

Genomic imprinting, sibling solidarity and the logic of collective action

David Haig* and Jon F. Wilkins

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

Genomic imprinting has been proposed to evolve when a gene's expression has fitness consequences for individuals with different coefficients of matrilineal and patrilineal relatedness, especially in the context of competition between offspring for maternal resources. Previous models have focused on pre-emptive hierarchies, where conflict arises with respect to resource allocation between present and future offspring. Here we present a model in which imprinting arises from scramble competition within litters. The model predicts paternal-specific expression of a gene that increases an offspring's fractional share of resources but reduces the size of the resource pool, and maternal-specific expression of a gene with opposite effects. These predictions parallel the observation in economic models that individuals tend to underprovide public goods, and that the magnitude of this shortfall increases with the number of individuals in the group. Maternally derived alleles are more willing than their paternally derived counterparts to contribute to public goods because they have a smaller effective group size.

Keywords: genomic imprinting; public goods; evolutionarily stable strategy; sibling rivalry; scramble competition

1. INTRODUCTION

Models of sibling competition address how a set of sibs—or rather the genes expressed in a set of sibs—divide something of value (such as resources or parental investment) between themselves. These models can be grouped into two broad classes: (i) models of pre-emptive hierarchies, and (ii) models of scramble competition (Mock & Parker 1997). The distinction can be illustrated by an analogy in which the resource to be divided is represented by a milkshake (Haig 1992). In hierarchy models, a queue of offspring wait in line to suck on a single straw. The more milk that is taken by the individual at the head of the queue, the less that is available for other offspring who are yet to take their turn. Such models apply to situations in which offspring are produced sequentially—an offspring yet to be conceived can do little to limit the resources consumed by an older sib—but also to competition within litters, if the most dominant sib is free to choose how much she takes, the next most dominant is free to choose how much of the remainder he takes, and so on. By contrast, in models of scramble competition, each offspring has its own straw and all suck at once. Such models apply to competition within litters when stronger offspring are unable to exclude weaker offspring from access to a resource.

The crucial difference between the two types of model is who bears the cost of increased consumption. If the queue of offspring in a single-straw model contains a random mixture of greedy and abstemious genotypes, members of both classes will suffer when the milkshake runs dry, whereas, if some offspring suck harder than

others in a multi-straw model, the offspring with higher sucking rates increase their intake at the expense of the poor suckers (Haig 1992). Roughly speaking, sibling greed is limited by relatedness in single-straw models, but not in multi-straw models (Metcalf *et al.* 1979).

Single-straw models predict the evolution of genomic imprinting (gene expression specific to parent of origin) at loci that influence an offspring's consumption of maternal resources if an allele's effects when maternally derived are evolutionarily separable from its effects when paternally derived (Haig 1992; Mochizuki *et al.* 1996). That is, whenever mothers have offspring by more than one father, two randomly chosen sibs will be more likely to share alleles inherited from their mother than alleles inherited from their father(s). Therefore, the inclusive-fitness cost of reduced maternal investment in future offspring will be greater for an offspring's maternally derived alleles than for its paternally derived alleles, and alleles of paternal origin will be selected to take a larger share of maternal investment than will alleles of maternal origin.

At first sight, multi-straw models seem inimical to the evolution of genomic imprinting. Any newly arisen allele that increases an offspring's demand is seemingly favoured by natural selection, regardless of the allele's parental origin, because litter-mates with the allele take resources from those without the allele. Conversely, any allele that reduces an offspring's demand, either by producing less of a demand enhancer or more of a demand inhibitor, is disadvantaged when rare because restraint by offspring with the allele merely frees resources for sibs without the allele. For this reason, multi-straw models predict an infinite escalation of demand if increased solicitation is without cost. However,

* Author for correspondence (dhaig@oeb.harvard.edu).

unbounded escalation of demand is prevented if costs of solicitation are included in the models (MacNair & Parker 1979). Such costs create an opportunity for natural selection to act differently on alleles of maternal and paternal origin.

