Korean J. Math. **32** (2024), No. 2, pp. 213–218
https://dx.doi.org/10.11568/kjm.2024.32.2.213

ON THE ADAPTED PARTIAL DIFFERENTIAL EQUATION FOR GENERAL DIPLOID MODEL OF SELECTION AT A SINGLE LOCUS

Won Choi

Abstract. Assume that at a certain locus there are three genotypes and that for every one progeny produced by an $I^A I^A$ homozygote, the heterozygote $I^A I^B$ produces. W. Choi found the adapted partial differential equations for the density and operator of the frequency for one gene and applied this adapted partial differential equations to several diploid model. Also, he found adapted partial differential equations for the diploid model against recessive homozygotes and in case that the alley frequency occurs after one generation of selection when there is no dominance. (see. [1, 2]).
In this paper, we find the adapted partial equations for the model of selection against heterozygotes and in case that the allele frequency changes after one generation of selection when there is overdominance. Finally, we shall find the partial differential equation of general type of selection at diploid model and it also shall apply to actual examples. This is a very meaningful result in that it can be applied **in any model**.

1. Introduction

Assume that at a certain locus there are three genotypes and that for every one progeny produced by an $I^A I^A$ homozygote, the heterozygote $I^A I^B$ produces. To calculate a genotypic frequency, we add up the number of individuals possessing a genotype and divide by the total number of individuals in the sample N. The gene pool of a population can be represented in terms of allelic frequencies. Allelic frequencies can be calculated from the numbers or the frequencies of the genotypes. To calculate the allelic frequency from the numbers of genotypes, we count the number of copies of a particular allele present among the genotypes and divide by the total number of all alleles in the sample. ([5])

Suppose that N_{AA}, N_{AB} and N_{BB} are the numbers of $I^A I^A$, $I^A I^B$ and $I^B I^B$ respectively and N is the total number of populations. For a locus with only two alleles, the frequencies of the alleles are usually represented by p and q, respectively. To calculate the number of copies of the allele in the numerator of the equation, we add twice the number of homozygotes ($2N_{AA}$ or $2N_{BB}$) to the number of heterozygotes

---

Received September 24, 2023. Revised April 17, 2024. Accepted April 26, 2024.
2010 Mathematics Subject Classification: 92D10, 92B05, 60G07.
Key words and phrases: adapted equation, alleles frequency, mean change, selection coefficient.
This research was supported by Incheon National University Research Grant, 2023-2024.

N_{AB} and divide by $2N$. Since the sum of the allelic frequencies always equal 1, we have p and q.

Also, the alleles frequencies can be calculated from the genotype frequencies. To calculate allelic frequencies from genotype frequencies, we add the frequency of the homozygote for each allele to half the frequency of the heterozygote([5]);

$$p = f(I^A) = f(I^A I^A) + \frac{1}{2} f(I^A I^B)$$
$$q = f(I^B) = f(I^B I^B) + \frac{1}{2} f(I^A I^B)$$

W. Choi found the adapted partial differential equations for the density and operator of the frequency for one gene and applied this adapted equations to several diploid model.([2]). Also, he found adapted partial differential equations for the model against recessive homozygotes and in case that the alley frequency occurs after one generation of selection when there is no dominance. ([1])

In this paper, we find adapted partial equations for the model of selection against heterozygotes and in case that the alley frequency changes after one generation of selection when there is overdominance. Finally, we shall find the partial differential equation of general type of selection at diploid model and it also shall apply to actual examples. This is a very meaningful result in that it can be applied **in any model**.

2. The Main Results

We begin with the probability about the frequency of allele;

LEMMA 1. *Denote $x(t, p)$ be the probability that I^A become fixed by time t-th generation, given that its initial frequency is p. Then we have adapted equation*

$$\frac{\partial x}{\partial t} = \frac{pq}{4N} \frac{\partial^2 x}{\partial p^2} + M(t) \frac{\partial x}{\partial p}$$

where $M(t)$ is mean change respectively for frequence of alleles.

Proof. This result follows easily from the property of dyploid model and the result of M. Kimura.([3]) □

We denote the frequency of the $I^A I^A$ genotype, $I^A I^B$ genotype and $I^B I^B$ genotype in diploid model by τ_1, τ_2 and τ_3, respectively.

