Mutation-Selection Balance at a Modifier-of-Imprinting Locus

Hamish G. Spencer and Janine A. Barnett

Department of Zoology, University of Otago, Dunedin, New Zealand

Manuscript received July 14, 1995
Accepted for publication May 3, 1996

ABSTRACT

We propose a pair of population genetic models for a modifier-of-imprinting locus for which different genotypes imprint different proportions of an imprintable target locus in their gametes. The two models examine the situations in which imprinting is advantageous or disadvantageous, and we discuss three cases for which the modifier is respectively partially dominant, dominant, or recessive. The models predict the stable equilibrium frequencies of the mutant modifier and functionally diploid individuals in a large population in terms of up to four parameters: the mutation rate at the modifier locus, ν; the selection coefficient against the disadvantageous phenotype, s; the proportion of unimprinted eggs produced by homozygotes for the mutant modifier, θ, and, in the partially dominant models, the dominance parameter, k. The equilibrium frequency of the mutant phenotypes is shown to be approximately twice that of standard Mendelian models: $2\nu/s$ or $4\nu/s$ when the modifier is recessive or dominant, respectively. Mathematical equivalences between these and nonimprinting models are noted.

A gene whose expression depends on which parent passed it on is said to be imprinted (HALL 1990). One of the commonest forms of genomic imprinting is inactivation of the gene derived from parents of one sex (BARLOW 1995). For example, the insulin-like growth factor II (IGF-2) locus in humans and mice is normally maternally imprinted, with only the paternally derived copy being active (OGAWA *et al.* 1993). Just how imprinting occurs is not currently known, although the methylation of DNA has been strongly implicated (GOLD and PEDERSEN 1994).

The population-level consequences of genomic imprinting have only recently begun to be explored. SVED and LAIRD (1990) and SPENCER and WILLIAMS (1995), for example, have examined models for particular genetic systems, Fragile-X syndrome and IGF-2, respectively. PEARCE and SPENCER (1992) proposed several viability-selection models of a single imprintable locus and noted numerous formal equivalences between their models and those for standard Mendelian loci. SPENCER and WILLIAMS (1996) examined two-locus models involving, in addition to the primary, imprintable locus, a modifier-of-imprinting locus that, along with the gender of the transmitting parent, determined the imprinting status at the primary locus. They showed that imprinting was unlikely to arise unless it conferred a direct selective advantage. Moreover, if imprinting were advantageous, any modifiers of imprinting were likely to result in high penetrance of imprinting. That is, the modifiers would cause all offspring to be imprinted, rather than some lesser proportion. Conversely, if imprinting were disadvantageous, modifiers would be selected that imprinted none of the gametes. In this paper we extend these models to examine the equilibrium between mutation at the modifier locus and viability selection arising from the imprinting status at the primary locus.

The models are, in the first instance, motivated by the case of IGF-2 in humans. Although most individuals are functionally haploid at IGF-2, some individuals have both paternally and maternally derived genes active. These individuals are predisposed to the childhood cancer Wilms' tumor (OGAWA *et al.* 1993). SPENCER and WILLIAMS (1995) put forward a modifier model to explain the expected population frequency of this failure to imprint. They hypothesized that a failure to imprint the IGF-2 locus was due to a recessive modifier maintained in the population by recurrent mutation. Offspring of females homozygous for the mutant were functionally diploid. We first generalize their model by allowing the mutant to have partial penetrance, examining in detail the most general case in which the mutant is partially dominant. We then look at the cases in which the mutant is fully dominant and recessive. In our second model we look at the three corresponding cases for the situation in which imprinting is disadvantageous, presumably the majority of loci.

