  $q^2$ ) is 0.645. Since  $p + q = 1$ , then  $p$  must equal  $1 - 0.645 = 0.355$ .
- B. The expected genotype frequencies. Answer: Well,  $AA = p^2 = (0.355)^2 = 0.126$ ;  $Aa = 2(p)(q) = 2(0.355)(0.645) = 0.458$ ; and finally  $aa = q^2 = (0.645)^2 = 0.416$  (you already knew this from part A above).
- C. The number of heterozygous individuals that you would predict to be in this population. Answer: That would be  $0.458 \times 953 =$  about 436.
- D. The expected phenotype frequencies. Answer: Well, the "A" phenotype =  $0.126 + 0.458 = 0.584$  and the "a" phenotype =  $0.416$  (you already knew this from part A above).
- E. Conditions happen to be really good this year for breeding and next year there are 1,245 young "potential" Biology instructors. Assuming that all of the Hardy-Weinberg conditions are met, how many of these would you expect to be red-sided and how many tan-sided? Answer: Simply put, The "A" phenotype =  $0.584 \times 1,245 = 727$  tan-sided and the "a" phenotype =  $0.416 \times 1,245 = 518$  red-sided (or  $1,245 - 727 = 518$ ).

**6. A very large population of randomly-mating laboratory mice contains 35% white mice. White coloring is caused by the double recessive genotype, "aa". Calculate allelic and genotypic frequencies for this population.**

A very large population of randomly-mating laboratory mice contains 35% white mice. White coloring is caused by the double recessive genotype, "aa". Calculate allelic and genotypic frequencies for this population. Answer: 35% are white mice, which = 0.35 and represents the frequency of the aa genotype (or  $q^2$ ). The square root of 0.35 is 0.59, which equals  $q$ . Since  $p = 1 - q$  then  $1 - 0.59 = 0.41$ . Now that we know the frequency of each allele, we can calculate the frequency of the remaining genotypes in the population (AA and Aa individuals).  $AA = p^2 = 0.41 \times 0.41 = 0.17$ ;  $Aa = 2pq = 2 (0.59) (0.41) = 0.48$ ; and as before  $aa = q^2 = 0.59 \times 0.59 = 0.35$ . If you add up all these genotype frequencies, they should equal 1.

7. After graduation, you and 19 of your closest friends (lets say 10 males and 10 females) charter a plane to go on a round-the-world tour. Unfortunately, you all crash land (safely) on a deserted island. No one finds you and you start a new population totally isolated from the rest of the world. Two of your friends carry (i.e. are heterozygous for) the recessive cystic fibrosis allele (c). Assuming that the frequency of this allele does not change as the population grows, what will be the incidence of cystic fibrosis on your island?

After graduation, you and 19 of your closest friends (lets say 10 males and 10 females) charter a plane to go on a round-the-world tour. Unfortunately, you all crash land (safely) on a deserted island. No one finds you and you start a new population totally isolated from the rest of the world. Two of your friends carry (i.e. are heterozygous for) the recessive cystic fibrosis allele (c). Assuming that the frequency of this allele does not change as the population grows, what will be the incidence of cystic fibrosis on your island? Answer: There are 40 total alleles in the 20 people of which 2 alleles are for cystic fibrosis. So,  $2/40 = .05$  (5%) of the alleles are for cystic fibrosis. That represents p. Thus, cc or  $p^2 = (.05)^2 = 0.0025$  or 0.25% of the F1 population will be born with cystic fibrosis.

8. You sample 1,000 individuals from a large population for the MN blood group, which can easily be measured since co-dominance is involved (i.e., you can detect the heterozygotes). They are typed accordingly:

BLOOD TYPE	GENOTYPE	NUMBER OF INDIVIDUALS	RESULTING FREQUENCY
M	MM	490	0.49
MN	MN	420	0.42
N	NN	90	0.09

Using the data provide above, calculate the following:

- The frequency of each allele in the population.
- Supposing the matings are random, the frequencies of the matings.
- The probability of each genotype resulting from each potential cross.



