More Precise Calculations of the Cost of Substitution

Walter ReMine*

Abstract

This paper extends the applicability and accuracy of the cost of substitution beyond its traditional range, and demonstrates a useful calculation method. Using my previous clarification of the fundamental cost concept, this paper derives a method for computing the cost of substitution under wide genetic circumstances, including haploids; and diploids with varying degrees of dominance, inbreeding, and with a sex-linked locus. Unlike the traditional approaches, this method is accurate even under fluctuations in parameter values (such as population size, selection coefficient, dominance, and inbreeding coefficient). To display general-purpose results, the parameters are then held constant, and the total cost of substitution is graphed. This includes cases where the selection coefficient is not small and where the traditional equations become highly inaccurate. It is shown that neither environmental change nor soft selection reduces cost problems, at least in single substitutions.

Note added in publication: This paper offers previously unpublished clarifications, derivations and graphs, and refutes widely accepted solutions to a central problem in evolutionary genetics known as Haldane's Dilemma. It was submitted to the journal *Theoretical Population Biology*, where all the peer-reviewers found no errors. Nonetheless, they rejected it from publication on the grounds that it is not a "sufficient advance," and "there is little interest in this subject today among population biologists; it is one of those subjects which has sunk almost beyond trace." This has all been very unfortunate, as there continues to be widespread misunderstanding within the scientific community regarding these important matters, even among those who have studied the cost literature for years. It is hoped that the clarifications presented in this paper will eventually reach the greater scientific community. ***Walter J. ReMine***

Introduction

J.B.S. Haldane (1957) introduced the cost of substitution concept, though its usage became hampered by various confusions, and it eventually fell into common disuse. It was criticized for requiring constant population size, and constant selective values, and for its reliance on "genetic death" and "genetic load," whose physical interpretation is dubious. Such issues severely limited its deployment. Those difficulties were addressed and eliminated in ReMine (2005), which clarified cost theory and generalized the cost of substitution to have a concrete physical interpretation, without reliance on genetic death or genetic load, and while allowing fluctuations in any parameters. This paper applies that conceptual clarification to obtain more precise mathematical results.

---

* Walter ReMine, St. Paul Science, P.O. Box 28006, Saint Paul, MN 55128, Email: ReMine@SaintPaulScience.com

Accepted for publication: July 20, 2006

For any given evolutionary scenario, cost theory calculates the required reproduction rate (referred to as the "cost of evolution") and compares it with the species actual reproduction rate (referred to as the "payment"). If the species cannot "pay the cost," then the scenario is not plausible. That concept is general, and can apply to any model of any evolutionary scenario, because they all require some level of reproduction rate. For ease of comprehension, calculation, and discussion, cost theory partitions the cost of evolution into a sum of various costs, with each cost named according to its specific role. Thus, the cost of substitution is one of many costs that each add extra reproduction rate to the amount required by the scenario (ReMine, 2005).

Evolutionary theory requires that some traits originate as rare beneficial mutations and then, through reproductive means, these increase in number of copies. This increase requires extra reproduction rate. Under the clarified definition, the cost of substitution (C_S) is the extra reproduction rate required to increase a trait (or traits) at the rate given by an evolutionary scenario. This paper uses that clarified cost concept to study single substitutions (non-overlapping in time), under genetic circumstances of broad interest (the same cases studied by Haldane, 1957), and derives equations that are more general, more precise, and well-grounded in concrete physical principles.

Let Q be a specific genotype. At the start of generation i, let P be the "effective starting count"—the effective number of individuals who produce genotype-Q progeny. As the cycle of that generation comes to a close, let P′ be the "ending count" of genotype-Q individuals due solely to the reproduction of the former group. (Throughout this paper, a primed quantity, such as P′, denotes a quantity as the generation comes to a close.) The increase is $\Delta P = P' - P$. Then the cost of substitution for genotype-Q, in generation i, is:

$$\text{Cost}_i = \frac{\Delta P}{P} = \frac{P'}{P} - 1, \text{ if} > 0; \quad \text{otherwise, Cost}_i = 0 \qquad (1)$$

Next this equation is restated in terms more traditional to population genetics. For generation i, define the following variables. Let N_e be the effective breeding population size. Let the **population growth factor** be $G = N_e'/N_e$, (which is 1.0 when the population size remains constant). For calculating the cost of a given genotype, let g be its **effective starting frequency** at the start of generation i, and let g′ be its **ending frequency** at the end of generation i. These are given by: $g = P/N_e$, and $g' = P'/N_e'$. Restating equation 1:

$$\text{Cost}_i = G\,\frac{g'}{g} - 1\,, \quad \text{if} > 0; \quad \text{otherwise, Cost}_i = 0 \qquad (2)$$

The **total cost of substitution** (the "total cost") merely sums Cost_i over all the generations of the substitution.

$$\text{Total cost of substitution} = \sum \text{Cost}_i \qquad (3)$$

The total cost of substitution is later divided by the average number of generations per substitution, to obtain the average cost per generation (C_S)—which is a focus of concern in cost arguments.

Next the equations for the cost of substitution are derived under various genetic circumstances, together with a method for iteratively calculating the total cost of substitution. In each case, the method is accurate even when the selection coefficient is large and when the parameters are arbitrarily changed each generation (unlike the traditional concept, which is too narrowly defined).

In all the following cases, A is the substituting allele, and *a* is the old-type allele, with frequencies p and q, respectively. The substitution's **starting frequency** is p_0; this is the frequency at the beginning of the entire substitution. (Note: p_0 is different from the "effective starting frequency" of a given generation i.) Genotype growth is specified by selection coefficients, where $0 < s \leq 1$. All cases assume random mating (except for the inbreeding fraction f). These are the same genetic circumstances studied in Haldane (1957).

To calculate the cost of a given genotype in a given generation, I am using its effective starting frequency and ending frequency. Throughout this paper these quantities are enclosed in boxes for easy identification.

