doi:10.1111/j.1558-5646.2010.01147.x

GENOMIC IMPRINTING AND CONFLICT-INDUCED DECANALIZATION

Jon F. Wilkins[1,2]

[1]*Santa Fe Institute, Santa Fe, New Mexico 87501*

[2]*E-mail: wilkins@santafe.edu*

Received January 26, 2010
Accepted August 20, 2010

Genomic imprinting is the phenomenon in which the expression pattern of an allele depends on its parental origin. When maternally expressed and paternally expressed imprinted loci affect the same trait, the result is an arms race, with each locus under selection to increase its level of expression. This article develops a model of the deleterious consequences of this escalation, deriving from an increase in the variance in gene expression level, and resulting increase in phenotypic variance in the population. This phenomenon is referred to here as "conflict-induced decanalization." Modifiers that canalize gene expression are selectively favored, but these induce further escalation from both loci, resulting in a net increase in phenotypic variance and a reduction in population mean fitness. This results in a feedback loop, where increasing canalization of gene expression leads to increasing decanalization of the phenotype. This phenomenon may explain the surprisingly high frequency of certain diseases. Disorders to which this decanalization process might contribute include growth- and metabolism-related phenomena such as preterm birth, as well as certain major psychiatric disorders, including schizophrenia and autism.

KEY WORDS: **Autism, canalization, genetic conflict, gene expression, schizophrenia.**

The term "canalization" refers to mechanisms that reduce the extent of phenotypic variation in the face of underlying genotypic or environmental variation (Waddington 1942, 1959; Siegal and Bergman 2002). Canalization provides benefits to the individual organism, and it is easy to imagine how natural selection might favor traits that enhance canalization. In fact, a number of specific canalization mechanisms have been proposed (Queitsch et al. 2002; Flatt 2005; Sangster et al. 2008a, b). Canalization mechanisms may also impact the nature and rate of adaptive evolution in response to environmental changes: cryptic genetic variation accumulates under conditions in which that variation is phenotypically or selectively neutral (Masel and Siegal 2009). Following an environmental change, this genetic variation may manifest in phenotypic differences that can be acted on by natural selection.

The evolutionary origin of mechanisms that reduce phenotypic variance is a matter of some controversy, however. One debate is over what types of canalization could be acted on directly by selection, and which types are better thought of as epiphenomena, or side effects of other evolutionary processes (Masel and Siegal 2009). For instance, it has been argued that natural selection can act directly on features that buffer environmental variation, and that these features will then buffer genetic variation as a byproduct (Meiklejohn and Hartl 2002). Other results have questioned whether specifically evolved canalization mechanisms exist at all, or if all observed canalization is merely a consequence of the system of complex interactions underlying the construction of the phenotype (Siegal and Bergman 2002; Bergman and Siegal 2003).

Whatever the evolutionary origin of the structural features and/or specific mechanisms that account for the canalization observed in existing organisms, there is little doubt that canalizing features impact the distribution of phenotypes in natural populations and the response of those populations to specific selective pressures. It is therefore of interest to understand evolutionary processes that may influence the extent to which phenotypes are canalized.

This article presents a simple model in which antagonistic coevolution between a pair of genes, each of which is subject

to genomic imprinting, results in an increase in the phenotypic variance. I will refer to this phenomenon as 'conflict-induced decanalization". It has been suggested that stochastic variation may contribute substantially to many human diseases (Feinberg and Irizarry 2010), and that decanalization may contribute specifically to certain complex diseases, including diabetes, as well as immune and psychological disorders (Gibson 2009). After using this simple model to illustrate the principle of conflict-induced decanalization, I discuss briefly how this phenomenon might contribute to the high frequencies of certain growth and metabolism-related disorders, such as preterm birth, as well as certain psychiatric disorders, such as schizophrenia and autism.

Antagonistic Coevolution of Imprinted Genes

Genomic imprinting refers to the phenomenon in which the expression of an allele depends on its parent of origin (Wilkins 2008). Typically, imprinting results in monoallelic expression, with either the maternally derived or paternally derived allele being transcriptionally silenced. The most successful theory of the evolutionary origins of this phenomenon is the Kinship Theory of Imprinting (Haig 2000; Wilkins and Haig 2003b), which is based on the fact that the inclusive fitness of a maternally derived allele is not identical to that of a paternally derived allele.

Much of the empirical and theoretical work on imprinted genes expression in mammals has focused on the role played by these genes in early growth and development. In the simplest models, the fitness asymmetry between maternally and paternally derived alleles stems from the fact that offspring play an active role in soliciting resources from the mother. Gene expression in fetal and placental tissues during pregnancy can influence the distribution of maternal resources among her offspring (both among litters, and, for multiparous species, within litters). Natural selection acting on an allele favors the level of resource demand that maximizes the allele's inclusive fitness. Due to the fact that not all of a female's offspring necessarily have the same father, selection favors alleles that place a greater demand when paternally derived than when maternally derived.

In models of a single imprinted locus, the allele favoring a lower level of overall expression becomes silenced, and expression from the allele favoring higher overall production of the gene product evolves to the level that maximizes that allele's inclusive fitness (Mochizuki et al. 1996). This evolutionarily stable state (ESS) has been dubbed the "loudest voice prevails" principle (Haig 1996). When two imprinted loci have opposite effects on the same phenotype, and intragenomic conflict over the phenotype value is responsible for imprinting at these loci, modeling predicts that one locus will be maternally expressed whereas the other is paternally expressed (Kondoh and Higashi 2000; Wilkins and Haig 2001). Thus, the simple prediction is that imprinted genes that enhance prenatal growth (concomitant with an increased demand on maternal resources) will be paternally expressed and maternally silenced, whereas growth-suppressing loci will be maternally expressed and paternally silenced.