THE MODELS

In all of our models we consider a primary locus, A, at which just one allele, A, is found. (Alternatively, different alleles at A can be considered strictly neutral and thus equivalent under selection.) The A locus may be subject to maternal imprinting, by which we mean that the maternally derived gene is not expressed. The

Corresponding author: Hamish G. Spencer, Department of Zoology, University of Otago, P.O. Box 56, Dunedin, New Zealand.
E-mail: h.spencer@otago.ac.nz

TABLE 1

Proportion of imprinted offspring for female genotypes in different cases

	Case					
Genotype	AR	AP	AD	DR	DP	DD
M_1M_1	1	1	1	0	0	0
M_1M_2	1	$1-k\theta$	$1-\theta$	0	$1-k\theta$	$1-\theta$
M_2M_2	$1-\theta$	$1-\theta$	$1-\theta$	$1-\theta$	$1-\theta$	$1-\theta$

modifier locus, M, has two alleles, M_1 and M_2, the polymorphism being maintained by a balance between selection against the M_2 allele (strictly speaking, against a proportion of the offspring of M_2-bearing females) and mutation from M_1 to M_2 at a rate ν. We classify the models according to whether imprinting is advantageous or disadvantageous (A or D) and examine three cases of each in which the mutant M_2 allele is, respectively, partially dominant, dominant or recessive (P, D or R). Both of the mother's M genes affect the imprinting status of her offspring, which means the modifier is a genotypic modifier in the terminology used in SPENCER and WILLIAMS (1996). In Model A (in which imprinting is advantageous), individuals usually express only their paternally derived A gene; when it is disadvantageous (Model D), most individuals are functionally diploid, expressing both A genes. An imprinted A gene is indicated by placing it in parentheses: (A). We explain the first model in detail as a template for the exploration of the other models.

MODEL A – ADVANTAGEOUS IMPRINTING

Case P: M_2 partially dominant (AP): The first model is motivated by the case of IGF-2, in that imprinting is advantageous: functional diploidy at the A locus is selected against. We first examine the most general case in which the mutant modifier M_2 is partially dominant by assuming that a proportion, θ, of offspring of females homozygous for M_2 express both their A genes and that the corresponding proportion of unimprinted offspring of heterozygous (M_1M_2) females is $k\theta$. The parameter θ can be considered to be the penetrance of the M_2M_2 genotype, whereas k can be thought of as the degree of dominance of M_2. (The imprinting status of offspring of the three possible M genotypes are shown for all cases of both models in Table 1.) Individuals who possess two active copies of the A gene have a relative viability of $1 - s < 1$. Under these assumptions, the frequency of the M_2 allele reaches an equilibrium at which the number of M_2 alleles removed by selection is exactly balanced by the number added by mutation.

Following the method in SPENCER and WILLIAMS (1995), we let g_1, g_2 and g_3 be, respectively, the frequencies of M_1M_1, M_1M_2 and M_2M_2 genotypes in the adult population after selection with the usual restraint that $\Sigma g_i = 1$. Males produce during meiosis just two types of gametes, AM_1 and AM_2, at respective frequencies of

$$y_1^m = g_1 + \tfrac{1}{2}g_2 \quad \text{and} \quad y_2^m = \tfrac{1}{2}g_2 + g_3, \tag{1}$$

whereas females produce four types of eggs: $(A)M_1$, $(A)M_2$, AM_1 and AM_2, respective frequencies

$$\begin{aligned} x_1^m &= g_1 + \tfrac{1}{2}(1 - k\theta)g_2, \\ x_2^m &= \tfrac{1}{2}(1 - k\theta)g_2 + (1 - \theta)g_3, \\ x_3^m &= \tfrac{1}{2}k\theta g_2 \quad \text{and} \\ x_4^m &= \tfrac{1}{2}k\theta g_2 + \theta g_3. \end{aligned} \tag{2}$$

We now incorporate mutation, which occurs from the M_1 allele to the M_2 allele at a rate ν. (We can ignore mutation in the other direction since the M_2 allele will be rare because of the selection operating against the offspring of M_2-bearing females.) The postmutation frequencies of the various gametes are thus given by

$$\begin{aligned} y_1 &= (1 - \nu)y_1^m \quad \text{and} \\ y_2 &= y_2^m + \nu y_1^m, \quad \text{and} \\ x_1 &= (1 - \nu)x_1^m, \\ x_2 &= x_2^m + \nu x_1^m, \\ x_3 &= (1 - \nu)x_3^m \quad \text{and} \\ x_4 &= x_4^m + \nu x_3^m. \end{aligned} \tag{3}$$