The model presented below (§2) shows that paternally derived alleles will be selected to produce more of a demand factor than maternally derived alleles, if increased demand reduces the amount of resources to be divided between a set of sibs but increases the share of the diminished pool received by offspring that are more demanding. Conversely, a maternally derived allele that reduces an offspring's level of demand, relative to litter-mates without the allele, can be favoured by natural selection if the reduction in demand causes all offspring (including those with reduced demand) to be better off.

2. A MULTI-STRAW MODEL OF SIBLING COMPETITION

An evolutionarily stable strategy (ESS) occurs at a locus when the strategy employed by most alleles in the population cannot be displaced by any alternative strategy that is initially rare (Maynard Smith & Price 1973). Our model will consider two loci: *A*, encoding a demand factor, and *B*, encoding a demand-factor antagonist. Two alleles will be considered at each locus: an established allele possessed in homozygous form by most individuals in the population, and a rare allele possessed by a few heterozygous individuals. The established allele's strategy will constitute an ESS if all rare alternative strategies are associated with lower fitness. For the purposes of investigating genomic imprinting, each allele's strategy will be represented by a two-element vector, the first element of which represents the allele's level of expression when maternally derived and the second element its level of expression when paternally derived. We shall assume that the value of the first element is evolutionarily unconstrained by the value of the second element (and vice versa).

At the demand-factor locus, the established allele *A*₁ has strategy {*x*₁^m, *x*₁^p} and the rare allele *A*₂ has strategy {*x*₂^m, *x*₂^p}. When *A*₂ is rare, *A*₂*A*₂ genotypes are very rare, and will be ignored. *A*₁*A*₁, *A*₂*A*₁, *A*₁*A*₂ offspring (maternally derived alleles listed first) produce amounts *X*₁, *X*_m, *X*_p, where

$$\left. \begin{aligned} X_1 &= x_1^m + x_1^p, \\ X_m &= x_2^m + x_1^p, \\ X_p &= x_1^m + x_2^p, \\ x_1^m, x_1^p, x_2^m, x_2^p &\geq 0. \end{aligned} \right\} \tag{1a}$$

At the antagonist locus, the established allele *B*₁ has strategy {*y*₁^m, *y*₁^p} the rare allele *B*₂ has strategy {*y*₂^m, *y*₂^p} and *B*₁*B*₁, *B*₂*B*₁ and *B*₁*B*₂ offspring produce *Y*₁, *Y*_m and *Y*_p:

$$\left. \begin{aligned} Y_1 &= y_1^m + y_1^p, \\ Y_m &= y_2^m + y_1^p, \\ Y_p &= y_1^m + y_2^p, \\ y_1^m, y_1^p, y_2^m, y_2^p &\geq 0. \end{aligned} \right\} \tag{1b}$$

An offspring's level of demand *Z*_• is assumed to be an increasing function of *X*_• but a decreasing function of *Y*_•

(the subscripted dot is used here, and below, as a dummy that can be replaced in any given equation by either p or m). Specifically,

$$\left. \begin{aligned} Z_{\bullet} &= X_{\bullet} f(Y_{\bullet}), \\ 0 < f(Y_{\bullet}) &\leq 1, \quad f(0) = 1, \\ \partial f(Y_{\bullet}) / \partial Y_{\bullet} &< 0. \end{aligned} \right\} \tag{2}$$

For simplicity, the model will consider 'average' litters containing a rare allele (rather than summing across all possible litter compositions). An average heterozygous carrier of a rare allele belongs to a litter in which he is one of *r*_• heterozygous sibs producing *X*_•, *Y*_•. The other members of the litter are (*n* - *r*_•) homozygous sibs producing *X*₁, *Y*₁ (the model assumes *n* > 1 because scramble competition is absent from 'litters' of a single offspring). All members of the litter are assumed to extract resources from a common pool of size *S*_• that is a decreasing function of *Z*_•, the level of demand of each of the heterozygous offspring:

$$S_{\bullet}(Z_{\bullet}) \geq 0, \quad \frac{\partial S_{\bullet}}{\partial Z_{\bullet}} < 0. \tag{3}$$

Thus, increased production of the demand factor increases an offspring's fractional share of maternal resources but decreases the total amount of resources available to its litter, whereas increased production of the antagonist has opposite effects. Production of the demand factor confers an individual benefit but a shared cost, whereas production of the antagonist confers a shared benefit but an individual cost.