Although prenatal growth is the most familiar and best-studied arena in which intragenomic conflict can give rise to imprinted gene expression, the Kinship Theory applies to any context in which natural selection acts differently on maternally and paternally derived alleles. In fact, imprinted gene expression can also be found in a variety of adult tissues, particularly the brain, where these genes affect various cognitive and behavioral phenotypes (Goos and Silverman 2006; Wilkinson et al. 2007; Davies et al. 2008; Goos and Ragsdale 2008). Many recent studies have focused on understanding the origins of these imprinted gene affects in (post-natal) juveniles and adults, often through extensions or generalizations of the prenatal models in which genes influence the distribution of parental resources. These models have demonstrated how evolutionary outcomes can be affected by paternal resource provisioning, sex-specific migration, and details of the reproductive structure (Úbeda 2008; Wild and West 2009; Brandvain 2010; Úbeda and Gardner 2010; Van Cleve et al. 2010).

Other adult phenotypes associated with imprinted gene expression may not lend themselves to straightforward extensions of the idea that genes in the offspring are affecting the distribution of resources coming from the parents. However, to the extent that the imprinted gene expression associated with these phenotypes is shaped by asymmetric consequences of the phenotype for matrilineal and patrilineal inclusive fitness, the Kinship Theory may still prove a useful framework for understanding the evolution of these genes, and the consequences of that evolution on observable distributions of organismal phenotypes.

The work presented here is not concerned with the evolutionary origins of imprinted gene expression. Rather, it focuses on the consequences of imprinted gene expression for the evolution of the phenotype. I focus on a system of two imprinted genes: one maternally expressed and one paternally expressed, and assume that increasing gene expression from these two loci has opposite effects on a particular aspect of phenotype. Furthermore, I assume that the phenotype value that maximizes the matrilineal inclusive fitness differs from the value that maximizes patrilineal inclusive fitness.

As noted above, in the case of a single imprinted locus, the evolutionarily stable expression pattern is one that maximizes the inclusive fitness of the active allele (the allele for which natural selection favors higher expression). However, the ESS for a system of two oppositely imprinted genes is not immediately obvious. The simplest possible analysis predicts an arms race between the

two loci, with each allele selected to increase its level of expression without limit. Clearly, infinite escalation is not biologically realistic, and at some point other factors or processes will limit this arms race in expression level.

One possibility is that this escalation terminates when one of the two loci reaches a mechanistic limit beyond which gene expression cannot be elevated. In this case, the ESS will be the one that maximizes the fitness of alleles at the locus that does not reach its mechanistic limit first. For example, if gene expression from the maternally expressed locus reaches its intrinsic maximum, expression from the paternally expressed locus will increase to that point where the patrilineal fitness is locally maximized.

In this article, I focus on another set of effects that can determine the ESS of the escalating, two-locus system, where escalation terminates as a result of (1) diminishing returns in the phenotypic effects of the gene products and/or (2) costs associated with increasing gene expression. These costs could come in the form of the energetic costs of protein production, or, more likely, fitness costs associated with deleterious side effects of escalating gene expression. In a previous paper, I have examined one particular form of fitness cost, resulting from pleiotropic effects of the imprinted genes (Wilkins 2009). In that model, escalation is driven by intragenomic conflict over one aspect of the phenotype, and that escalation drives other aspects of the phenotype away from their optimal values.

Here, I focus on a different fitness cost associated with increased gene expression: stochastic variation around the optimal phenotype. This model assumes that increased expression from a locus is associated with increased variance in gene expression among individuals. The empirical data supporting this assumption are described, as are the evolutionary consequences if the assumption does not hold. The analysis finds that the increase in expression variance can bring an end to the antagonistic coevolutionary process, and that the resulting ESS will often be associated with increased phenotypic variance and reduced population mean fitness.

The Model

I consider two loci that exhibit opposite effects on a single aspect of phenotype, represented by the scalar value ϕ. Let X represent the total level of expression from the first locus, which is the sum of the expression levels x_m and x_p from the maternally and paternally derived alleles, respectively. Similarly, $Y = y_m + y_p$ is the expression from the second locus. Without loss of generality, assume that increased expression from the first locus increases the phenotype value, whereas increased expression from the second locus decreases it, that is,

$$\frac{\partial \phi}{\partial X} > 0, \quad \frac{\partial \phi}{\partial Y} < 0. \tag{1}$$

Also without loss of generality, I assume that the phenotype value that maximizes the patrilineal fitness is greater than the value maximizing the matrilineal fitness. Previous analyses have shown that this model results in imprinted expression from both loci, with the first locus being maternally silenced ($x_m = 0$, $X = x_p$), and the second being paternally silenced ($y_p = 0$, $Y = y_m$) (Kondoh and Higashi 2000; Wilkins and Haig 2001). Phenotype values can be rescaled such that the phenotype maximizing the patrilineal fitness is $\phi = \alpha$, and the value maximizing the matrilineal fitness is $\phi = -\alpha$. For example, much of the work on imprinted genes has focused on the phenotype of birth weight, where the paternal optimum is greater than the maternal optimum. In this context, the phenotype value ϕ would be a rescaled representation of the birth weight, where $\phi = 0$ is defined to be the birth weight at the midpoint between the two optima.

For phenotype values close to these optima, the patrilineal and matrilineal fitness associated with a particular organism (w_p and w_m, respectively) can be approximated using a Taylor's series expansion. This series expansion approach has the advantage that the results derived here will apply to many different fitness functions, so long as the fitness functions have a single, smooth peak. The ensuing caveat is that the approximate solutions derived below will be less valid for phenotype values far from these optima, where the results of a more exact analysis would depend on details of the exact shape(s) of the fitness functions. For simplicity, I will assume that these two fitness functions have the same fitness maximum ($w_{p\,max} = w_{m\,max} = 1$) and the same local curvature (i.e., that $\partial^2 w_p/\partial\phi^2 = \partial^2 w_m/\partial\phi^2 < 0$). It is then possible, without further loss of generality, to rescale the phenotype values such that $\partial^2 w_p/\partial\phi^2 = \partial^2 w_m/\partial\phi^2 = -2$. The approximate fitness are then given by

$$w_p \approx 1 - (\phi - \alpha)^2 \quad \text{and} \tag{2}$$

$$w_m \approx 1 - (\phi + \alpha)^2 . \tag{3}$$

This fitness model is illustrated in Figure 1.