Haploids, clonal, or self-fertilizing organisms, or for maternally inherited cytoplasmic characters

$$\begin{array}{lcccl}
\text{Genotype:} & A & & a & \\
\text{Selective value:} & 1 & & 1\text{-}s & \\
\text{Frequency:} & \boxed{p} & + & q & = 1 \\
\text{Apply Selection:} & p & + & (1\text{-}s)q & = 1 - sq \\
\text{Normalize:} & \boxed{\dfrac{p}{1\text{-}sq}} & + & \dfrac{(1\text{-}s)q}{1\text{-}sq} & = 1
\end{array} \qquad (4)$$

That gives the new allele frequencies:

$$p' = \frac{p}{1-sq} \quad ; \quad q' = 1 - p' \tag{5}$$

By equation 2:

$$\text{Cost}_i = \frac{G}{(1-sq)} - 1 \text{ , if} > 0; \quad \text{otherwise, Cost}_i = 0 \tag{6}$$

Equations 3, 5, and 6 allow iterative calculation of the total cost of substitution, while allowing changes in any parameters at the start of each generation. With parameters held constant, results are shown in Graph 1.

Diploids

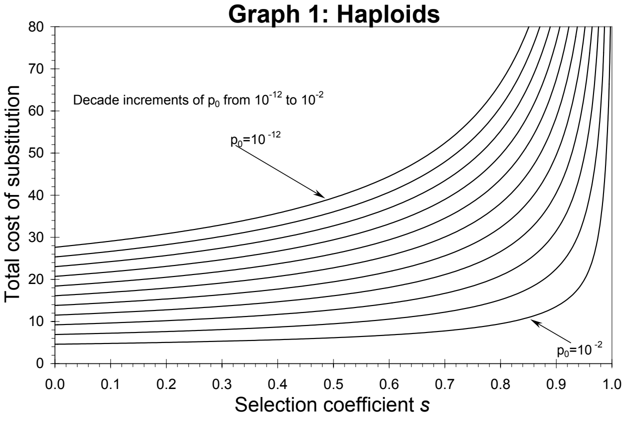
In diploids, Mendelian segregation (in combination with some mating scheme, such as random mating or inbreeding) can affect the genotype frequencies, while tending to leave allele frequencies unaffected. Such change is due solely to the passive remixing of alleles at the gene level, and does not require extra reproduction rate from whole-bodies; therefore we do not tally it into the reproduction rate that whole-bodies are required to produce. So, we disallow it from our cost computation, that is, we calculate the cost after the effects of the mating scheme and Mendelian segregation have been allotted.

The adult population has a reproductive capacity, which is redistributed toward progeny of each given genotype at a predictable rate. That distribution is given by the "effective starting frequencies" of the genotypes, as follows:

$$g_{AA} + g_{Aa} + g_{aa} = 1 \tag{7}$$

That is, a fraction (g_{AA}) of the population's reproductive capacity goes toward producing the AA genotype. Another fraction (g_{Aa}) goes towards producing the *Aa* genotype. And so forth.

As the cycle of one generation completes, let the genotypes' actual "ending frequencies" be labeled as follows:



$$g_{AA}' + g_{Aa}' + g_{aa}' = 1 \tag{8}$$

Each genotype has its own cost, given by equation 2 as:

$$\text{Cost\_}AA_i = G\frac{g_{AA}'}{g_{AA}} - 1 \text{ ,} \quad \text{if} > 0;$$
$$\text{otherwise Cost\_}AA_i = 0 \tag{9}$$

$$\text{Cost\_}Aa_i = G\frac{g_{Aa}'}{g_{Aa}} - 1 \text{ ,} \quad \text{if} > 0;$$
$$\text{otherwise Cost\_}Aa_i = 0 \tag{10}$$

$$\text{Cost\_}aa_i = G\frac{g_{aa}'}{g_{aa}} - 1 \text{ ,} \quad \text{if} > 0;$$
$$\text{otherwise Cost\_}aa_i = 0 \tag{11}$$

I am merely applying equation 2 to each genotype after discounting the passive redistribution effects of mating and Mendelian segregation. In this way, I calculate only the requirements on the reproduction rate of individuals (whole-bodies)—the cost. The producers of a given genotype are required to pay the cost of that genotype. It is usually sufficient to focus solely on the greatest cost, as this almost always forms the most stringent test of the scenario, thus:

$$\text{Cost}_i = \text{Maximum}(\text{ Cost\_}AA_i, \text{Cost\_}Aa_i, \text{Cost\_}aa_i) \tag{12}$$

For a well-behaved substitution, the Cost_*AA* $_i$ always dominates. Therefore,

$$\text{Total cost of substitution} = \sum_{i=1}^{N} \text{Cost\_}AA_i \tag{13}$$

The following case studies demonstrate this method.

An Autosomal Locus in a Diploid

The diploid cases allow dominant and recessive substitutions, specified by recessivity h, where $0 \le h \le 1$. When h=0 this represents a fully dominant substitution.

$$\begin{array}{lllllll} \text{Genotype:} & AA & & Aa & & aa & \\ \text{Selective value:} & 1 & & 1\text{-}hs & & 1\text{-}s & \\ \text{Frequency:} & \boxed{p^2} & + & 2pq & + & q^2 & = 1 \\ \text{Selection:} & p^2 & + & (1\text{-}hs)2pq & + & (1\text{-}s)q^2 & = 1\text{-}2hspq\text{-}sq^2 \end{array} \tag{14}$$

Normalize:

$$\boxed{\frac{p^2}{1\text{-}2hspq\text{-}sq^2}} + \frac{(1\text{-}hs)2pq}{1\text{-}2hspq\text{-}sq^2} + \frac{(1\text{-}s)q^2}{1\text{-}2hspq\text{-}sq^2} = 1$$

That gives the new genotype frequencies. The new allele frequencies are:

$$p' = \frac{p^2 + (1\text{-}hs)pq}{1 - 2hspq - sq^2}, \quad q' = 1 - p' \tag{15}$$

By equation 9 and the above boxed quantities:

$$\text{Cost\_}AA_i = \frac{G}{1\text{-}2hspq\text{-}sq^2} - 1, \quad \text{if} > 0;$$

$$\text{otherwise, Cost\_}AA_i = 0 \tag{16}$$

The total cost of substitution can be iteratively calculated by equations 13, 15, 16, while allowing changes in any parameters at the start of each generation. With parameters held constant, results are shown in Graphs 2 through 5.