When selection is for one genotype, selection is automatically against at least one other genotype. The selection coefficien is defined by the relative intensity of selection against a genotype([4]).

The related variable for selection coefficient is the fitness. The fitness means the relative reproductive success of a genotype in case of natural selection. The natural selection occurs when individuals with adaptive traits produce a greater numbers of offspring than that produced by others in the population. If the adaptive traits have a genetic basis, they are inherited by the offspring and appear with greater frequency in the next generation ([4], [5]).

In diplod model, suppose that p is the frequency of allele I^A and q is the frequency of allele I^B. Assume that the allelic frequencies do not change and the genotypic frequencies will not change after one generation in the proportion p^2 (the frequency of I^AI^A), $2pq$ (the frequency of I^AI^B) and q^2 (the frequency of I^BI^B). The population is said to be in Hardy-Weinburg equilibrium when genotype are in the expected proportions of $p^2,\ 2pq,\ q^2$,([4], [5]).

Firstly, we consider situations in which the heterozygotes have upper selection coefficient than either homozygote. In this case, we can suppose that the two types of homozygotes have equal fitness 1.

Then we have;

THEOREM 2. *Assume that the allele frequency changes after one generation of selection against heterozygotes with selection coefficient* α. *The adapted equation*

$$\frac{\partial x}{\partial t} = \frac{pq}{4N}\frac{\partial^2 x}{\partial p^2} + \frac{\alpha pq(2p-1)}{1-2\alpha pq}\frac{\partial x}{\partial p}$$

for frequence of alleles I^A *and*

$$\frac{\partial x}{\partial t} = \frac{pq}{4N}\frac{\partial^2 x}{\partial p^2} + \frac{\alpha pq(2q-1)}{1-2\alpha pq}\frac{\partial x}{\partial p}$$

for frequence of alleles I^B.

Proof. Since the fitness for genotype I^AI^A, I^AI^B and I^BI^B are $1,\ 1-\alpha,\ 1$, respectively, we have the average fitness

$$p^2 + 2pq(1-\alpha) + q^2 = 1 - 2\alpha pq.$$

The frequency of the I^A allele after one generation of selection is calculated by $\tau_1 + \frac{1}{2}\tau_2$. Therefore The frequency of the I^A allele is

$$\frac{p^2 + pq(1-\alpha)}{1-2\alpha pq} = \frac{p(1-\alpha q)}{1-2\alpha pq}.$$

Since the mean change for allele I^A is

$$\frac{p(1-\alpha q)}{1-2\alpha pq} - p = \frac{\alpha pq(2p-1)}{1-2\alpha pq},$$

we have the first result from Lemma 1.

On the other hand, the frequency of the I^B allele after one generation of selection is calculated by $\frac{1}{2}\tau_2 + \tau_3$. Therefore the frequency of the I^B allele

$$\frac{pq(1-\alpha) + q^2}{1-2\alpha pq} = \frac{q(1-\alpha p)}{1-2\alpha pq}.$$

Since the mean change for allele I^B is

$$\frac{q(1-\alpha p)}{1-2\alpha pq} - q = \frac{\alpha pq(2q-1)}{1-2\alpha pq},$$

we have the second result from Lemma 1. □

Selection in favor of the heterozygotes over both homozygotes is known as overdominance. With overdominance, both alleles are favored in the heterozygotes and neither allele is eliminated from the population.

Then we have ;

THEOREM 3. Assume that the allele frequency changes after one generation of selection with the selection coefficient α of heterozygote when there is overdominance. The adapted equations

$$\frac{\partial x}{\partial t} = \frac{pq}{4N}\frac{\partial^2 x}{\partial p^2} + \frac{pq(tq-\alpha p)}{1-\alpha p^2 - tq^2}\frac{\partial x}{\partial p}$$

for frequence of alleles I^A and

$$\frac{\partial x}{\partial t} = \frac{pq}{4N}\frac{\partial^2 x}{\partial p^2} + \frac{pq(\alpha p - tq)}{1-\alpha p^2 - tq^2}\frac{\partial x}{\partial p}$$

for frequence of alleles I^B.