These gametes combine to produce zygotes, which are then subject to viability selection according to their imprinting status. These postselection zygote frequencies are the genotype frequencies of the next generation:

$$\begin{aligned} \bar{w}g_1' &= x_1y_1 + (1 - s)x_3y_1, \\ \bar{w}g_2' &= x_1y_2 + x_2y_1 + (1 - s)(x_4y_1 + x_3y_2) \quad \text{and} \\ \bar{w}g_3' &= x_2y_2 + (1 - s)x_4y_2, \end{aligned} \tag{4}$$

in which

$$\bar{w} = 1 - s\theta(g_3 + kg_2). \tag{5}$$

Here $\bar{w}$ is the mean viability of the population, the sum of the right side of Equations 4, so that $\Sigma g_i' = 1$. The value for $\bar{w}$ makes intuitive sense because θ of the offspring of M_2M_2 females and $k\theta$ of the offspring of M_1M_2 females are selected against, both with selection coefficient s. This constraint means that there are just two independent recursion equations in this system. Substituting $p = g_1 + \tfrac{1}{2}g_2$ gives two recursion equations in two variables, chosen for their computational convenience:

$$\begin{aligned} p' &= (p(1 - \tfrac{1}{2}s\theta(1 - 2p + g_1)) - \tfrac{1}{2}sk\theta((p - g_1) \\ &\quad \times (1 + 2p)))(1 - v)/\bar{w} \quad \text{and} \\ g_1' &= p(p - sk\theta(p - g_1))(1 - \nu)^2/\bar{w}. \end{aligned} \tag{6}$$

The mean viability reduces to

$$\bar{w} = 1 - s(\theta(1 - 2p + g_1) + 2k\theta(p - g_1)). \quad (7)$$

An equilibrium is reached when both $p' = p$ and $g_1' = g_1$, and there are three possible solutions. The first is a trivial equilibrium ($\hat{p} = 0$, $\hat{g}_1 = 0$) and not biologically interesting. (Indeed, this equilibrium only exists because we ignore back mutation. In any real situation in which p was equal or close to zero, back mutation would be important.) The second is unfeasible (*i.e.*, $\hat{p} \notin (0,1)$), except when k is close to one, in which case it remains biologically irrelevant because it is unstable. [Indeed, this second equilibrium corresponds to the universally ignored second equilibrium of a standard one-locus Mendelian partial dominant (CROW and KIMURA 1970: 260)]. The third, but interesting, equilibrium point is given by

$$\hat{p} = 1 - \frac{2\nu}{sk\theta} \quad \text{and}$$

$$\hat{g}_1 = 1 - \frac{4\nu}{sk\theta} \quad (8)$$

with terms of ν^2 and above ignored. The equilibrium proportion of unimprinted female gametes (and hence also functionally diploid individuals) is given by

$$\hat{x}_3 + \hat{x}_4 = k\theta\hat{g}_2 + \theta\hat{g}_3 \approx \frac{4\nu}{s}, \quad (9)$$

which does not depend on θ or k. This independence is because (at least to linear order) the effect of the size of k and θ on the frequency of phenotypic expression of M_2 is exactly balanced by the strength of selection on these phenotypes.

To examine the stability of this equilibrium, we linearize Equations 6 around the approximated equilibrium 8, (see, *e.g.*, ROUGHGARDEN 1979 for a description of this technique) and obtain a 2×2 matrix for which the leading eigenvalue, λ, is

$$\lambda = 1 - \tfrac{1}{2}sk\theta, \quad (10)$$

ignoring terms in ν and above. Clearly, for biologically realistic parts of parameter space we have that $0 < \lambda < 1$ and our equilibrium is stable. This result should fit with our intuition: selection is removing the M_2 allele, while mutation is increasing its frequency, and these two forces should balance in a stable equilibrium.

Case D: M_2 dominant (AD): We now examine the case of a completely dominant mutant. The parameter θ now describes the proportion of offspring of females bearing at least one mutant modifier, M_2, that expresses both their A genes. Thus, M_1 is recessive to M_2 (see Table 1). In effect, the AD case is the AP case with $k = 1$. Indeed, letting $k = 1$ does not cause any mathematical singularities, and without any further calculations, therefore, we have the M_2 allele equilibrium frequency of

$$\hat{q} = \frac{2\nu}{s\theta} \quad (11)$$

and the frequency of M_1M_1 homozygotes is

$$\hat{g}_1 = 1 - \frac{4\nu}{s\theta}, \quad (12)$$

again with terms of ν^2 and above ignored. The equilibrium proportion of unimprinted female gametes is given by

$$\hat{x}_3 + \hat{x}_4 = \theta(1 - \hat{g}_1) \approx \frac{4\nu}{s}, \quad (13)$$

the same value as for the AP case.