In the language of trait-group selection (Wilson 1977), *r*_m/*n* is the 'average subjective frequency' of heterozygotes who inherit the rare allele from their mother, whereas *r*_p/*n* is the corresponding frequency for heterozygotes who inherit the rare allele from their father. If all litters are sired by a single male, *r*_p = *r*_m = ½(*n* + 1), otherwise *r*_p < *r*_m = ½(*n* + 1).

An offspring's fractional share of *S*_• is assumed to be the ratio of its own level of demand to the aggregate demand of the litter. Thus, the amount of resources acquired by heterozygous offspring considered as a group (*R*_•) is given by

$$R_{\bullet} = \frac{r_{\bullet} Z_{\bullet}}{(n - r_{\bullet}) Z_1 + r_{\bullet} Z_{\bullet}} S_{\bullet}(Z_{\bullet}). \tag{4}$$

Finally, we shall assume that a heterozygous offspring's fitness is a monotonically increasing function of *R*_•. Maxima of fitness will therefore occur when *R*_• is maximized.

Before presenting a formal analysis of ESS conditions, it is worth discussing two qualitative aspects of the model. First, the size of the pool of resources, *S*, is maximal when the demand function, *Z*, is zero for all offspring. *Z* is therefore a measure of the inefficiencies that arise from sibling rivalry. Second, in a model of pure scramble competition, such as this, a rare allele can increase in frequency only if it increases the amount of resources, *R*, obtained by heterozygous offspring. This condition applies to both the demand factor and its antagonist. An allele *B*₂ that increases production of the antagonist will decrease the fractional share of resources received by

heterozygotes. If such an allele is to increase in frequency, this decrease in fractional share must be more than compensated for by an increase in the size of the pool. A similar restriction need not apply in models that include effects on subsequent litters because an increase in inclusive fitness is possible—despite the decrease in individual fitness—if the benefit to future siblings is sufficiently great.

(a) Demand-factor ESS

For purposes of describing the ESS at the demand-factor locus, all alleles at the antagonist locus will be assumed to be B_1 . As a result, all offspring produce the same amount of antagonist, and an offspring's share of resources is proportional to its own level of production of demand factor relative to the aggregate production by its litter. For simplicity, $S_\bullet(X_\bullet)$ is defined to be the size of the pool of resources for a litter containing $(n - r_\bullet)$ members producing X_1 , and r_\bullet members producing X_\bullet . Therefore, heterozygous offspring considered as a group will receive

$$R_\bullet = \frac{r_\bullet X_\bullet}{(n - r_\bullet)X_1 + r_\bullet X_\bullet} S_\bullet(X_\bullet). \tag{5}$$

The best strategic response of A_2 to A_1 will occur at maxima of equation (5). The relevant partial derivatives are

$$\frac{\partial R_\bullet}{\partial x_2^\bullet} = \frac{r_\bullet X_\bullet}{(n - r_\bullet)X_1 + r_\bullet X_\bullet} S'_\bullet(X_\bullet) + \frac{r_\bullet(n - r_\bullet)X_1}{[(n - r_\bullet)X_1 + r_\bullet X_\bullet]^2} S_\bullet(X_\bullet). \tag{6}$$

An infinite population fixed for A_1 cannot be invaded by any allele with a different pattern of expression if $\{x_1^m, x_1^p\}$ is the best response to itself. Therefore, a candidate ESS can be found by evaluating equation (6) when $\{x_2^m, x_2^p\} = \{x_1^m, x_1^p\}$:

$$\left. \frac{\partial R_m}{\partial x_2^m} \right|_{x_m=x_1} = \frac{r_m}{n} \left[S'(X) + (n - r_m) \frac{S(X)}{nX} \right], \tag{7a}$$

$$\left. \frac{\partial R_p}{\partial x_2^p} \right|_{x_p=x_1} = \frac{r_p}{n} \left[S'(X) + (n - r_p) \frac{S(X)}{nX} \right]. \tag{7b}$$

Subscripts are dropped from $S_\bullet(X_\bullet)$ and X_\bullet because $S_m(X_m) = S_p(X_p)$ when $X_p = X_m = X_1$. The term within the square brackets is larger for equation (7b) than for equation (7a), except when $r_m = r_p$ (single paternity). With this one exception, (7a) < 0 when (7b) = 0 and (7b) > 0 when (7a) = 0. Therefore, the maternally derived allele will be silent at an imprinted ESS of the form $\{0, X^*\}$. This is an expression of the 'loudest-voice-prevails' principle (Haig 1996, 1997a).