Proof. Since the fitness for each genotype I^AI^A, I^AI^B and I^BI^B are $1-\alpha,\ 1,\ 1-t$ respectively, we have the average fitness

$$p^2(1-\alpha) + 2pq + q^2(1-t) = 1 - \alpha p^2 - tq^2.$$

The frequency of the I^A allele after one generation of selection is calculated by $\tau_1 + \frac{1}{2}\tau_2$. Therefore the frequency of the I^A allele is

$$\frac{p^2(1-\alpha)+pq}{1-\alpha p^2 - tq^2} = \frac{p(1-\alpha p)}{1-\alpha p^2 - tq^2}.$$

Since the mean change for allele I^A is

$$\frac{p(1-\alpha p)}{1-\alpha p^2 - tq^2} - p = \frac{\alpha p^2(p-1)+tpq^2}{1-\alpha p^2 - tq^2} = \frac{pq(tq-\alpha p)}{1-\alpha p^2 - tq^2},$$

we have the first result from Lemma 1.

Similarly, the frequency of the I^B allele after one generation of selection is calculated by $\frac{1}{2}\tau_2 + \tau_3$. Since the mean change for allele I^B is

$$\frac{pq+q^2(1-t)}{1-\alpha p^2 - tq^2} - q = \frac{tq^2(q-1)+\alpha p^2 q}{1-\alpha p^2 - tq^2} = \frac{pq(\alpha p - tq)}{1-\alpha p^2 - tq^2},$$

and we have the second result from Lemma 1. □

An example of overdominance is stick-cell anemia in humans, a disease that results from a mutation in one of the gene that encodes hemoglobin.([4])

So far, Choi have showed the adapted partial differential equation of various model of selection. They are all special cases of more general model of selection at single locus. We shall find the partial differential equation of general model at diploid model.

We conclude with general type of selection;

THEOREM 4. *Assume that the selection coefficients for each genotype I^AI^A, I^AI^B and I^BI^B are $\alpha_1,\ \alpha_2,\ \alpha_3$ respectively. The adapted equations*

$$\frac{\partial x}{\partial t} = \frac{pq}{4N}\frac{\partial^2 x}{\partial p^2} + pq\frac{p(\alpha_1-\alpha_2)+q(\alpha_2-\alpha_3)}{\hat{\alpha}}\frac{\partial x}{\partial p}$$

for frequence of alleles I^A and

$$\frac{\partial x}{\partial t} = \frac{pq}{4N}\frac{\partial^2 x}{\partial p^2} + pq\frac{p(\alpha_2-\alpha_1)+q(\alpha_3-\alpha_2)}{\hat{\alpha}}\frac{\partial x}{\partial p}$$

for frequence of alleles I^B. Here $\hat{\alpha} = p^2(1-\alpha_1) + 2pq(1-\alpha_2) + q^2(1-\alpha_3)$.

Proof. Since the selection coefficients for each genotype I^AI^A, I^AI^B and I^BI^B are $\alpha_1,\ \alpha_2,\ \alpha_3$ respectively, the average fitness is

$$\hat{\alpha} = p^2(1-\alpha_1) + 2pq(1-\alpha_2) + q^2(1-\alpha_3).$$

The frequency of the I^A allele after one generation of selection is

$$\frac{p^2(1-\alpha_1) + pq(1-\alpha_2)}{\hat{\alpha}}.$$

Since the mean change for allele I^A is

$$\begin{aligned}\frac{p^2(1-\alpha_1) + pq(1-\alpha_2)}{\hat{\alpha}} - p &= \frac{pq(1-\alpha_2)(q-p) + p^2q(1-\alpha_1) - pq^2(1-\alpha_3)}{\hat{\alpha}} \\ &= pq\frac{p(\alpha_1-\alpha_2) + q(\alpha_2-\alpha_3)}{\hat{\alpha}},\end{aligned}$$

we have the first result from Lemma 1
Similarly, the frequency of the I^B allele after one generation of selection is calculated by adding the frequency of the I^BI^B genotype and half of the frequency of the I^AI^B genotype. Therefore, the mean change for allele I^B is

$$\begin{aligned}\frac{pq(1-\alpha_2) + q^2(1-\alpha_3)}{\hat{\alpha}} - q &= \frac{pq(1-\alpha_2)(p-q) + pq^2(1-\alpha_3) - p^2q(1-\alpha_1)}{\hat{\alpha}} \\ &= pq\frac{p(\alpha_2-\alpha_1) + q(\alpha_3-\alpha_2)}{\hat{\alpha}}\end{aligned}$$