Note that this rescaling implies a particular interpretation of the conflict parameter α as the magnitude of the conflict relative to the strength of stabilizing selection. For example, if a particular phenotype is under very strong stabilizing viability selection, meaning that there is only a narrow range of viable phenotype values, this rescaling will result in a very small value of α. Put another way, the matrilineal and patrilineal fitness functions used here contain direct viability effects as well as indirect effects (e.g., through social interactions or effects on the mother's residual reproductive value). The larger those direct viability effects (in relation to whatever indirect effects may exist), the more similar the matrilineal and patrilineal fitness functions will be, and the smaller the magnitude of α.

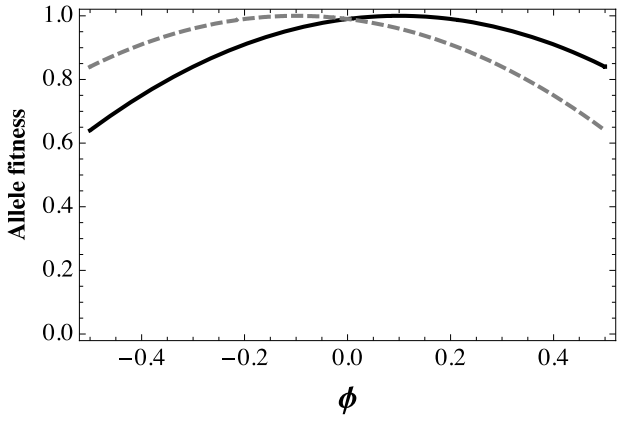

Figure 1. The matrilineal and patrilineal fitness functions are indicated with dashed and solid lines, respectively. Matrilineal fitness is maximized when the phenotype value equals $-\alpha(= -0.1$ in this case), whereas the patrilineal fitness is maximized at $\alpha(= 0.1)$. Increasing expression from the *X* locus increases the phenotype value (a shift to the right), whereas increasing expression from *Y* reduces the phenotype value.

Expressions (2) and (3) give the inclusive fitness for a paternally or maternally derived allele in a single individual with phenotype value ϕ. The overall fitness of an allele will be equal to its average inclusive fitness across all individuals, and therefore depends on the distribution of phenotypes associated with that allele:

$$\langle w_p \rangle \approx 1 - \int f(\phi)\,(\phi - \alpha)^2 \quad \text{and} \tag{4}$$

$$\langle w_m \rangle \approx 1 - \int f(\phi)\,(\phi + \alpha)^2, \tag{5}$$

where $f(\phi)$ is the probability that an organism containing the allele has phenotype value ϕ. These fitness functions can be rewritten in terms of the mean and variance of the phenotype distribution function $f(\phi)$, denoted by ϕ_0 and σ^2_ϕ, respectively.

$$\langle w_p \rangle \approx 1 - \left(\sigma^2_\phi + (\phi_0 - \alpha)^2\right) \quad \text{and} \tag{6}$$

$$\langle w_m \rangle \approx 1 - \left(\sigma^2_\phi + (\phi_0 + \alpha)^2\right). \tag{7}$$

The analysis presented here will focus on evolutionary stability with respect to changes in gene expression level. For purposes of this analysis, an ESS will be defined as a pattern of expression for which a small change in any of the four expression levels (x_m, x_p, y_m, and y_p) results in a decrease in the relevant allelic fitness. That is, a small change in x_m or y_m decreases the matrilineal fitness w_m, and a small change in x_p or y_p decreases the patrilineal fitness w_p. A detailed mathematical description of these ESS conditions is provided in the Appendix.

By considering the general ESS conditions in the context of the fitness relationships described above, it is possible to draw certain general conclusions. These conclusions do not depend on a specific model of the interaction between the two gene products, nor do they make any assumptions regarding the relationship between the mean gene expression level associated with an allele and the phenotypic variance among individuals carrying that allele. However, it is difficult to provide a clear biological interpretation for these very general conclusions, which have therefore been relegated to the Appendix. The rest of the article will focus on a more concrete model of these interactions and relationships, to present results to which biological interpretations can more easily be attached.

Noise Susceptivity

The evolutionary stability of gene expression levels depends on the derivatives of the fitness functions presented in (6) and (7) with respect to the expression levels X and Y. It is useful to define a quantity that relates the change in the mean phenotype value among individuals carrying an allele to changes in the phenotypic variance among those individuals. I therefore define the "noise susceptivity" of the phenotype with respect to the X or Y locus as

$$\nu_x \equiv \frac{\partial \sigma^2_\phi / \partial X}{\partial \phi_0 / \partial X} \quad \text{and} \tag{8}$$

$$\nu_x \equiv \frac{\partial \sigma^2_\phi / \partial Y}{\partial \phi_0 / \partial Y}. \tag{9}$$

In many contexts, it will be reasonable to expect that the phenotypic variance function $\sigma^2_\phi(X, Y)$ is monotonically increasing in both X and Y (see below), in which case both ν_X and ν_Y will be greater than zero. A negative noise susceptivity would result from a situation in which the absolute phenotypic variance actually decreases with increasing expression level from a gene. If both noise susceptivities are negative, the escalation process will never lead to an ESS within this model, meaning that escalation will only halt through other mechanisms.

One of the conclusions from the general model (see appendix) is that greater noise susceptivity of the X–Y system (larger values of ν_X and ν_Y) will favor ESS conditions in which one of the two loci is transcriptionally silenced, whereas a more intense conflict (a larger value of α) will favor solutions in which both loci are expressed. If both ν_X and ν_Y are negative, both X and Y should have nonzero expression even in the absence of any genetic conflict (when $\alpha = 0$).

In the specific case in which both loci have nonzero expression at the ESS, the equilibrium mean phenotype value can be written in terms of the noise susceptivities

$$\phi_0 = \frac{\nu_Y - \nu_X}{4}. \quad (10)$$

The evolutionarily stable mean phenotype value ϕ_0 can be interpreted in some sense as a measure of which of the two alleles "wins" in the intragenomic conflict. Note that this value will be closer to the optimum of the (maternally or paternally) expressed allele at the locus with the lower noise susceptivity. For example, if $\nu_X = 0$ and $\nu_Y > 0$, the equilibrium phenotype value will occur at $\phi_0 = \alpha$, the patrilineal optimum.