Put into words rather than equations, both alleles at a locus contribute their products to a common pool. Whenever the combined level of demand factor is greater than the (lower) maternal optimum, maternally derived alleles would benefit from producing less. Conversely, whenever the combined level is less than the (higher) paternal optimum, paternally derived alleles would benefit from producing more. Each increase in paternal production can be matched by a decrease in maternal production, until maternal production reaches zero, at which point no

further reduction is possible. Paternally derived alleles are then free to produce their favoured amount.

In the case of single paternity, $r_m = r_p = \frac{1}{2}(n + 1)$. Therefore, (7a) = (7b) = 0 at an ESS. This implies

$$S'(X) + \frac{1}{2}(n - 1)S(X)/nX = 0. \tag{8}$$

Condition (8) specifies a continuum of ESSs of the form $\{x_m^*, x_p^*\}$, where $x_m^* + x_p^* = X^*$. The unimprinted strategy $\{\frac{1}{2}X^*, \frac{1}{2}X^*\}$ lies at the midpoint of this continuum.

The assumption that an offspring's fractional share of resources is proportional to its value of Z relative to the aggregate Z in its litter ensures that $X > 0$ at an ESS. The evolutionary instability of a population in which the established allele has strategy $\{0, 0\}$ can be seen by considering the effect of an initially rare allele that produced an infinitesimally small amount δX . Such an allele would have little effect on the total amount of resources to be divided, but, in mixed litters, all resources would be obtained by offspring that produced δX and none by offspring that produced zero. This implication is clearly unrealistic. Nevertheless, our model retains the assumption of pro rata shares for reasons of mathematical tractability, and because similar assumptions have been employed in most previous models of scramble competition within litters (Mock & Parker 1997).

(b) Antagonist ESS

For purposes of describing the ESS at the antagonist locus, all alleles at the demand-factor locus will be assumed to be A_1 . In this section, $S_\bullet(Y_\bullet)$ will be defined to be the size of the pool of resources available to a litter containing $(n - r_\bullet)$ members producing Y_1 and r_\bullet members producing Y_\bullet . Therefore, heterozygous offspring as a group will receive

$$R_\bullet = \frac{r_\bullet f(Y_\bullet)}{(n - r_\bullet)f(Y_1) + r_\bullet f(Y_\bullet)} S_\bullet(Y_\bullet). \tag{9}$$

The ESS is obtained by evaluating partial derivatives of equation (9) when $\{y_1^m, y_1^p\} = \{y_2^m, y_2^p\}$:

$$\left. \frac{\partial R_m}{\partial y_2^m} \right|_{r_m=r_1} = \frac{r_m}{n} \left[S'(Y) + (n - r_m) \frac{f'(Y)S(Y)}{nf(Y)} \right], \tag{10a}$$

$$\left. \frac{\partial R_p}{\partial y_2^p} \right|_{r_p=r_1} = \frac{r_p}{n} \left[S'(Y) + (n - r_p) \frac{f'(Y)S(Y)}{nf(Y)} \right]. \tag{10b}$$

Subscripts are dropped from $S_\bullet(Y_\bullet)$ and Y_\bullet because $S_m(Y_m) = S_p(Y_p)$ when $Y_p = Y_m = Y_1$. The term within the square brackets is larger for equation (10a) than equation (10b), except when $r_m = r_p$, because $f'(Y) < 0$. Therefore, maternally derived alleles are predicted to produce their favoured amount of antagonist at an ESS of the form $\{Y^*, 0\}$. For single paternity, $r_m = r_p$, the model predicts a continuum of ESSs of the form $\{y_m^*, y_p^*\}$, where $y_m^* + y_p^* = Y^*$. The unimprinted strategy $\{\frac{1}{2}Y^*, \frac{1}{2}Y^*\}$ lies at the midpoint of this continuum.