Using the data provide above, calculate the following:

- A. The frequency of each allele in the population. Answer: Since  $MM = p^2$ ,  $MN = 2pq$ , and  $NN = q^2$ , then  $p$  (the frequency of the M allele) must be the square root of 0.49, which is 0.7. Since  $q = 1 - p$ , then  $q$  must equal 0.3.
- B. Supposing the matings are random, the frequencies of the matings. Answer: This is a little harder to figure out. Try setting up a "Punnett square" type arrangement using the 3 genotypes and multiplying the numbers in a manner something like this:

	<b>MM (0.49)</b>	<b>MN (0.42)</b>	<b>NN (0.09)</b>
<b>MM (0.49)</b>	<b>0.2401*</b>	<b>0.2058</b>	<b>0.0441</b>
<b>MN (0.42)</b>	<b>0.2058</b>	<b>0.1764*</b>	<b>0.0378</b>
<b>NN (0.09)</b>	<b>0.0441</b>	<b>0.0378</b>	<b>0.0081*</b>

- C. Note that three of the six possible crosses are unique (\*), but that the other three occur twice (i.e. the probabilities of matings occurring between these genotypes is TWICE that of the other three "unique" combinations. Thus, three of the possibilities must be doubled.
- D.  $MM \times MM = 0.49 \times 0.49 = 0.2401$   
 $MM \times MN = 0.49 \times 0.42 = 0.2058 \times 2 = 0.4116$   
 $MM \times NN = 0.49 \times 0.09 = 0.0441 \times 2 = 0.0882$   
 $MN \times MN = 0.42 \times 0.42 = 0.1764$   
 $MN \times NN = 0.42 \times 0.09 = 0.0378 \times 2 = 0.0756$   
 $NN \times NN = 0.09 \times 0.09 = 0.0081$
- E. The probability of each genotype resulting from each potential cross. Answer: You may wish to do a simple Punnett's square monohybrid cross and, if you do, you'll come out with the following result:

$MM \times MM = 1.0 \text{ MM}$   
 $MM \times MN = 0.5 \text{ MM } 0.5 \text{ MN}$   
 $MM \times NN = 1.0 \text{ MN}$   
 $MN \times MN = 0.25 \text{ MM } 0.5 \text{ MN } 0.25 \text{ NN}$   
 $MN \times NN = 0.5 \text{ MN } 0.5 \text{ NN}$   
 $NN \times NN = 1.0 \text{ NN}$

- 9. Cystic fibrosis is a recessive condition that affects about 1 in 2,500 babies in the Caucasian population of the United States. Please calculate the following.**
- A. The frequency of the recessive allele in the population.**
  - B. The frequency of the dominant allele in the population.**
  - C. The percentage of heterozygous individuals (carriers) in the population.**

Cystic fibrosis is a recessive condition that affects about 1 in 2,500 babies in the Caucasian population of the United States. Please calculate the following.

- A. The frequency of the recessive allele in the population. Answer: We know from the above that  $q^2$  is  $1/2,500$  or  $0.0004$ . Therefore,  $q$  is the square root, or  $0.02$ . That is the answer to our first question: the frequency of the cystic fibrosis (recessive) allele in the population is  $0.02$  (or  $2\%$ ).
- B. The frequency of the dominant allele in the population. Answer: The frequency of the dominant (normal) allele in the population ( $p$ ) is simply  $1 - 0.02 = 0.98$  (or  $98\%$ ).
- C. The percentage of heterozygous individuals (carriers) in the population. Answer: Since  $2pq$  equals the frequency of heterozygotes or carriers, then the equation will be as follows:  $2pq = (2)(.98)(.02) = 0.04$  or 1 in 25 are carriers.