Ignoring terms in ν and above, the leading eigenvalue for the linearized equilibrium system is

$$\lambda = 1 - \tfrac{1}{2}s\theta. \quad (14)$$

Our equilibrium is again stable for biologically realistic parts of parameter space.

Case R: M_2 recessive (AR): The final case is that in which M_2 is recessive to M_1 and corresponds to fixing k to be 0. That is, a proportion, θ, of offspring of females homozygous for a mutant modifier, M_2, express both their A genes, whereas females who bear at least one M_1 allele have children with their maternally derived A gene inactivated. As before, individuals who possess two active copies of the A gene, in this case θ of those whose mothers were homozygous for the M_2 allele, have a relative viability of $1 - s < 1$. This case is a generalization of the model proposed for IGF-2 by SPENCER and WILLIAMS (1995), in which θ was assumed to be 1.

Equations 1–5 are unchanged, although with $k = 0$, some variables (*e.g.*, x_3, which is always zero) drop out. It is more convenient to use g_3 rather than g_1 in the two variable recursion, which is

$$p' = p(2 - s\theta g_3)(1 - \nu)/2\bar{w} \quad \text{and}$$

$$g_3' = (1 - p(1 - \nu))(1 - p(1 - \nu) - s\theta g_3)/\bar{w}. \quad (15)$$

Some algebra shows that $\bar{w} = 1 - s\theta g_3$ (which makes sense since only θ of the offspring of M_2M_2 females are selected against). A trivial equilibrium ($\hat{p} = 0$, $\hat{g}_3 = 1$) is not biologically interesting (again existing only because we ignore back mutation). Once we have that $p \neq 0$, the first of these equations immediately gives the stationary value for the proportion of M_2M_2 adults in the population

$$\hat{g}_3 = \frac{2\nu}{s\theta(1 + \nu)}. \quad (16)$$

Ignoring terms in ν^2 and higher powers of ν, we obtain the approximation

$$\hat{g}_3 = \frac{2\nu}{s\theta}, \quad (17)$$

and, again ignoring a biologically unfeasible solution for $\hat{p}$ (*i.e.*, one that is >1),

$$\hat{p} = 1 - \sqrt{\frac{2\nu}{s\theta}}, \tag{18}$$

ignoring terms in $\nu^{3/2}$ and greater. The equilibrium proportion of unimprinted female gametes (and hence also functionally diploid individuals) is given by

$$\hat{x}_3 + \hat{x}_4 = \theta\hat{g}_3 \approx \frac{2\nu}{s}, \tag{19}$$

half that for the dominant cases, but again independent of θ.

The leading eigenvalue for the linearization of Equations 15 around the approximated Equilibrium 17 and 18 is

$$\lambda = 1 - \sqrt{2\nu s\theta} + 2\nu, \tag{20}$$

ignoring terms in $\nu^{3/2}$ and above. Again, for biologically realistic parts of parameter space (*i.e.*, when selection is sufficiently strong, $s > 2\nu/\theta$), we have that $0 < \lambda < 1$, and our equilibrium is stable.

MODEL D – DISADVANTAGEOUS IMPRINTING

Case P: M_2 partially dominant (DP): We now consider a model in which imprinting is selectively disadvantageous, a situation we expect pertains to almost all loci. Selection thus acts against functionally haploid individuals. In this first case, DP, the mutant M_2 is partially dominant over M_1. The wild-type allele, M_1, is now a Mendelizing modifier, and no females bearing two copies of this allele imprint their eggs. [We can assume this proportion is zero because of the result of SPENCER and WILLIAMS (1996) that wild-type modifiers will quickly evolve to prevent any imprinting when it is disadvantageous.] A proportion, θ, of offspring of M_2M_2 females and a proportion, $k\theta$, of offspring of heterozygous (M_1M_2) females express both their A genes (see Table 1). All other details are unchanged.