A necessary condition for maternally derived alleles to favour non-zero production of the antagonist is that (10a) > 0 when evaluated for $Y = 0$. That is,

$$\frac{S'(0)}{S(0)} > - \left(\frac{n - r_m}{n} \right) \frac{f'(0)}{f(0)} = - \left(\frac{n - 1}{2n} \right) f'(0). \tag{11}$$

The left-hand side of this inequality represents the proportional increase in the pool of resources caused by the first small increment in antagonist. The right-hand side is the corresponding decrease in a B_2B_1 offspring's level of demand multiplied by the offspring's average subjective frequency of B_1B_1 sibs (who produce zero antagonist). For a litter of two, this translates into a requirement that the proportional increase of the pool be greater than a quarter of the offspring's decrease in demand. For large litters, condition (11) specifies that the proportional increase in the pool must exceed half the decrease in demand.

(c) *Competition within and between litters*

The model presented above excludes effects on future offspring to isolate the theoretical consequences of scramble competition. This limits the kinds of gene action that are favoured by natural selection to those that enhance individual fitness. In particular, the model cannot represent gene actions that reduce individual fitness for the benefit of future offspring. For many real organisms, competition within litters can have consequences for future litters, but few models have attempted to combine single- and multi-straw competition, probably because such models are messy and lack simple generalizations.

A model that considers both kinds of interaction has been presented by Haig (1996). In his model, increased production of a nutrient-enhancing hormone results in increased resources available to a litter at the cost of decreased resources available for future siblings. The multi-straw component of the model is particularly simple because all members of a litter release the hormone into a common pool (the maternal circulation) and all obtain an equal share of the resulting resources (from which each member subtracts her own costs of production).

In this model, multiple paternity within litters had opposite effects to changes of paternity between litters. Production of the hormone by offspring was of communal benefit within litters, but had an individual cost. Thus, an increase in the number of fathers *within litters* increased the free-rider problem for paternally derived alleles—but not for maternally derived alleles—and consequently reduced the level of hormone favoured by paternally derived alleles. (This effect is present in the current model and accounts for the higher level of antagonist favoured by maternally derived alleles.) However, increased resources obtained by the current litter occurred at the expense of future litters. As a consequence, increased turnover of fathers *between litters* caused paternally derived alleles to favour higher levels of hormone production but had no effect on the level favoured by maternally derived alleles. (Effects on future offspring are absent from the current model.)

The placental-hormone model did not predict imprinted expression of the hormone by maternally derived alleles of offspring for any combination of multiple paternity within litters and changes of paternity between litters because it included the possibility that alleles expressed in the mother (as well as in offspring) could release hormone into the common pool. As a result, whenever maternally derived alleles of offspring favoured

greater production than paternally derived alleles, alleles expressed in the mother favoured even higher levels and took over all production. However, if the possibility of production by the mother had been excluded, the model would have predicted imprinted expression of the hormone by either maternally derived or paternally derived alleles, depending on details of the mating system.

The placental-hormone model illustrates some of the complexities of the interaction between competition within and between litters, but is not a general treatment of this problem. Other possibilities that remain to be formally modelled include a demand factor that increases an offspring's share at the expense of current and future offspring, or a demand-factor antagonist that takes from the current litter to give to future litters.

3. PUBLIC GOODS AND COLLECTIVE ACTION

The model of scramble competition presented in this paper can be interpreted as an illustration of the economic principle that groups tend to underprovide public goods because a public good (by definition) is available to all, whether or not an individual contributes to its provision (Olson 1961). From this perspective, the pool of resources shared by the members of a litter is a public good that is underprovided at evolutionary equilibrium. The size of the pool would be larger, and each sib would receive more, if all sibs demanded less, but unilateral restraint by a subset of sibs would reduce their own share for the benefit of sibs who do not show restraint. Such inefficiencies are a general feature of models of sib competition.