- 10. In a given population, only the "A" and "B" alleles are present in the ABO system; there are no individuals with type "O" blood or with O alleles in this particular population. If 200 people have type A blood, 75 have type AB blood, and 25 have type B blood, what are the allelic frequencies of this population (i.e., what are  $p$  and  $q$ )?**

In a given population, only the "A" and "B" alleles are present in the ABO system; there are no individuals with type "O" blood or with O alleles in this particular population. If 200 people have type A blood, 75 have type AB blood, and 25 have type B blood, what are the allelic frequencies of this population (i.e., what are  $p$  and  $q$ )? Answer: To calculate the allele frequencies for A and B, we need to remember that the individuals with type A blood are homozygous AA, individuals with type AB blood are heterozygous AB, and individuals with type B blood are homozygous BB. The frequency of A equals the following:  $2 \times (\text{number of AA}) + (\text{number of AB})$  divided by  $2 \times (\text{total number of individuals})$ . Thus  $2 \times (200) + (75)$  divided by  $2 (200 + 75 + 25)$ . This is  $475/600 = 0.792 = p$ . Since  $q$  is simply  $1 - p$ , then  $q = 1 - 0.792$  or  $0.208$ .

**11. The ability to taste PTC is due to a single dominant allele "T". You sampled 215 individuals in biology, and determined that 150 could detect the bitter taste of PTC and 65 could not. Calculate all of the potential frequencies.**

The ability to taste PTC is due to a single dominant allele "T". You sampled 215 individuals in biology, and determined that 150 could detect the bitter taste of PTC and 65 could not. Calculate all of the potential frequencies. Answer: First, let's go after the recessives (tt) or  $q^2$ . That is easy since  $q^2 = 65/215 = 0.302$ . Taking the square root of  $q^2$ , you get 0.55, which is  $q$ . To get  $p$ , simply subtract  $q$  from 1 so that  $1 - 0.55 = 0.45 = p$ . Now then, you want to find out what TT, Tt, and tt represent. You already know that  $q^2 = 0.302$ , which is tt.  $TT = p^2 = 0.45 \times 0.45 = 0.2025$ . Tt is  $2pq = 2 \times 0.45 \times 0.55 = 0.495$ . To check your own work, add 0.302, 0.2025, and 0.495 and these should equal 1.0 or very close to it. This type of problem may be on the exam.

**12. What allelic frequency will generate twice as many recessive homozygotes as heterozygotes?**

What allelic frequency will generate twice as many recessive homozygotes as heterozygotes? Answer: We need to solve for the following equation:  $q^2 (aa) = 2 \times$  the frequency of Aa. Thus,  $q^2 (aa) = 2(2pq)$ . Or another way of writing it is  $q^2 = 4 \times p \times q$ . We only want  $q$ , so let's trash  $p$ . Since  $p = 1 - q$ , we can substitute  $1 - q$  for  $p$  and, thus,  $q^2 = 4(1 - q)q$ . Then, if we multiply everything on the right by that lone  $q$ , we get  $q^2 = 4q - 4q^2$ . We then divide both sides through by  $q$  and get  $q = 4 - 4q$ . Subtracting 4 from both sides, and then  $q$  (i.e.  $-4q$  minus  $q = -5q$ ) also from both sides, we get  $-4 = -5q$ . We then divide through by  $-5$  to get  $-4/-5 = q$ , or otherwords the answer which is  $0.8 = q$ . I cannot imagine you getting this type of problem in this general biology course although if you take algebra good luck.

1. About 70 percent of all white North Americans can taste the chemical phenylthiocarbamide, and the remainder cannot. The ability to taste is determined by the dominant allele T, and the inability to taste is determined by the recessive allele t. If the population is assumed to be in Hardy-Weinberg equilibrium, what are the genotypic and allelic frequencies in this population?