Using the same symbols for the three genotype frequencies as in Model A, sperm are produced at the same frequencies (Equations 1). The four types of eggs, $(A)M_1$, $(A)M_2$, AM_1 and AM_2, are now produced at frequencies

$$\begin{aligned} x_1^m &= \tfrac{1}{2}(1 - k\theta)g_2, \\ x_2^m &= \tfrac{1}{2}(1 - k\theta)g_2 + (1 - \theta)g_3, \\ x_3^m &= g_1 + \tfrac{1}{2}k\theta g_2 \quad \text{and} \\ x_4^m &= \tfrac{1}{2}k\theta g_2 + \theta g_3, \end{aligned} \tag{21}$$

which are Equations 2 with the roles of x_1^m and x_2^m exchanged, respectively, with x_3^m and x_4^m, and θ and $k\theta$ replaced, respectively, by $1 - \theta$ and $1 - k\theta$.

The equations for the genotype frequencies in the next generation are now

$$\begin{aligned} \bar{w}g_1' &= (1 - s)x_1y_1 + x_3y_1, \\ \bar{w}g_2' &= (1 - s)(x_1y_2 + x_2y_1) + x_3y_2 + x_4y_1 \quad \text{and} \\ \bar{w}g_3' &= (1 - s)x_2y_2 + x_4y_2, \end{aligned} \tag{22}$$

in which

$$\bar{w} = 1 - s(g_2(1 - k\theta) + g_3(1 - \theta)). \tag{23}$$

Again substituting $p = g_1 + \tfrac{1}{2}g_2$ and using the constraint that $\Sigma g_i' = 1$ gives two recursion equations in two variables,

$$\begin{aligned} p' = {} & (p(1 - \tfrac{1}{2}s(1 - \theta)(1 - 2p + g_1)) \\ & - \tfrac{1}{2}s(1 - k\theta)(p - g_1)(1 + 2p))(1 - \nu)/\bar{w}, \\ g_1' = {} & p(p - s(1 - k\theta)(p - g_1))(1 - \nu)^2/\bar{w}, \end{aligned} \tag{24}$$

in which

$$\begin{aligned} \bar{w} = 1 - s((1 - \theta)(1 - 2p + g_1) \\ + 2(1 - k\theta)(p - g_1)), \end{aligned} \tag{25}$$

which are Equations 6 and 7 with the $k\theta$ term replaced by a $(1 - k\theta)$ term and the θ term replaced by one involving $(1 - \theta)$. This formal equivalence between Model A and Model D arises because these terms are the frequency of selectively disadvantaged zygotes produced by, respectively, the heterozygous M_1M_2 females and homozygous M_2M_2 females. In Model A, these are unimprinted individuals; in Model D, they are imprinted. These equations lead to one nontrivial feasible equilibrium of biological interest:

$$\begin{aligned} \hat{g}_1 &= 1 - \frac{4\nu}{s(1 - k\theta)}, \\ \hat{p} &= 1 - \frac{2\nu}{s(1 - k\theta)}, \end{aligned} \tag{26}$$

with terms of ν^2 and above ignored. The equilibrium proportion of imprinted female gametes is given by

$$\hat{x}_1 + \hat{x}_2 = (1 - k\theta)\hat{g}_2 + (1 - \theta)\hat{g}_3 \approx \frac{4\nu}{s}. \tag{27}$$

Equilibrium 26 is stable because the leading eigenvalue for the linearization of Equations 24 and 25 around it is

$$\lambda = 1 - \tfrac{1}{2}s(1 - k\theta), \tag{28}$$

ignoring terms in ν and above.