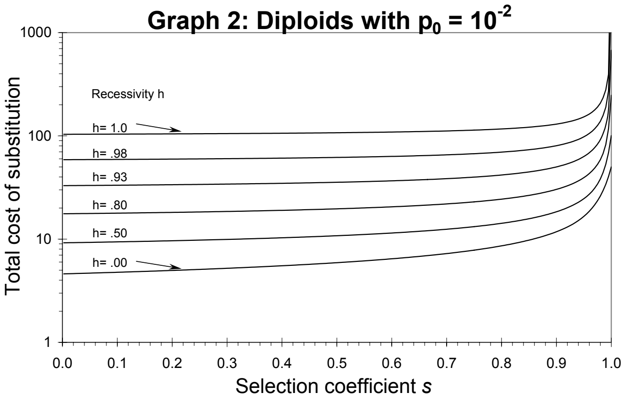
An Autosomal Locus in a Diploid with Inbreeding

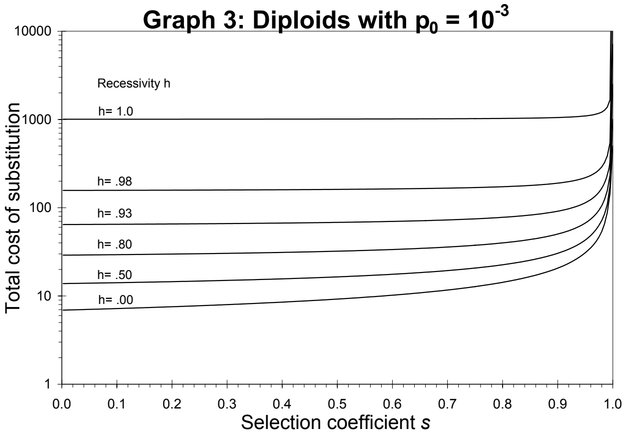
Assume an inbreeding coefficient f, $(0 \le f \le 1)$, where a fraction, f, of the population mates only with its own genotype, and the remainder of the population mates randomly.

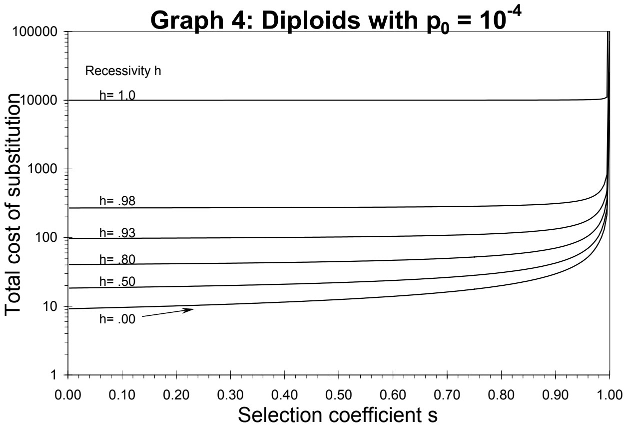
$$\begin{array}{lllllll} \text{Genotype:} & AA & & Aa & & aa & \\ \text{Selective value:} & 1 & & 1\text{-}hs & & 1\text{-}s & \\ \text{Random mate:} & p^2 & + & 2pq & + & q^2 & = 1 \\ \text{Inbreeding:} & 1+\frac{fq}{p} & & 1\text{-}f & & 1+\frac{fp}{q} & \end{array}$$

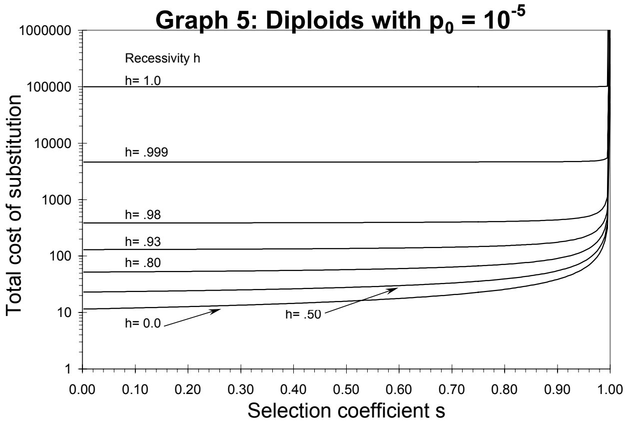
Combining the effects of random mating and inbreeding, the effective frequency of reproducers for each genotype is obtained:

$$\begin{array}{lllllll} \text{Frequency:} & \boxed{p^2+fpq} & + & 2pq(1\text{-}f) & + & q^2+fpq & = 1 \\ \text{Selection:} & p^2+fpq & + & (1\text{-}hs)2pq(1\text{-}f) & + & (1\text{-}s)(q^2+fpq) & = \text{RHS} \end{array}$$


Graph 2: Diploids with $p_0 = 10^{-2}$


Graph 3: Diploids with $p_0 = 10^{-3}$


Graph 4: Diploids with $p_0 = 10^{-4}$


Graph 5: Diploids with $p_0 = 10^{-5}$

where the right-hand side is:

$$\text{RHS} \equiv 1 - spq(2h - 2fh + f) - sq^2 \tag{17}$$

Normalize to obtain new genotype frequencies:

$$\boxed{\frac{p^2+fpq}{\text{RHS}}} + \frac{(1-hs)2pq(1-f)}{\text{RHS}} + \frac{(1-s)(q^2+fpq)}{\text{RHS}} = 1$$