Because 70 percent are tasters ( $T/T$ ), 30 percent must be nontasters ( $t/t$ ). This homozygous recessive frequency is equal to  $q^2$ ; so, to obtain  $q$ , we simply take the square root of 0.30:

$$q = \sqrt{0.30} = 0.55$$

Because  $p + q = 1$ , we can write

$$p = 1 - q = 1 - 0.55 = 0.45$$

Now we can calculate

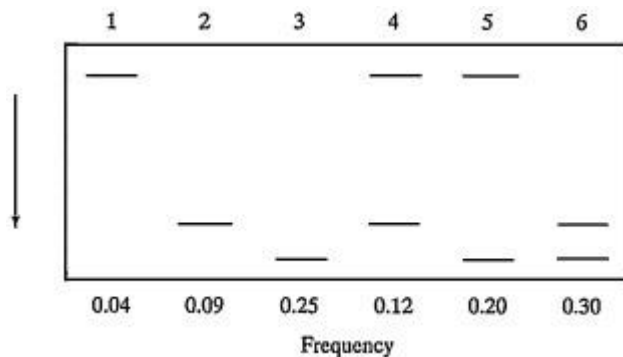
$$p^2 = 0.45^2 = 0.20(T/T)$$

$$2pq = 0 \times 0.45 \times 0.55 = 0.50(T/t)$$

$$q^2 = 0.3(t/t)$$

2. In a large natural population of *Mimulus guttatus*, one leaf was sampled from each of a large number of plants. The leaves were crushed and subjected to gel electrophoresis. The gel was then stained for a specific enzyme X. Six different banding patterns were observed, as shown in the diagram.

- Assuming that these patterns are produced by a single locus, propose a genetic explanation for the six types.
- How can you test your idea?
- What are the allelic frequencies in this population?
- Is the population in Hardy-Weinberg equilibrium?



- Inspection of the gel reveals that there are only three band positions: we will call them slow, intermediate, and fast. Furthermore, any individual can show either one band or two. The simplest explanation is that there are three alleles of one locus (let's call them  $A^S$ ,  $A^I$ , and  $A^F$ ) and that the individuals with two bands are heterozygotes. Hence, 1 =  $S/S$ , 2 =  $I/I$ , 3 =  $F/F$ , 4 =  $S/I$ , 5 =  $S/F$ , and 6 =  $I/F$ .
- The hypothesis can be tested by making controlled crosses. For example, from a self of type 5, we can predict  $1/4 S/S$ ,  $1/2 S/F$ , and  $1/4 F/F$ .
- The frequencies can be calculated by a simple extension of the two-allele formulas. Hence:

$$f(S) = 0.04 + \frac{1}{2}(0.12) + \frac{1}{2}(0.20) = 0.20$$

$$f(I) = 0.09 + \frac{1}{2}(0.12) + \frac{1}{2}(0.30) = 0.30$$

$$f(F) = 0.25 + \frac{1}{2}(0.20) + \frac{1}{2}(0.30) = 0.50$$

d.

The Hardy-Weinberg genotypic frequencies are:

$$\begin{aligned}(p + q + r)^2 &= p^2 + q^2 + r^2 \\ &\quad + 2pq + 2pr + 2qr \\ &= 0.04 + 0.09 + 0.25 + 0.12 \\ &\quad + 0.20 + 0.30\end{aligned}$$

which are precisely the observed frequencies. So it appears that the population is in equilibrium.

3. In a large experimental *Drosophila* population, the fitness of a recessive phenotype is calculated to be 0.90, and the mutation rate to the recessive allele is  $5 \times 10^{-5}$ . If the population is allowed to come to equilibrium, what allelic frequencies can be predicted?

Here mutation and selection are working in opposite directions, so an equilibrium is predicted. Such an equilibrium is described by the formula

$$\hat{q} = \sqrt{\frac{\mu}{s}}$$

In the present question,  $\mu = 5 \times 10^{-5}$  and  $s = 1 - W = 1 - 0.9 = 0.1$ . Hence

$$\hat{q} = \sqrt{\frac{5 \times 10^{-5}}{10^{-1}}} = 2.2 \times 10^{-2} = 0.022$$

$$\hat{p} = 1 - 0.022 = 0.978$$