This observation suggests a principle for the coevolution of two antagonistic imprinted loci that is analogous to the "loudest voice prevails" principle for a single imprinted gene—the "least susceptive voice prevails" principle. Note, however that this conflict is at risk of overshoot. If increased expression from one of the two loci actually has the effect of reducing phenotypic variance (i.e., that locus has a negative noise susceptivity), the evolutionarily stable phenotype value (if it exists for nonzero expression from both loci) will lie outside the range defined by the matrilineal and patrilineal optima. More explicitly,

$$\phi_0 < \alpha \quad \textit{iff} \quad \nu_X > 0 \quad \text{and} \quad (11)$$

$$\phi_0 > -\alpha \quad \textit{iff} \quad \nu_Y > 0. \quad (12)$$

If both noise susceptivities are positive, the equilibrium phenotype value will be bounded between the matrilineal and patrilineal optima ($-\alpha < \phi_0 < \alpha$). If both are negative, there will be no ESS for the system, and both loci will be under selection to increase expression level without bound. As noted in the introduction, this boundless increase should not be taken as an empirical prediction, as some other mechanistic or selective factor will eventually bring this escalation to a halt (e.g., pleiotropic effects of the genes (Wilkins 2009)). However, if ν_X and ν_Y are both less than zero, phenotypic decanalization will not serve to stop the escalation.

Cost of Imprinting

One quantity that will be of interest in the analysis that follows is the fitness cost at evolutionary equilibrium. Because alleles are, on average, maternally and paternally derived each half of the time, we can combine equations (6) and (7) to find the mean fitness at equilibrium

$$\langle w \rangle = 1 - \alpha^2 - \phi_0^2 - \sigma_\phi^2. \quad (13)$$

Equation (13) makes it easy to understand three distinct fitness consequences of interlocus, intragenomic conflict. The α^2 term is a fixed cost resulting from whatever environmental and/or demographic factors are responsible for creating the asymmetric selection on maternally and paternally inherited alleles. The ϕ_0^2 term indicates the fitness cost when the population average phenotype deviates from the average of the matrilineal and patrilineal optima (recall that these two optima were at $-\alpha$ and α, respectively, so that their average is zero). Finally, the σ_ϕ^2 term represents the fitness cost due to phenotypic variation around the population mean.

In the analysis that follows, I partition σ_ϕ^2 term into three parts, which are assumed to be additive: the parts of the phenotypic variance resulting from stochastic variation in gene expression from X and Y, and σ_0^2, which includes all other sources of variation in the phenotype, including genetic or transcriptional variation at other loci, as well as environmental variation. Note that the α^2 term and the σ_0^2 contribution to the σ_ϕ^2 term represent fixed fitness costs that do not depend on the expression levels X and Y, and do not depend on whether those loci are imprinted. Thus, although the maximum fitness for a given allele is 1, the maximum mean fitness in the population is only $1 - \alpha^2 - \sigma_0^2$. I will define the fitness cost for the system of imprinted loci relative to this maximum as ξ_O, which is given by

$$\xi_O = \phi_0^2 + \sigma_\phi^2 - \sigma_0^2. \quad (14)$$

I will use ξ_U to refer to the analogous fitness cost when both loci are unimprinted, calculated in the same way as ξ_O. The portion of the fitness reduction attributable to imprinting, the "imprinting load," can then be defined as $\xi_I = \xi_O - \xi_U$.

It is instructive at this point to pause and consider the effect of natural selection on unimprinted loci. In the absence of genomic imprinting, alleles will be under selection to maximize their average fitness. Equivalently, they will be selected to minimize the fitness cost ξ_U. Natural selection will drive ϕ_0 toward zero, and σ_ϕ^2 will be minimized to the extent possible. The unimprinted optimum is therefore one that maximizes canalization around the optimal mean phenotype.

In contrast, the evolutionarily stable outcome for two antagonistic imprinted genes will not necessarily minimize the phenotypic variance σ_ϕ^2. Furthermore, to the extent that one of the two loci "wins," the mean phenotype will deviate from zero, introducing an additional cost. To draw more specific conclusions, it is necessary to develop a specific model for the noise susceptivity function of a locus, as well as a more completely specified model of how the gene expression levels from the two loci influence the phenotype. These specifications will be developed in the following sections.

Variance in Phenotype and Gene Expression Level

Assume that the phenotype of interest is a function of the realized expression levels from each of the two loci ($\phi = \phi\ (X,Y)$). The

mean and variance of the phenotype distribution can then be approximated in terms of the means (x_0 and y_0) and variances (σ_x^2 and σ_y^2) of the two gene expression levels using a Taylor expansion

$$\phi_0 = \phi(x_0, y_0) + \frac{\sigma_x^2}{2}\frac{\partial^2\phi}{\partial x^2}\bigg|_{x_0,y_0} + \frac{\sigma_y^2}{2}\frac{\partial^2\phi}{\partial y^2}\bigg|_{x_0,y_0} \quad (15)$$

$$\sigma_\phi^2 = \sigma_0^2 + \sigma_x^2\left(\frac{\partial\phi}{\partial x}\bigg|_{x_0,\,y_0}\right)^2 + \sigma_y^2\left(\frac{\partial\phi}{\partial y}\bigg|_{x_0,\,y_0}\right)^2, \quad (16)$$

where σ_0^2 incorporates all other sources of phenotypic variation.

The relationship between mean expression level and expression variance is fundamentally an empirical question. Despite the wealth of recent quantitative data on gene expression levels, however, few studies have addressed this relationship directly. One study of an inducible green fluorescent protein construct in *Bacillus subtilis* (Ozbudak et al. 2002) suggests that the noise strength is greater than Poisson, with the variance increasing linearly with the mean (i.e., $\sigma_z^2 \propto (1 + c)z_0$). A more recent flow-cytometry study (Newman et al. 2006) measured the distribution of protein levels across single cells in *Saccharomyces cerevisiae*, finding that the coefficient of variation was relatively constant across four orders of magnitude of protein abundance (i.e., $\sigma_z^2 \propto z_0^2$).