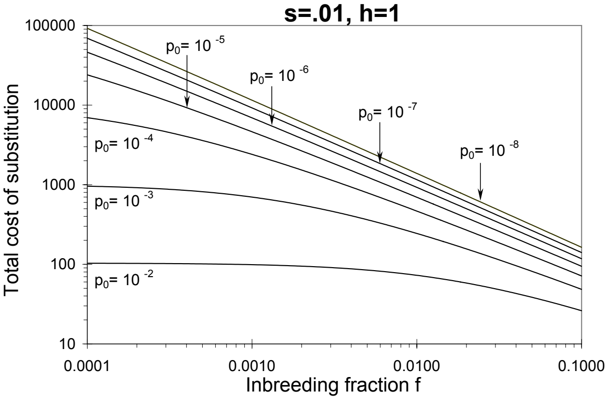
$$p' = \frac{p^2 + pq(1 - hs + fhs)}{\text{RHS}} \quad ; \quad q' = 1 - p' \tag{18}$$

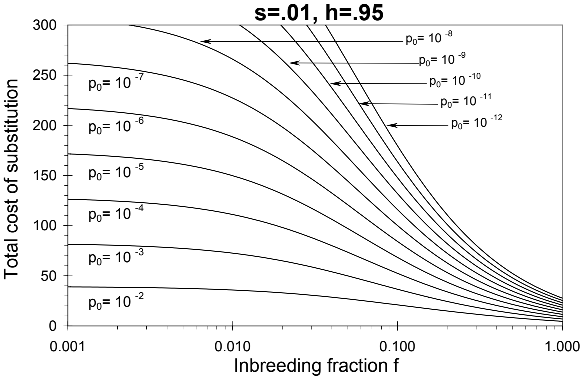
Apply equation 9.

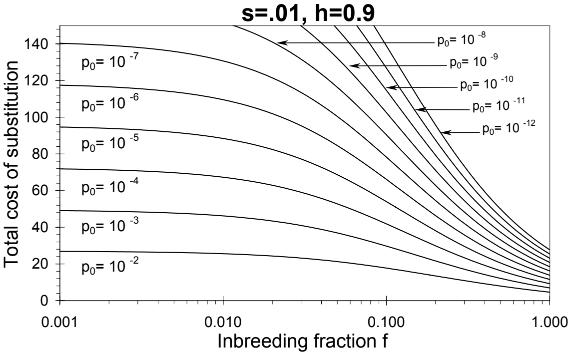
$$\text{Cost\_}AA_i = \frac{G}{\text{RHS}} - 1, \quad \text{if} > 0;$$

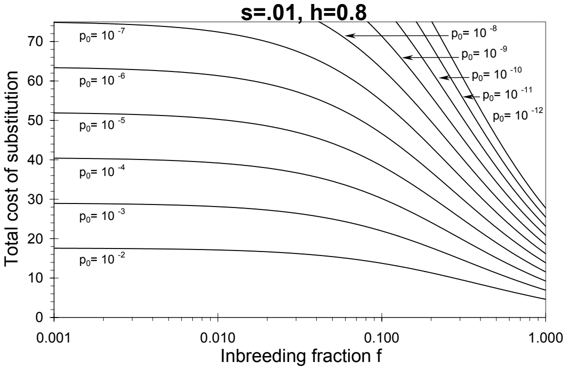
$$\text{otherwise, Cost\_}AA_i = 0 \tag{19}$$

Equations 13, 17, 18, 19 allow iterative calculation of the total cost of substitution, while allowing changes in any parameters at the start of each generation. With parameters held constant, results are shown in Graphs 6 through 9.


Graph 6: Diploids with inbreeding, s=.01, h=1


Graph 7: Diploids with inbreeding, s=.01, h=.95


Graph 8: Diploids with inbreeding, s=.01, h=0.9


Graph 9: Diploids with inbreeding, s=.01, h=0.8

For additional insight, examine a fully dominant substitution, while allowing any inbreeding fraction f, ($0 \le f \le 1$). Define P_A as the number of copies of allele A at the start of generation i. So, $P_A = p \cdot 2N_e$, and at the end of the generation $P_A' = p' \cdot 2N_e \cdot G$, for an increase of $\Delta P_A = p' \cdot 2N_e \cdot G - p \cdot 2N_e$. (By equation 12, $\text{Cost}_i = \text{Cost\_}AA_i$. Using h=0 with equations 17, 18, and 19, and the identity $p^2+pq=p$, it is easily shown that $\text{Cost}_i/\Delta P_A = 1/P_A$.) Therefore:

$$\text{Cost}_i = \frac{\Delta P_A}{P_A}, \quad \text{if} > 0; \quad \text{otherwise, Cost}_i = 0 \tag{20}$$

Thus, in any given generation, the cost of substitution (Cost_i) is given by the percent increase in the number of copies of the substituting allele (regardless of the number of copies of the old allele). This happens to be the same equation as in haploids, and because of that equality, a proof previously established for haploids (ReMine, 2005, appendix) applies to this case in diploids, and the following result is proven. For a fully dominant substitution (with or without any amount of inbreeding fraction f, and allowing a non-constant population size), for a substitution of any given duration, the minimum total cost of substitution is achieved only when the cost each generation (Cost_i) is constant throughout the substitution. In cases of haploids, or diploids with full dominance, cost constancy provides the minimum total cost for a substitution of any given duration. This proof opposes the notion that the total cost can be lowered by arbitrary fluctuations in selection coefficients (as in some versions of soft selection), or fluctuations in population size.

A Sex-Linked Locus in a Diploid

For a sex-linked locus, males and females have different selection coefficients (s_m and s_f) and different costs. Assume the males to be the heterogametic sex, but the results are the same (only swapped) if females are so. Define the male and female allele frequencies: ($p_m+q_m=1$) and ($p_f+q_f=1$). We here assume the *A* allele is neutral or beneficial in both sexes, and disallow less fortunate circumstances (such as beneficial in one gender and harmful in the other). Let the female population size have a growth factor of G_f, and for the male population it is G_m. Assume male and female progeny are produced at the same rate (a sex ratio of 1:1), as is ordinarily the case.