In contrast with models in which genes expressed in offspring determine the distribution of a collective good among sibs, efficient outcomes are predicted if genes expressed in a shared parent determine the distribution of the good. This is because the effect on offspring of an allele expressed in a parent is (usually) independent of whether the offspring inherits the allele. That is, if genes in parents are uninformed about which offspring inherit which alleles, the best they can do is to maximize the combined fitness of offspring considered as a group (cf. Harsanyi 1953). The extra information available to genes in offspring—that an allele is definitely present in the offspring in which it is expressed—prevents these genes from achieving an efficient outcome. It is not always better to be better informed.

In *The logic of collective action*, Olson (1961) argued that the larger a group, the greater the shortfall from optimal provision of public goods, other things being equal. If the analogy between sib competition and economic models is to be extended, one must ask who or what in the multi-straw model takes the place of the rational self-interested individual of economics. The unit of strategic innovation in our model is not an individual sib—or even an individual copy of a gene within a sib—but rather an 'allele', in the collective sense of all identical-by-descent copies of a DNA sequence within a litter (compare the discussion of the 'strategic gene' in Haig (1997b)). By this definition, the expected number of maternally derived alleles in a litter (two) is smaller than the expected number of paternally derived alleles (more than two if there is some

possibility of multiple paternity). That is, maternally derived alleles comprise a smaller group than paternally derived alleles and are therefore predicted to contribute more to the provision of public goods. Conversely, paternally derived alleles are predicted to invest more in the acquisition of selfish benefits that reduce the supply of public goods.

Multiple paternity within litters is not the only reason why alleles of different parental origin might interact in groups of different sizes. For example, consider a species that forms social groups of matrilineal female kin whose reproduction is monopolized (temporarily) by each of a series of unrelated males. Offspring born during the reproductive tenure of any given male would typically possess fewer paternally derived alleles than maternally derived alleles, especially at X-linked loci. If so, one might expect the expression of alleles of paternal origin to favour greater cooperation between individuals of similar age than would the expression of alleles of maternal origin (Trivers & Burt 1999; Haig 2000).

Robert Trivers and two anonymous referees have commented helpfully on the manuscript.

REFERENCES

- DeChiara, T. M., Robertson, E. J. & Efstratiadis, A. 1991 Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64**, 849–859.
- Haig, D. 1992 Genomic imprinting and the theory of parent-offspring conflict. *Semin. Dev. Biol.* **3**, 153–160.
- Haig, D. 1996 Placental hormones, genomic imprinting, and maternal-fetal communication. *J. Evol. Biol.* **9**, 357–380.
- Haig, D. 1997a Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proc. R. Soc. Lond. B* **264**, 1657–1662.
- Haig, D. 1997b The social gene. In *Behavioural ecology*, 4th edn (ed. J. R. Krebs & N. B. Davies), pp. 284–304. Oxford, UK: Blackwell Scientific.
- Haig, D. 2000 Genomic imprinting, sex-biased dispersal, and social behavior. *Annals NY Acad. Sci.* **907**, 149–163.
- Harsanyi, J. C. 1953 Cardinal utility in welfare economics and in the theory of risk-taking. *J. Polit. Econ.* **61**, 434–435.
- Macnair, M. R. & Parker, G. A. 1979 Models of parent-offspring conflict. III. Intra-brood conflict. *Anim. Behav.* **27**, 1202–1209.
- Maynard Smith, J. & Price, G. R. 1973 The logic of animal conflict. *Nature* **246**, 15–18.
- Metcalf, R. A., Stamps, J. A. & Krishnan, V. V. 1979 Parent-offspring conflict that is not limited by degree of kinship. *J. Theor. Biol.* **76**, 99–107.
- Mochizuki, A., Takeda, Y. & Iwasa, Y. 1996 The evolution of genomic imprinting. *Genetics* **144**, 1283–1295.
- Mock, D. W. & Parker, G. A. 1997 *The evolution of sibling rivalry*. Oxford University Press.
- Olson, M. 1961 *The logic of collective action. Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Trivers, R. & Burt, A. 1999 Kinship and genomic imprinting. In *Genomic imprinting. An interdisciplinary approach* (ed. R. Ohlsson), pp. 1–21. Berlin: Springer.
- Wilson, D. S. 1977 Structured demes and the evolution of group-advantageous traits. *Am. Nat.* **111**, 157–185.