To complement these results, I have reanalyzed published data on population-level variation in human lymphoblast cells (Spielman et al. 2007). The relationship between expression mean and variance is illustrated in Figure 2. For each locus, the interindividual variance in gene expression level was estimated by subtracting the variance among replicates from the variance in the entire dataset. The relationship between log mean and log variance was approximately linear, and the slope was estimated by two methods appropriate to cases in which there is noise in both regressed quantities (log mean and log variance in this case). The two methods gave similar results: slope = 1.992; 95% CI = (1.973, 2.012) for major axis regression, and slope = 1.889; 95% CI = (1.878, 1.900) for standardized major axis regression. Slopes and confidence intervals were calculated as described in Warton et al. (2006). This analysis reveals a relationship similar to that found in yeast.

Taken together, these results suggest a simple model in which the variance in gene expression level is linearly proportional to some power of the mean. Letting z stand for either x or y, this model can be written as

$$\sigma_z^2 = \beta_z z_0^2 \quad \text{and} \quad (17)$$

$$\frac{\partial\sigma_z^2}{\partial z_0} = 2\beta_z z_0, \quad (18)$$

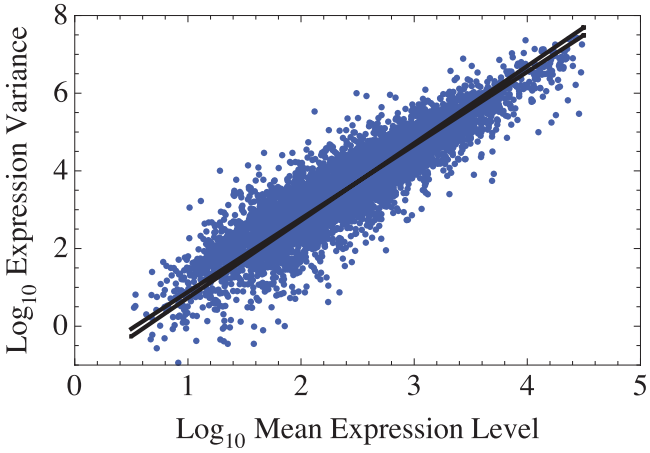

Figure 2. This figure plots the relationship between the mean level of gene expression and the variance in gene expression. Data come from Affymetrix arrays containing approximately 8500 annotated genes, applied to lymphoblast cells from 142 individuals. The original analysis was published by Spielman et al. (2007), and the data are drawn from the Gene Expression Omnibus (GSE5859). The lines indicate the best linear fits based on major axis regression and standardized major axis regression, as described in the text.

where β_z is a constant coefficient relating the variance to the square of the mean. Interestingly, in both the yeast flow-cytometry study and in the study of variation in human gene expression levels, this coefficient is of the order 10^{-1}.

Although the expressions in (15) and (16) are generally applicable, at least for phenotypes that are sufficiently smoothly varying functions of the two gene expression levels, it will be difficult to generate significant further insights without a commitment both to a specific model relating gene expression level to phenotype and to a relationship between the mean and variance in gene expression. Therefore, in the following sections, where a simple, linear model of gene interaction is analyzed, solutions will be presented in terms of the noise model embodied in (17) and (18).

This form of the noise function results in the analytically simple expressions presented below, but the results hold qualitatively for any model in which the expression variance is a greater than linear function of the mean ($\sigma_z^2 \sim z^\gamma$, where $\gamma > 1$), or, equivalently, any model for which the noise susceptivity is an increasing function of gene expression ($\partial\nu_z/\partial z > 0$). It is also worth noting that the use of expressions (15) and (16) does not imply an assumption that expression noise is normally distributed. Because the fitness expressions used here are based on series expansions around the fitness maxima, only the first two moments of the distributions of expression levels matter, rather than the full shape of the distribution.

Linear Gene Effects on Phenotype

To derive explicit conclusions regarding the decanalizing effects of antagonistic coevolution between imprinted loci, I will consider a model in which the phenotype depends linearly on each of the expression levels. Although typical gene interactions are not well described by this simple, linear model, this construction leads to certain qualitative conclusions that will carry over to more complex systems. The issue of nonlinear effects of gene expression level on the phenotype is addressed explicitly in the next section.

Explicitly, the linear effects of gene expression from the two loci on the phenotype are modeled here as

$$\phi = \eta_b + \eta_x X - \eta_y Y, \tag{19}$$

where η_x and η_y are both positive. The term η_b indicates the phenotype value in the absence of expression from either locus, and will be referred to as the baseline phenotype. The baseline phenotype is assumed to have mean η_0 and variance σ_0^2, where that variance comes from genetic polymorphism and/or stochastic variation in gene expression at other loci, as well as from environmental variation. Under this model, the mean and variance of the phenotype distribution are simply

$$\phi_0 = \eta_0 + \eta_x x_0 - \eta_y y_0 \quad \text{and} \tag{20}$$

$$\sigma_\phi^2 = \sigma_0^2 + \eta_x^2 \sigma_x^2 + \eta_y^2 \sigma_y^2 = \sigma_0^2 + \eta_x^2 \beta_x x_0^2 + \eta_y^2 \beta_y y_0^2. \tag{21}$$

The noise susceptivities in this system are

$$\nu_x = 2\eta_x \beta_x x_0 \tag{22}$$

$$\nu_y = 2\eta_y \beta_y y_0. \tag{23}$$

Substituting expressions (20) and (21) into the fitness functions (6) and (7), and taking the relevant first and second derivatives yields

$$\frac{\partial \langle w_p \rangle}{\partial x_0} = 2\eta_x(\alpha - \eta_0 - \eta_x x_0 + \eta_y y_0 - \eta_x \beta_x x_0) \quad \text{and} \tag{24}$$

$$\frac{\partial \langle w_p \rangle}{\partial x_0} = 2\,\eta_y\,(\alpha + \eta_0 + \eta_x\, x_0 - \eta_y\, y_0 - \eta_y\, \beta_y\, y_0) \tag{25}$$

$$\frac{\partial^2 \langle w_p \rangle}{\partial x_0^2} = -2\eta_x^2(1 + \beta_x) \quad \text{and} \tag{26}$$

$$\frac{\partial^2 \langle w_m \rangle}{\partial y_0^2} = -2\eta_y^2(1 + \beta_y). \tag{27}$$

Note that both second derivatives (26) and (27) are always negative, guaranteeing that equilibrium states in which the first derivatives (24) and (25) are equal to zero are both stable and unique.