For the case of female progeny, random mating produces the generating function, $(p_m+q_m)(p_f+q_f)=1$, which determines the genotype frequencies given in equation 21.

$$\begin{array}{lllllll} \text{Genotype:} & AA & & Aa & & aa & \\ \text{Selective value:} & 1 & & 1\text{-}hs_f & & 1\text{-}s_f & \\ \text{Frequency} & \boxed{p_m\, p_f} & + & (p_m\, q_f + q_m\, p_f) & + & q_m\, q_f & = 1 \end{array} \tag{21}$$

$$\begin{aligned} \text{Selection: } & p_m\, p_f + (1\text{-}hs_f)(p_m\, q_f + q_m\, p_f) + (1\text{-}s_f)q_m\, q_f \\ & = \text{RHS2} \end{aligned}$$

$$\text{where:} \quad \text{RHS2} \equiv 1 - hs_f\,(p_m\, q_f + q_m\, p_f) - s_f\, q_m\, q_f \tag{22}$$

This is normalized to obtain the new genotype frequencies for females:

$$\boxed{\frac{p_m\, p_f}{\text{RHS2}}} + \frac{(1\text{-}hs_f)(p_m\, q_f + q_m p_f)}{\text{RHS2}} + \frac{(1\text{-}s_f)q_m\, q_f}{\text{RHS2}} = 1 \tag{23}$$

The new allele frequencies for females are:

$$\begin{aligned} & p_f' = \frac{p_m\, p_f + 0.5\,(1\text{-}hs_f)(p_m\, q_f + q_m p_f)}{\text{RHS2}} \quad ; \\ & q_f' = 1 - p_f' \end{aligned} \tag{24}$$

Applying equation 9 gives the cost for genotype AA females. This is the extra reproduction rate necessary to produce enough genotype AA females to satisfy the scenario (including males too, at a progeny sex ratio of 1:1).

$$\begin{aligned} & \mathit{Cost\_AA\_females}_i = \frac{G_f}{\text{RHS2}} - 1\,, \quad \text{if} > 0; \\ & \text{otherwise, } \mathit{Cost\_AA\_females}_i = 0 \end{aligned} \tag{25}$$

In the case of male progeny, for alleles A and *a*, the males are effectively haploid. Male progeny inherit this chromosome from their mother; therefore p_f is the effective frequency of mating that produces the A genotype in male progeny.

$$\begin{array}{lllll} \text{Genotype:} & A & & a & \\ \text{Selective value:} & 1 & & 1\text{-}s_m & \\ \text{Frequency:} & \boxed{p_f} & + & q_f & = 1 \end{array} \tag{26}$$

$$\text{Apply Selection:} \quad p_f + (1-s_m)\, q_f = 1 - s_m\, q_f \tag{27}$$

This is normalized to obtain the new genotype frequencies for males:

$$\boxed{\frac{p_f}{1 - s_m\, q_f}} + \frac{(1 - s_m)\, q_f}{1 - s_m\, q_f} = 1 \tag{28}$$

The new allele frequencies for males are:

$$p_m' = \frac{p_f}{1 - s_m\, q_f} \quad ; \quad q_m' = 1 - p_m' \tag{29}$$

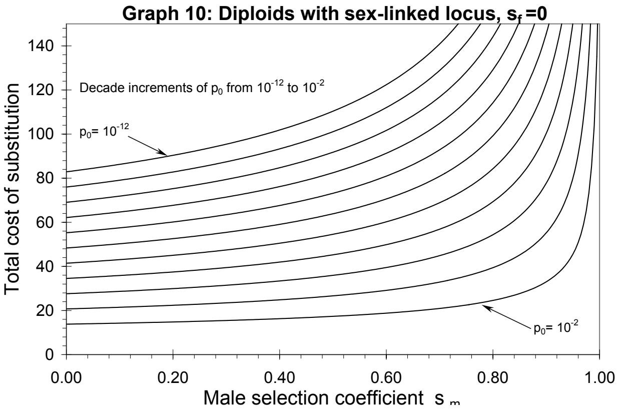
Applying equation 2 gives the cost for genotype A males. Physically this is the extra reproduction rate necessary to produce enough genotype A males to satisfy the scenario requirements (including females too, at a progeny sex ratio of 1:1).

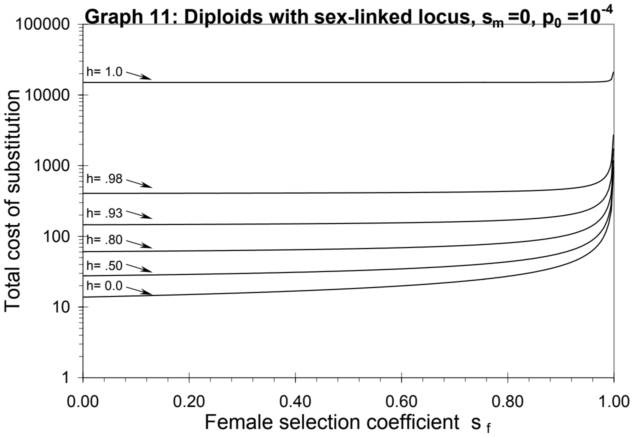
$$\begin{aligned} & \mathit{Cost\_A\_males}_i = \frac{G_m}{1 - s_m\, q_f} - 1\,, \quad \text{if} > 0; \\ & \text{otherwise, } \mathit{Cost\_A\_males}_i = 0 \end{aligned} \tag{30}$$