and we have the second result from Lemma 1. □

Remark 1. The results of selection depend on the fitnesses or selection coefficients of the genotypes in a population as we've already seen in the Theorem 2 and 3. For example, in case that dominant allele I^A confers a fitness advantage, the fitnesses of genotypes I^AI^A and I^AI^B are equal and higher than the fitness of I^AI^B. Also, in case that the heterozygotehas lower fitness than either homozygote, we called this type of selection by underdominance. Underdominance means that allelic frequencies do not change as long as they are at equilibrium. Theorem 4 is a very meaningful result in that it can be applied **in any model**.

The next example is underdominance type of selection in which the heterozygotes has upper selection coefficient than either homozygotes.

Example. Suppose $N_{AA} = 135$, $N_{AB} = 44$ and $N_{BB} = 11$ which means the number of zygotes in one generation. If the frequency of the I^A allele equals p and the frequency of the I^B allele equals q, the frequency of the I^A alleke is

$$p = f(I^A) = \frac{2N_{AA} + N_{AB}}{2N} = 0.826.$$

Also, the frequency of the I^B allele is

$$q = f(I^B) = \frac{2N_{BB} + N_{AB}}{2N} = 1 - p = 0.174$$

Suppose that the average number of viable offspring produced by three genotype I^AI^A, I^AI^B and I^BI^B are $10, 2$ and 5, respectively. This means the average number of progeny per individual in next generation. We find the fitness for each genotype as following;

$$\text{the fitness of } I^A I^A = \frac{10}{10} = 1$$
$$\text{the fitness of } I^A I^B = \frac{20}{10} = 0.2$$
$$\text{the fitness of } I^B I^B = \frac{5}{10} = 0.5$$

and selection coefficients are $\alpha_1 = 0,\ \alpha_2 = 0.8, \alpha_3 = 0.5$. Obviously, the heterozygotes has upper selection coefficient than either homozygotes and we know the type of selection is underdominance. Also $\hat{\alpha} = 0.0726276$ and therefore from Theorem 4, we have adapted equations

$$\frac{\partial x}{\partial t} = 0.000189\frac{\partial^2 x}{\partial p^2} - 1.2043769\frac{\partial x}{\partial p}$$

for frequence of alleles I^A and

$$\frac{\partial x}{\partial t} = 0.000189\frac{\partial^2 x}{\partial p^2} + 1.204369\frac{\partial x}{\partial p}$$

for frequence of alleles I^B. Here the coefficients were rounded to six decimal places.

REMARK 2. In case of a locus with three alleles A_1, A_2 and A_3(triploid model), we have six genotypes and add the frequency of the homozygote to half the frequency of each heterozygous genotype that possesses the allels. To calculate allele frequencies, we add twice the number of homozygotes to the number of heterozygotes that possess the allele and divide $2N$. We can use the same principles for diploid model to determine the mean change for three alleles A_1, A_2 and A_3. But in order to derive the adapted partial differential equation for three alleles, we must first create a basic equation for triploid model, such as Lemma 1 for diploid model. In triploid case with three alleles, as a follow-up to this paper, we will try to proceed in the next study.

References

[1] W.Choi, *On the adapted equations in various dyploid model and Hardy-Weinburg equilibrium in a triploid model* , Korean J. Mathematics **31** (1) (2023).
`https://doi.org/10.11568/kjm.2023.31.1.17`

[2] W.Choi,*On the adapted equations for several dyploid model in population genetics*, Korean J. Mathematics, **30** (1) (2022).
`https://doi.org/10.11568/kjm.2022.30.1.67`

[3] M.Kimura, *A Stochastic model of Compensatory Neutral Evolution* , Proceedings of a Workshop held in Nagoya, July 8-12. Stochastic Methods in Biology (1985).

[4] R. Lewis, *Human Genetics : Concepts and Applications*, McGraw-Hill Education (2016).

[5] B.A. Pierce, *Genetics Essentials : Concepts and Connections*, W.H.Freeman and Company (2014), 216-240.

Won Choi
Department of Mathematics, Incheon National University, Incheon, Korea.
E-mail: `choiwon@inu.ac.kr`