Case D: M_2 dominant (DD): We now allow the imprinting modifier, M_2, to be completely dominant by fixing k to be 1. The parameter θ now describes the proportion of offspring of females bearing at least one mutant modifier, M_2, that expresses both their A genes (see Table 1). Fixing $k = 1$ introduces no novel mathematics, the recurrence equations reducing to

$$p' = (p + \tfrac{1}{2}s(1 - \theta)(g_1 - 2p + g_1p))(1 - \nu)/\bar{w}$$

TABLE 2

Summary of equilibrial frequencies and eigenvalues in different cases

	Case					
	AR	AP	AD	DR	DP	DD
Mutant allele (M_2)	$\sqrt{\frac{2\nu}{s\theta}}$	$\frac{2\nu}{sk\theta}$	$\frac{2\nu}{s\theta}$	$\sqrt{\frac{2\nu}{s(1-\theta)}}$	$\frac{2\nu}{s(1-k\theta)}$	$\frac{2\nu}{s(1-\theta)}$
Mutant phenotype	$\frac{2\nu}{s}$	$\frac{4\nu}{s}$	$\frac{4\nu}{s}$	$\frac{2\nu}{s}$	$\frac{4\nu}{s}$	$\frac{4\nu}{s}$
Leading eigenvalue	$1-\sqrt{2\nu s\theta}$	$1-\frac{1}{2}sk\theta$	$1-\frac{1}{2}s\theta$	$1-\sqrt{2\nu s(1-\theta)}$	$1-\frac{1}{2}s(1-k\theta)$	$1-\frac{1}{2}s(1-\theta)$

and

$$g_1' = (p - s(1-\theta)(p-g_1))(1-\nu)^2/\bar{w}, \tag{29}$$

in which

$$\bar{w} = 1 - s(1-\theta)(g_2+g_3) = 1 - s(1-\theta)(1-g_1). \tag{30}$$

In this case there is a second nontrivial, feasible equilibrium, but, like the second (and ignored) solution to the standard Mendelian dominant case (CROW and KIMURA 1970: 260), it is not biologically realistic with the mutant being all but fixed in the population. All the above results for the DP case hold with the $k = 1$ substitution (see Table 2): ignoring terms in ν^2 and above we obtain the approximations to the remaining equilibrium

$$\hat{g}_1 = 1 - \frac{4\nu}{s(1-\theta)},$$

$$\hat{p} = 1 - \frac{2\nu}{s(1-\theta)}. \tag{31}$$

The equilibrium proportion of imprinted female gametes is given by

$$\hat{x}_1 + \hat{x}_2 = (1-\theta)(\hat{g}_2+\hat{g}_3) \approx \frac{4\nu}{s}. \tag{32}$$

The leading eigenvalue, λ, of the linearized system is

$$\lambda = 1 - \tfrac{1}{2}s(1-\theta), \tag{33}$$

ignoring terms in ν and above.

Case R: M_2 recessive (DR): In this last case we have M_1 dominant over the mutant M_2: no females bearing an M_1 allele imprint their eggs and a proportion $1-\theta$ of offspring of homozygous M_2M_2 females imprint their A genes (see Table 1). All other details are unchanged.

Sperm are produced at the same frequencies as above (Equations 1), and the three types of eggs, $(A)M_2$, AM_1 and AM_2, at respective frequencies of

$$x_2^m = (1-\theta)g_3$$

$$x_3^m = g_1 + \tfrac{1}{2}g_2 \quad \text{and}$$

$$x_4^m = \tfrac{1}{2}g_2 + \theta g_3, \tag{34}$$

which are Equations 21 with $k\theta = 1$ (which in turn implies $x_1 = 0$). The equations for the genotype frequencies in the next generation are Equations 22 with $x_1 = 0$ in which

$$\bar{w} = 1 - s(1-\theta)g_3. \tag{35}$$

Substituting $p = g_1 + \tfrac{1}{2}g_2$ and using the constraint that $\Sigma g_i' = 1$ gives two recursion equations in two variables:

$$p' = p(2 - s(1-\theta)g_3)(1-\nu)/2\bar{w} \quad \text{and}$$

$$g_3' = (1 - p(1-\nu))(1 - p(1-\nu) - s(1-\theta)g_3)/\bar{w}, \tag{36}$$

which are Equations 15 with θ replaced by $1-\theta$. This formal equivalence arises because the proportion of disadvantaged offspring is, respectively, θ and $1-\theta$ in the AR and DR cases. Thus, ignoring a trivial and an unfeasible equilibrium, we obtain the nontrivial feasible equilibrium point

$$\hat{g}_3 = \frac{2\nu}{s(1-\theta)(1+\nu)} \approx \frac{2\nu}{s(1-\theta)},$$

$$\hat{p} = 1 - \sqrt{\frac{2\nu}{s(1-\theta)}} \tag{37}$$

(with, respectively, terms of ν^2 and $\nu^{3/2}$ and above ignored). The equilibrium proportion of imprinted female gametes is given by

$$\hat{x}_2 = (1-\theta)\hat{g}_3 \approx \frac{2\nu}{s}. \tag{38}$$

The leading eigenvalue for the linearization of Equations 36 around Equilibrium 37 is

$$\lambda = 1 - \sqrt{2\nu s(1-\theta)} + O(\nu). \tag{39}$$

As expected, our equilibrium is stable for biologically realistic parts of parameter space.