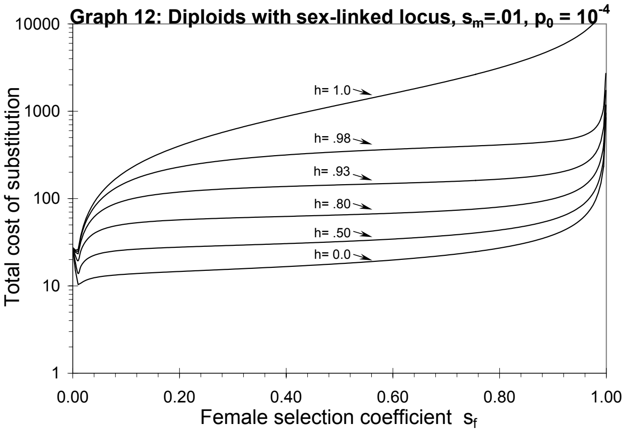
The scenario requires the larger of the two costs, therefore:

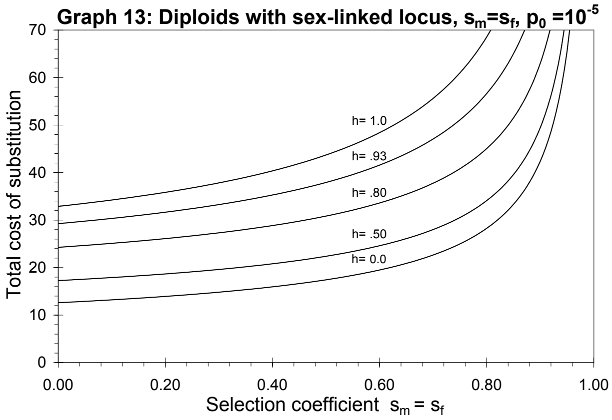
$$\text{Cost}_i = \text{Maximum}(\,\text{Cost\_}A\text{\_}males_i\,,\ \text{Cost\_}AA\text{\_}females_i\,) \tag{31}$$

The total cost of substitution can be iteratively calculated (from equations 3, 22, 24, 25, 29, 30, and 31), while allowing changes in any parameters at the start of each generation. With parameters held constant, results are shown in Graphs 10 through 13.


Graph 10: Diploids with sex-linked locus, s_f =0


Graph 11: Diploids with sex-linked locus, s_m =0, p_0 =10^{-4}


Graph 12: Diploids with sex-linked locus, s_m=.01, p_0 = 10^{-4}


Graph 13: Diploids with sex-linked locus, s_m=s_f, p_0 =10^{-5}

Haldane (1957) assumed tiny s-values (s_m & $s_f \rightarrow 0^+$), where the substitution is very slow, with ample time for allele frequencies to equalize between the sexes, so he assumed p_m=p_f. Unlike Haldane (1957), my iterative calculation allows the sexes to have unequal allele frequencies, and their *appropriate* values are recalculated each generation.

The sex with the higher selection coefficient tends to dominate the substitution process. When the heterogametic sex (male in this study) has the higher selective value (s_m) (as in Graph 10), then the total cost tends to be haploid-like and depends little on the recessivity in females. On the other hand, when the female selective value (s_f) is the higher (as in Graph 11), then the cost takes on the behavior of diploids—for example, the cost increases rapidly with recessivity h. (Note: If the substituting allele is beneficial in one sex, and detrimental or heterosis in the other, then the total cost of substitution will be higher and its behavior will be more complex. However, the cost interactions between the two sexes are too numerous for discussion here.)

For scenarios with constant male and female population sizes (i.e., G_f =G_m =1), the sex with the lower cost will have some reproduction rate leftover. The leftover must be accounted through elimination of individuals of that gender; otherwise it would cause that gender's population size to increase (making G_f >1 or G_m >1). (Also the elimination must be accounted without altering allele frequencies for that gender, which is reserved for the selection coefficients.) In short, the leftover individuals of that gender must be eliminated, even though some of them possess the substituting allele—their production is wasted. On the other hand, when the costs for the two sexes are equal (such that Cost_A_*males* $_i$ = Cost_AA_*females* $_i$), then there is no leftover to waste, so the situation is more efficient at converting the reproduction rate into higher growth for the substituting allele—and this reduces the total cost of substitution. This higher efficiency is seen by the slight dip in total cost at the left end of Graph 12, which occurs when the costs for the two genders tend to be equal throughout the substitution. When selection coefficients for the two sexes are always equal, as in Graph

13, then the total cost is lower than when one selection coefficient is always zero, as in Graphs 11 or 12, but still higher than cases without the sex-linked locus.

Comparison with Haldane's Cost Concept

Haldane (1957) defined the cost of substitution in terms of "genetic death" (originally called "selective death"), which has been a source of much confusion. He did not give an explicit physical definition of genetic death; instead he gave a mathematical equation for it, a different equation for each specific case. In each case, his genetic death concept is identified as the amount subtracted from 1 on the far right-hand side of my equations 4, 14, 17, 22, and 27. In effect, that is how Haldane obtained his equations for genetic death. For example, for an autosomal locus in a diploid, his equation is $2hspq+sq^2$ (see it on the right side of my equation 14), which he interpreted as the genetic deaths of the heterozygote ($2hspq$), plus the genetic deaths of the disfavored homozygote (sq^2).

As pointed out in ReMine (2005), Haldane's genetic death concept exists only *between the mathematical steps* of (1) selection and (2) normalization. But in nature those happen physically in one event, not in separate steps. Therefore, Haldane's genetic death concept is a mathematical phantom that does not exist in physical reality. For this and many additional reasons, genetic death has been a continual source of confusion.

My cost concept uses quantities (enclosed in boxes here) that appear immediately before selection is applied and immediately after normalization is applied. Thus, these values straddle the line used by Haldane. Unlike Haldane, I use quantities (the genotype frequencies enclosed in boxes) whose physical meaning is concrete, and whose use is already ubiquitous in population genetics. I suggest his concept of genetic death is physically unclear and should be abandoned.

Haldane's concept worked correctly only under his two assumptions: (1) constant population size, and (2) small selection coefficients ($s \rightarrow 0^+$). His 1957 paper studied the same cases examined here. For each case, and for each generation, after applying Haldane's two assumptions, my cost of substitution is identical to his. (That is, under his two assumptions, my cost equations 6, 16, 19, 25, and 30 reduce to his cost equations 1, 5, 10, and the second equation on his page 518.) Under his two assumptions, my cost *each generation* is identical to his, therefore the *total* cost of substitution is also identical. However, my cost equations for any given generation are accurate for all selection coefficients (even large selection coefficients), and for fluctuations in any parameters (such as population size)—whereas Haldane's equations are not.