Evolutionary stability of complete silencing of both loci would require both first derivatives to be negative when $x_0 = y_0 = 0$. Because $\alpha > 0$, this condition is never satisfied simultaneously for both loci. This leaves three possible solutions: nonzero expression from the X locus only, nonzero expression from the Y locus only, and nonzero expression from both loci. Which solution applies to a particular system depends on the relationship among the baseline mean phenotype value (η_0), the magnitude of the conflict (α), and the scaling coefficients relating mean and variance of gene expression level (β_x and β_y). Specifically, nonzero expression from both X and Y will hold for

$$-\alpha\left(1 + \frac{2}{\beta_x}\right) < \eta_0 < \alpha\left(1 + \frac{2}{\beta_y}\right), \tag{28}$$

Basically, if the mean baseline phenotype value η_0 is very negative, there will be strong selection to increase expression from the X locus. However, the accumulation of phenotypic noise due to the variance in expression from X will set the equilibrium phenotype below the value favored by a maternally expressed allele at the Y locus ($\phi_0 < -\alpha$ at equilibrium), and selection will favor complete silencing of both alleles at Y. Similarly, if η_0 is large and positive, the equilibrium phenotype will lie above the patrilineal optimum ($\phi_0 > \alpha$), and the X locus will be silent. For β values of order 10^{-1}, only one locus will be expressed if the deviation of the mean baseline phenotype (η_0) from zero is more than approximately 10-fold greater than the magnitude of the conflict between the two sets of alleles (2α). Thus, nonzero expression from both loci will be more likely if the conflict is greater (large α), and if the phenotype in the absence of expression from either locus is closer to optimal (small absolute value of η_0). The derivation of equation (28) is presented in the Appendix, along with analysis of the case in which only one locus is expressed at the ESS.

For systems in which both loci are have nonzero expression at the ESS, stability conditions require that the first derivatives given in equations (24) and (25) simultaneously equal zero, yielding evolutionarily stable values of x_0 and y_0:

$$x_0 = \frac{2\alpha + (\alpha - \eta_0)\beta_y}{\eta_x(\beta_x + \beta_y + \beta_x\beta_y)} \tag{29}$$

$$y_0 = \frac{2\alpha + (\alpha + \eta_0)\beta_x}{\eta_y(\beta_x + \beta_y + \beta_x\beta_y)} \tag{30}$$

as well as the phenotypic mean and variance

$$\phi_0 = \frac{\eta_0\beta_x\beta_y + \alpha(\beta_y - \beta_x)}{\beta_x + \beta_y + \beta_x\beta_y} \tag{31}$$

$$\sigma_\phi^2 = \sigma_0^2 + \frac{\beta_x\beta_y^2\,(\alpha-\eta_0)^2 + \beta_x^2\beta_y(\alpha+\eta_0)^2 + 4\alpha^2(\beta_x+\beta_y+2\beta_x\beta_y)}{\left(\beta_x+\beta_y+\beta_x\,\beta_y\right)^2}, \tag{32}$$

and the cost

$$\xi_O = \frac{\alpha^2(4+\beta_x+\beta_y)+\eta_0^2\,\beta_x\,\beta_y}{\beta_x+\beta_y+\beta_x\,\beta_y}. \tag{33}$$

It is straightforward, then, to calculate the imprinting load ξ_I. For example, assuming $\eta_0 < 0$, the appropriate comparison is with the equilibrium fitness associated with unimprinted expression from the *X* locus (given by equation (A23) in the Appendix)

$$\xi_I = \frac{\alpha^2(4+\beta_x+\beta_y)}{\beta_x+\beta_y+\beta_x\beta_y} - \frac{\beta_x^2\eta_0^2}{(1+\beta_x)(\beta_x+\beta_y+\beta_x\beta_y)}. \tag{34}$$

The corresponding equation for the $\eta_0 > 0$ case can be obtained by simply exchanging the x and y subscripts in (34). The dependence of these quantities ($\langle w\rangle$, ϕ_0, σ_ϕ^2, ξ_I, x_0, and y_0) on η_0 is illustrated in Figure 3 both with and without imprinting. In this figure, $\beta_x < \beta_y$, partially illustrating the effect of changes in the β coefficients on the system quantities.