DISCUSSION

The algebraic results of our models are summarized in Table 2. For cases in which the mutant modifier is

recessive, its equilibrium frequency is given by the square root of twice the mutation rate divided by the product of the selection coefficient and the proportion of disadvantageous gametes (which are produced only by homozygous females). For all of the partially dominant and completely dominant models, the mutant allele's frequency is given by twice the mutation rate divided by the product of the selection coefficient and the proportion of disadvantageous gametes produced by heterozygous females. These results parallel those for mutation-selection balance at standard Mendelian loci (CROW and KIMURA 1970: 259–260): recessive mutants occur at a frequency of $\sqrt{\mu/s}$, partially dominant ones with level of dominance h at μ/hs, and completely dominant ones at μ/s (in which μ is the mutation rate). In all cases, a recessive mutant will be found at a greater frequency (hence the surd) because selection fails to act against heterozygotes. The extra factor of two in the imprinting models arises because imprinted individuals are functionally haploid.

The equilibrium frequencies of the mutant phenotype (unimprinted individuals for Model A and imprinted individuals for Model D) are even simpler: twice the ratio of the mutation rate to the selection coefficient for the recessive cases (AR and DR), and twice that for the dominant ones (AP, AD, DP and DD). These values are twice those of the corresponding standard one-locus Mendelian models (CROW and KIMURA 1970), this factor of two again arising because of the functionally haploid nature of imprinting systems. It is also of interest that neither the penetrance of the modifier (θ for Model A; $1 - \theta$ for Model D) nor its degree of dominance (k) affects the equilibrium phenotype frequency. Thus the observed frequencies of any such imprinting defects in natural populations imply little about the action of the modifier.

The third line of Table 2 shows, as would be expected, that for the biologically relevant parts of parameter space, all these mutation-selection balance equilibria are stable. The biologically feasible equilibria discounted in the AP, AD, DP and DD cases are uninteresting precisely because they are unstable. The trivial (edge) equilibria, at which M_2 is fixed, are also biologically unreasonable because we have ignored back mutation, on the assumption that M_2 will be rare. The stability of edge equilibria to genotypic modifiers-of-imprinting mutants is considered in more detail in SPENCER and WILLIAMS (1996).

One aspect of our results that deserves comment is the mirroring of θ and $k\theta$ in the advantageous model (cases AP, AD and AR) by, respectively, $1 - \theta$ and $1 - k\theta$ in the disadvantageous model (cases DP, DD and DR). Indeed, these latter cases are mathematically equivalent to the former ones, if these substitutions are made. The biological motivation for the latter three cases, however, is much easier since presumably imprinting is disadvantageous at the vast majority of loci because it nullifies many of the benefits of diploidy (CROW and KIMURA 1970; SPENCER and WILLIAMS 1996).

We may also compare our results with other models of mutation-selection balance under imprinting. In LAIRD's (1987) model of the fragile-X syndrome, for instance, the equilibrium frequency of affecteds is $3\mu_T/(1 + s_f)$, in which μ_T is the weighted average of mutation rates in males and females, s_f is the selection coefficient against imprinted females, and imprinted males do not reproduce (SVED and LAIRD 1990). The details of LAIRD's model are considerably different from ours (in part because of sex-linked inheritance), and the predicted frequency is of a very different form and always smaller than our value. The recessive single-locus model for IGF-2 of SPENCER and WILLIAMS (1995) gave an equilibrium of $2\mu/s$, the same form as for our recessive two-locus models, but where μ is the mutation rate at the imprinted locus rather than the modifier locus. These authors noted several mathematical equivalences among biologically distinct models (see also DENNISTON and CROW 1990; PEARCE and SPENCER 1992).