In the case of a sex-linked locus, Haldane's focus on genetic death led to error. In each generation, he calculated the genetic deaths of disfavored males, and totaled that over the entire substitution. He did likewise for females. Apparently to calculate the average rate of genetic death, he then averaged those two totals together (instead of taking the larger of the two, as in my equation 31), to obtain his total cost of substitution. The genetic death concept focuses on death (not reproduction rate), so it failed to inform Haldane that the costs for males and females are linked together, because male and female progeny are produced at the same rate. This is yet another example of how Haldane's concept of genetic death caused confusion and error.

Discussion

Under the above variety of genetic circumstances, the parameters were held constant, and the graphs display the total cost of substitution. Each curve plots 201 data points, (200 intervals), equally spaced visually across the horizontal range. In Graphs 1 through 5, there is a high cost impulse at $s=1$, which goes off-scale vertically.

The graphs show that the total cost increases with selection coefficient s. (That holds for all cases, except for some complicated interactions involving a sex-linked locus in diploids, as discussed above.) The traditional total cost formulas assumed infinitesimally small selection coefficients ($s \rightarrow 0^+$), and therefore underestimate the total cost in realistic situations. However, the underestimate is less than ten percent error as long as $s<0.17$, which covers the most typical evolutionary substitutions.

The total costs for "haploids" and "fully dominant substitutions in diploids" are nearly identical (so only the former is graphed here). They diverge from each other only slightly as the selection coefficient increases, with diploids having the slightly lower total cost. Using the haploid case as a reference, the diploid case has lower total cost by no more than 1.4 percent (9.2 percent) for all selection coefficients less than 0.2 (0.8).

Traditionally the total cost of semi-dominant substitutions ($h=0.5$) is given as double that of dominants ($h=0$). That is accurate to within ten percent for selection coefficients $s<0.35$, according to these results.

It is known that almost all beneficial mutations fail to substitute; instead they are eliminated by genetic drift without ever reaching fixation. In such instances, they only partially substitute within a population, and they incur a portion of the total cost of substitution. Even though such substitutions are unsuccessful, they are nonetheless part of the scenario, and they incur a cost that the species must pay if the scenario is to be plausible. Haldane did not account this *cost of unsuccessful substitution* (C_U). When a single beneficial mutation is lost, this cost may be small, but it becomes significant when multiplied by its overwhelming prevalence in nature. Kimura and Maruyama (1969) estimated that this adds 2 to the cost of each successful substitution. Their calculation should be revisited, especially in light of: (a) recent clarifications of the cost concept, and (b) modern measurements concerning the mutation rate and the frequency of nondominant mutations (which have a high cost of unsuccessful substitution).

Of special concern is the high cost of recessive substitution. Their total cost is quite high, (easily in the thousands), and substantially constant over any selective value s. Graphs 6 through 9 show the effects of inbreeding, which may dramatically lower the cost of recessive substitutions. For example, at a starting frequency of $p_0=10^{-5}$ and no inbreeding (f=0), the total cost is 100,000, but is reduced to 7,000 by an inbreeding fraction, f=0.001. However, for a high starting frequency of $p_0=10^{-2}$, that same inbreeding fraction has virtually no effect. In other words, a given level of inbreeding may produce dramatic cost reductions when starting frequencies are low, but at high starting frequencies there is scarcely any reduction. This is shown in Graph 6.

For the same reason, however, inbreeding simultaneously increases costs. Inbreeding causes an increased production of homozygotes that express the beneficial recessive mutation and aid its substitution. Thus, some beneficial recessive mutations (that would be lost at lower cost under no inbreeding) will successfully substitute under inbreeding—thus incurring a high cost of substitution. Also, some beneficial recessive mutations (that would substitute only slightly under no inbreeding) will substitute further under inbreeding—thus increasing the cost of unsuccessful substitution. In other words, it is not immediately obvious whether inbreeding will increase or decrease the costs associated with substitution, or by how much. The role of inbreeding is a double-edged sword that cuts in both directions, and needs further research.

The costs derived in this paper are mechanical and unavoidable. In realistic situations the actual cost of substitution can be higher, but cannot be lower. The equations give a lower bound, regardless of the type of selection (such as "soft selection" versus "hard selection"), and regardless of environmental change. This lower bound in *any single generation* (as shown in equations 6, 16, 19, 25, 30, and 31) is independent of those factors, and therefore so is the total, summed over all generations. This means that, contrary to common assumption, environmental change (such as employed in Felsenstein, 1971 and 1972), and soft selection (such as employed in Grant and Flake, 1974, and Wallace, 1991) cannot even slightly reduce cost problems for single substitutions. Those errors were caused by confusion factors identified in ReMine (2005), such as genetic death, genetic load, extinction, and the environment. Those factors are largely irrelevant to the central issue, which is the *growth* (in number of copies) of the substituting allele, and the extra reproduction rate necessary to achieve it.

Haldane (1957) estimated that the average value for the total cost of substitution is 30. To rationalize that low figure, he employed an environmental-change scenario to bypass the earliest stages of substitution, where most of the total cost would ordinarily be incurred. The scenario (given in Haldane, 1957 and 1961) assumes that neutral mutations and slightly deleterious mutations occasionally drift upward to maintain a pool of these at elevated frequencies—call these "pool mutations"—and when the environment changes, some of these are converted to beneficial, which then begin their substitution at an elevated starting frequency, thereby dramatically lowering their total cost of substitution. This scenario is awkward to justify for most substitutions (much less for all substitutions, as assumed by Haldane), and especially during multiple-concurrent substitutions, because the same environmental change alleged to start one substitution would likely disrupt or reverse the ongoing substitutions.