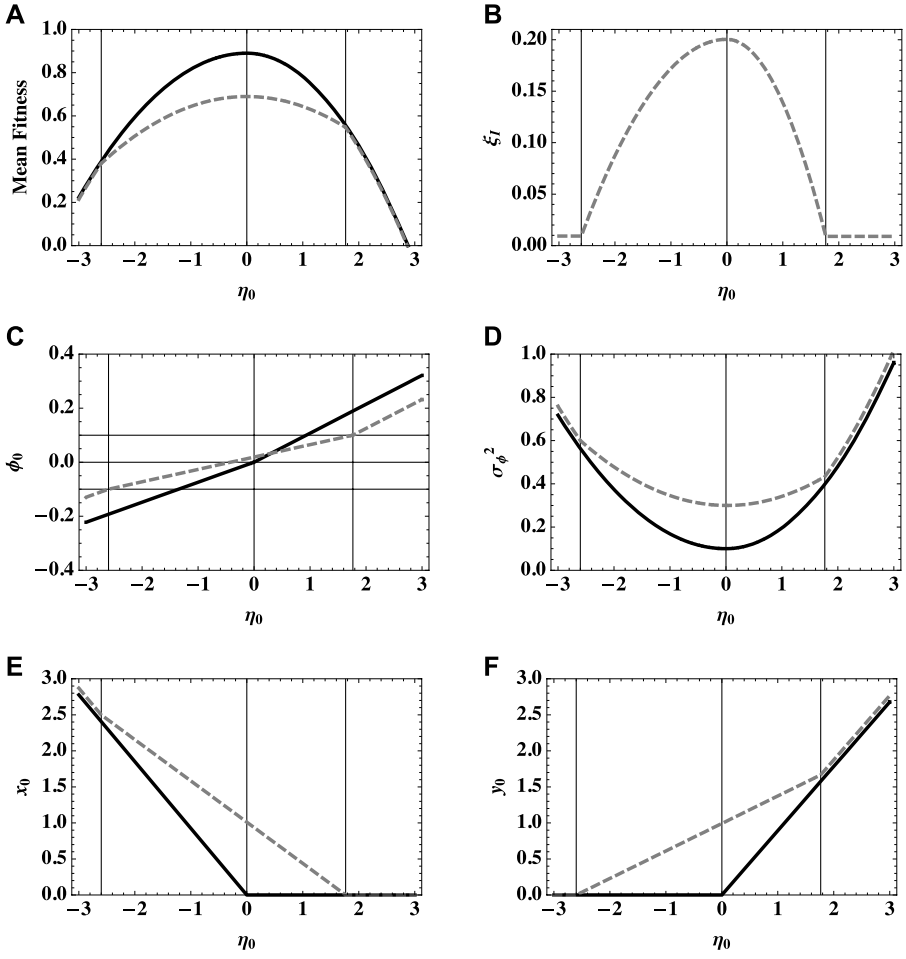

Figure 3. These figures illustrate the behavior of key quantities, each as a function of η_0. In this figure, $\alpha = 0.1$, $\sigma_0^2 = 0.1$, $\beta_X = 0.08$ and $\beta_Y = 0.12$. These values set the range over which both loci are expressed to $-2.6 < \eta_0 < 1.76667$. The quantities represented are (A) population mean fitness, (B) imprinting load, (C) population mean phenotype, (D) phenotypic variance, (E) mean expression level from the *X* locus, and (F) mean expression level from the *Y* locus. The value of each quantity with imprinting is indicated in each panel by the dotted line. The solid line indicates the value of the same quantity in the case in which both loci are unimprinted. The imprinting load (panel B) in the absence of imprinting is, by definition, zero for all values of η_0. Note that for most quantities, these two values diverge significantly only in the region where both imprinted loci have nonzero expression ($-2.6 < \eta_0 < 1.76667$ for these values of α, σ_0^2, β_X, and β_Y). The one exception is the mean phenotype ϕ_0 (panel C). This difference is due to the fact that natural selection at an unimprinted locus is driving ϕ_0 toward zero, as opposed to either α or $-\alpha$ for an imprinted locus. For each panel in this figure, the horizontal axis indicates the value of η_0, and the vertical lines indicate the critical values (-2.6 and 1.76667) for this quantity.

Nonlinearity in the Phenotypic Effects of Gene Expression

The results developed above assume that the phenotypic consequences of changes in gene expression from the two loci are both linear and additive. In many cases, the gene products will interact in a nonadditive way to produce the phenotype, for instance, if the two gene products compete for binding to a particular receptor. A detailed analysis of more realistic interaction models is beyond the scope of the current article, and will be the subject of future work.

It is possible, however, within the current context, to consider the case in which each locus has an independent nonlinear effect on the phenotype. For example, in many cases, it may be reasonable to assume that the phenotypic consequences of changes in gene expression are subject to diminishing returns. The conclusions developed for the linear model remain qualitatively unchanged, so long as the phenotypic effects and noise susceptivities are monotonic functions of the expression level of both genes. Specifically:

$$\frac{\partial \phi}{\partial X} > 0\,;\ \frac{\partial \phi}{\partial Y} < 0\,;\ \frac{\partial \nu_X}{\partial X} > 0\,;\quad \text{and}\quad \frac{\partial \nu_Y}{\partial X} > 0\,. \tag{35}$$

So long as the noise susceptivities are both increasing functions of expression level, the accumulation of phenotypic noise will halt the escalation at the point where

$$\nu_X + \nu_Y = 4\alpha \tag{36}$$

(see general conclusions in the Appendix). That is, it is the ratio of the change in phenotypic variance to the change in the mean phenotype that brings the escalation to a halt. If the effect of gene expression on the phenotype is subject to diminishing returns, then as expression increases, the consequences will diminish for both the mean and the variance. In fact, to first order—where the phenotypic consequences of variation in expression level are treated by a linear approximation around the mean phenotype—the ratio of the two is independent of the magnitude of the effect on the phenotype.

The generality of the conclusions can also be inferred from the fact that none of the phenotypic and fitness values derived from the linear model contain η_x or η_y, which represent the extent to which changes in expression level affect the phenotype. Those coefficients appear only in the expressions for x_0 and y_0, the equilibrium level of expression from the two loci. Thus, to first approximation, diminishing returns of the phenotypic effects of gene expression from the two loci can simply be viewed as a reduction in η_x and η_y, which will increase the level of expression at equilibrium, but will not alter the equilibrium mean phenotype value at which decanalization brings the escalation to a halt.

There are, however, two important areas where the values of η_x and η_y, and the consequences of diminishing returns, may have a substantial effect. The first is in the relationship between the decanalization process described here and other factors that might limit the escalation of gene expression. The less significant the consequences of increased gene expression for the phenotype, the more likely that the escalation is halted by a mechanistic limit to gene expression from one of the two loci, for instance. The second major effect is in the extent of the phenotypic variation at the ESS. As the phenotypic effects of changes in gene expression level decline, so will the magnitude of the phenotypic variance σ_ϕ^2. These consequences can be illustrated through a brief analysis of a slightly more general model of phenotype construction:

$$\phi = \eta_b + f_x(X) - f_y(Y)\,, \tag{37}$$

Assuming, as before, a constant coefficient of variation for the level of expression from each locus, mean expression levels x_0 and y_0, yields the following approximate expression (based on the Taylor's series expansion used above) for the variance of the distribution of phenotype values

$$\sigma_\phi^2 = \sigma_0^2 + \frac{1}{2}(\beta_x x_0 f_x'(x_0))^2 + \frac{1}{2}(\beta_y y_0 f_y'(y_0))^2. \tag{38}$$

The critical insight from equation (38) is that the phenotypic variance depends on the product of the expression level (x_0 or y_0) with the sensitivity of the phenotype to changes in the expression level ($f_x'(x_0)$ or $f_y'(y_0)$). In the linear model, these sensitivities are constant, and are represented by η_x and η_y. For a model with diminishing effects of gene expression, these sensitivities will be decreasing functions of x_0 and y_0, respectively. If the sensitivities fall off rapidly (faster than $1/x_0$ or $1/y_0$), the phenotypic variance will actually decrease as a result of the escalation. It is therefore possible that modest escalation, resulting from modest conflict, could decanalize the phenotype, whereas more extreme escalation, resulting from more intense conflict, could reverse the trend, leading to recanalization.