The assumption we make that homozygotes for the wild-type modifier (and also heterozygotes in the AR and DR cases) should imprint all (Model A) or none (Model D) of their eggs says that the wild-type allele has complete penetrance. We can justify this value using the results of SPENCER and WILLIAMS (1996) that when imprinting is directly advantageous, selection will favor modifiers that imprint a greater proportion of their bearers' gametes, but that modifiers imprinting fewer gametes will be selected when imprinting is directly disadvantageous. Thus in a system in which imprinting has a strong selective advantage (or disadvantage), wild-type alleles are likely to be completely penetrant.

One of the most important assumptions of all the models in this paper is that the imprinting status of offspring depends on both of the genes at the mother's modifier locus. SPENCER and WILLIAMS (1996) refer to this sort of imprinting modifier as a genotypic or *trans*-acting modifier. This assumption implies that imprinting occurs at an early stage in gametogenesis, before the second stage of meiosis when the cells become haploid. A conceptually different sort of modifier would act after this point, the occurrence of imprinting of the primary locus depending only on the modifier present in the haploid cell. This sort of modifier, hypothesized by SAPIENZA (1989), can be called a *cis*-acting or gametic modifier (SPENCER and WILLIAMS 1996).

Our models further assume that the population-size effects of genetic drift and inbreeding are negligible. For finite single-locus Mendelian populations, WRIGHT (1937) showed that the frequency of affecteds is likely to be lower than that predicted by "infinite population" models, especially for populations with an effective size of $< \sim 10^4$. Occasionally, however, the effect of finiteness is to increase the frequency because the

standard error of the correction is large. WRIGHT's result will also pertain to our imprinting modifier models.

TONY REEVE has always been ready to discuss different aspects of genomic imprinting, and we are grateful for his enthusiastic contributions. ROBERT POULIN kindly commented on the manuscript. ANDY CLARK and an anonymous reviewer made several helpful suggestions. Mathematica Enhanced Version 2.2 for Windows was used to perform the algebra and calculus. This work was partially supported by the New Zealand Lottery Grants Board Health Sciences Committee and the Division of Sciences, University of Otago.

LITERATURE CITED

BARLOW, D., 1995 Gametic imprinting in mammals. Science **270:** 1610–1613.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Burgess, Minneapolis.

DENNISTON, D., and J. F. CROW, 1990 Alternative fitness models with the same allele frequency dynamics. Genetics **125:** 201–205.

GOLD, J. D., and R. A. PEDERSEN, 1994 Mechanisms of genomic imprinting in mammals. Curr. Top. Dev. Biol. **29:** 227–280.

HALL, J., 1990 Genomic imprinting: review and relevance to human diseases. Am. J. Hum. Genet. **46:** 857–873.

LAIRD, C. D., 1987 Proposed mechanism of inheritance and expression of the human fragile-X syndrome of mental retardation. Genetics **117:** 587–599.

OGAWA, O., M. R. ECCLES, J. SZETO, L. A. MCNOE, K. YUN *et al.*, 1993 Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature **362:** 749–751.

PEARCE, G. P., and H. G. SPENCER, 1992 Population genetic models of genomic imprinting. Genetics **130:** 899–907.

ROUGHGARDEN, J., 1979 *Theory of Population Genetics and Evolutionary Ecology: An Introduction.* MacMillan, New York.

SAPIENZA, C., 1989 Genome imprinting and dominance modification. Ann. N. Y. Acad. Sci. **564:** 24–38.

SPENCER, H. G., and M. J. M. WILLIAMS, 1995 Failure of imprinting of IGF-2: two models of mutation-selection balance. Am. J. Hum. Genet. **56:** 434–437.

SPENCER, H. G., and M. J. M. WILLIAMS, 1996 The evolution of genomic imprinting: two modifier-locus models. Theor. Popul. Biol. (in press).

SVED, J. A., and C. D. LAIRD, 1990 Population genetic consequences of the fragile-X syndrome, based on the X-inactivation imprinting model. Am. J. Hum. Genet. **46:** 443–451.

WRIGHT, S., 1937 The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci. USA **23:** 307–320.

Communicating editor: A. G. CLARK