As discussed in ReMine (1993), the environmental-change scenario tries to obtain a "free lunch" from nature. Environmental-change, like mutation, is random concerning the organism, and therefore is far more harmful than beneficial. For every "neutral or slightly deleterious" mutation that converts to "beneficial," there will be vastly more converted to become more harmful, with these also tending to have a larger effect—for a net effect that is overwhelmingly harmful. If one pool mutation converts from neutral to beneficial, this would be overwhelmed by thousands that become detrimental and must be eliminated—and precisely because these eliminations start at elevated frequencies, this increases the cost of mutation (which is the extra reproduction rate required for coping with the elimination of harmful mutations). That makes cost problems worse. In a similar way, for every pool mutation that converts to beneficial, there would be many beneficial alleles at fixation that convert to harmful, and must be substituted just to break even. To avoid a theoretical Ponzi scheme, almost all of these must be substituted by a directly new mutation (not a pool mutation), thereby incurring a full (non-discounted) cost of substitution, and usually resulting in a net fitness loss. When fully accounted, the environmental-change scenario makes cost problems worse. There is no free lunch.

In other words, Haldane gave poor justification for his figure of 30, and we must seek a more realistic figure, which will surely be higher. Even after removing the environmental-change scenario, the value of 30 implicitly assumes: (1) a nearly nonexistent role for recessive substitutions, (2) no cost for unsuccessful substitutions, and (3) a high reliance on small population sizes (which receive beneficial mutations at a much lower rate and slows down beneficial evolution). It is doubtful those assumptions are realistic or consistent with other considerations from evolutionary theory.

Conclusions

Compared to the traditional cost derivations, the equations described in this paper are general-purpose, simpler, and easier to understand. The total cost of substitution is then calculated iteratively, generation by generation, for improved accuracy. This method allows arbitrary changes to any parameters at the start of each generation, which can be conveniently handled by computers. In addition, this method is transparently clear, whereas Haldane's derivations are opaque.

Though the environmental-change scenario was traditionally used, it was never given a proper cost analysis. When fully tallied it actually makes the cost problems worse. The same applies to the concept of soft selection (at least for single substitutions).

The following areas are suggested for further research in cost theory: the frequency and role of nondominant or recessive mutation, the overall effect of inbreeding, the cost of unsuccessful substitutions, and the effect of unequal sex ratios at conception or in the breeding population.

Most importantly, this paper calls for the average total cost of substitution (such as Haldane's figure of 30) to be given an explicit justification within some widely accepted evolutionary model. The graphs and methods presented in this paper can help estimate it across a blend of circumstances anticipated in nature.

Appendix: Notes on Haldane's Derivations

Haldane's equations for the cost *each generation* are typographically correct (Haldane, 1957). However, I detect the following typographical errors in his derivations for the *total* cost of substitution. Most of these are clearly typographical errors, because they disappear in his subsequent equations. Some of them, however, are in his final equations for the total cost of substitution, so researchers should be wary of that. (A) In the calculation prior to his equation 7, there should be a factor (1–2h) multiplying the second p_0. (B) In his equation 9, the first *K* in the denominator should be *k*. (C) In the calculation following his equation 9, there should be an opening bracket immediately between the μ and the second natural logarithm, mated with a closing bracket immediately after the last parenthesis of that equation. (D) In his equation 11, it may be that the *K* should be 2*K*, because the missing "2" shows up again in his next equation. (E) In the equation before his equation 12, the sign before the ½ should be negative. (F) In his equation 12, the sign before his second natural logarithm should be negative. (G) In his equation 13 (in his equation for female, and in his separate equation for male), the first p_0 should be a factor in the denominator, not the numerator. For the reader's convenience, my paper uses modern notation rather than Haldane's.

Haldane's subsequent paper (Haldane, 1961) attempts to derive "more precise" cost equations when selection coefficients are not small. I regard that paper as a decrease in clarity and accuracy from his previous attempt (1957), and largely irrelevant to understanding cost theory today. Absent is his previous emphasis on "reproductive capacity," which should be the central issue. Instead the 1961 paper focuses exclusively on genetic death, which should now be abandoned as unnecessary confusion. In effect, his 1961 paper takes each of his previous equations for genetic death (only this time allowing a large selection coefficient) and pursues the task of deriving a closed-form equation for its sum over all generations of the substitution (to obtain the total cost of substitution). The paper's lengthy mathematical contrivances are opaque and conceptually ill founded. For example in haploids, in a given generation, his cost of substitution, sq, is accurate only for a *small* selection coefficient. When his 1961 paper imposes a *large* selection coefficient into that same equation, the result is useless or erroneous for its intended purpose (that of establishing a limit on the substitution rate). His 1961 paper thereby arrives at erroneous cost equations, and mistaken conclusions, such as the notion of "negative cost."

Acknowledgements

This research was supported in part by a grant from Discovery Institute.

References

Felsenstein, J. 1971. On the biological significance of the cost of gene substitution. *Am. Nat.* 105:1–11.

Felsenstein, J. 1972. The substitutional load in a finite population. *Heredity* 28:57–69.

Grant, V., and R.H. Flake. 1974. Solutions to the cost-of-selection dilemma. *Proceedings of the National Academy of Science, U.S.A.* 71:10:3863–3865.

Haldane, J. B. S. 1957. The cost of natural selection. *Journal of Genetics* 55:511–524.

Haldane, J. B. S. 1961. More precise expressions for the cost of natural selection. *Journal of Genetics* 57:2–3:351–360.

Kimura, M., and T. Maruyama. 1969. The substitutional load in a finite population. *Heredity* 24:101–114.

ReMine, W.J. 1993. *The Biotic Message.* Saint Paul Science publishers, Saint Paul, MN.

ReMine, W.J. 2005. Cost theory and the cost of substitution—a clarification. *TJ* 19:1:113–125.

Wallace, B. 1991. *Fifty Years of Genetic Load—An Odyssey*. Cornell University Press, Cornell, NY.