The generic ESS conditions (for the case in which both loci have nonzero expression) are found by simultaneously solving the following two stability conditions:

$$\beta_x^2 x_0(x_0 f_x''(x_0) - f_x'(x_0)) + \eta_0 + f_x(x_0) - f_y(y_0) - \alpha = 0 \quad \text{and} \tag{39}$$

$$\beta_x^2 y_0\,(y_0 f_y''(y_0) - f_y'(y_0)) - \eta_0 - f_x(x_0) + f_y(y_0) - \alpha = 0. \tag{40}$$

Once the equilibrium values of x_0 and y_0 have been identified, these values can then be inserted into the other expressions presented here to produce the various quantities discussed for the linear model (phenotypic variance, imprinting load, etc.). However, for most biologically relevant functions f_x and f_y, these

expressions do not lend themselves to simple, closed-form analytic solutions.

Furthermore, biologically realistic models will necessarily consider more than the two loci discussed here. For example, if X and Y produce ligands that affect the phenotype through binding to one or more receptors, alleles encoding those receptors will be coevolving simultaneously, and will be subject to stochastic variation in expression level. The case in which X and Y act independently with diminishing returns might reflect a system in which each acts through a different receptor, and binding becomes saturated, at which point the phenotypic variance would be determined by the expression variance for the two receptors. A detailed analysis of specific systems will be the subject of future work.

Fitness Effects of Canalization Modifiers

Within the simple, linear model, it is also possible to examine the fitness effects of canalization modifiers. The analysis that follows assumes the possibility of a mutation within the *cis*-acting regulatory region of one of the two loci that affects the coefficient of the linear relationship between the variance and mean of the gene expression level at that locus. I will consider both the short-term and long-term fitness effects of changes in the noise susceptivity coefficients β_x and β_y. The short-term effect is defined as the fitness change $\Delta w_{ST} = w(\beta_\bullet + \Delta\beta_\bullet) - w(\beta_\bullet)$ holding x_0 and y_0 constant. The long-term effect Δw_{LT} is the change in fitness after x_0 and y_0 have reached their new equilibrium values. Each of these fitness changes will be evaluated for three conditions: (1) no imprinting, (2) imprinted expression from one locus only, and (3) imprinted gene expression from both loci. In the first two cases, I will focus on the $\eta_0 < 0$ case, where expression is from the X locus, although analogous results hold for expression from Y with $\eta_0 > 0$.

First, the case without imprinting

$$\Delta w_{ST} = -\frac{\Delta\beta_x \eta_0^2}{(1+\beta_x)^2}. \tag{41}$$

$$\Delta w_{LT} = -\frac{\Delta\beta_x \eta_0^2}{(1+\beta_x)(1+\beta_x+\Delta\beta_x)}. \tag{42}$$

These results can be interpreted intuitively as natural selection favoring canalization of gene expression. Note that both Δw_{ST} and Δw_{LT} are opposite in sign from $\Delta\beta_x$, meaning that modifiers that reduce β_x are favored. It is also clear that $\Delta w_{LT} > \Delta w_{ST}$ for any nonzero value of $\Delta\beta_x$ (although this difference is of order $\Delta\beta_x^2$). Thus a canalizing modifier (one for which $\Delta\beta_x < 0$) results in an immediate increase in fitness due to a reduction in phenotypic variance, and then induces a further fitness increase as the mean expression level x_0 adjusts to its new optimum.

Similar results hold for a single imprinted locus, where

$$\Delta w_{ST} = -\frac{\Delta\beta_x(\alpha-\eta_0)^2}{(1+\beta_x)^2} \tag{43}$$

$$\Delta w_{LT} = -\frac{\Delta\beta_x\left(\alpha^2-\eta_0^2\right)}{(1+\beta_x)(1+\beta_x+\Delta\beta_x)}. \tag{44}$$

Clearly, Δw_{ST} is opposite in sign from $\Delta\beta_x$, again indicating that a modifier that canalizes gene expression level will be selectively favored. Similarly, Δw_{LT} is opposite in sign from $\Delta\beta_x$, because the conditions for transcriptional silencing of Y require that $\eta_0^2 > \alpha^2$. Furthermore, for permitted values of η_0 (where (28) is not satisfied), Δw_{LT} is always greater than Δw_{ST}.

A canalization modifier in the system with two oppositely imprinted loci, however, produces quite different results. The direct effects of selection will favor a canalizing modifier, just as in the two cases considered above. For example, considering again a modifier that changes the value of β_x

$$\Delta w_{ST} = -\frac{\Delta\beta_x(2\alpha+\beta_y(\alpha-\eta_0))^2}{(\beta_x+\beta_y+\beta_x\beta_y)^2}. \tag{45}$$

As was true in the previous cases, Δw_{ST} is opposite in sign from $\Delta\beta_x$, meaning that natural selection will favor a canalizing modifier ($\Delta\beta_x < 0$), which will reduce the phenotypic variance. However, the long-term fitness consequences of the canalizing modifier will be, for most parameter values, a reduction in mean fitness as x_0 and y_0 take on their new evolutionarily stable values.

$$\Delta w_{LT} = \frac{\Delta\beta_x\left(\alpha^2(2+\beta_y)^2-\beta_y^2\eta_0^2\right)}{(\beta_x+\beta_y+\beta_x\beta_y)(\beta_x+\Delta\beta_x+\beta_y+(\beta_x+\Delta\beta_x)\beta_y)}. \tag{46}$$

Specifically, Δw_{LT} will have the same sign as $\Delta\beta_x$ under the following condition:

$$\eta_0^2 < \alpha^2\left(1+\frac{2}{\beta_y}\right)^2. \tag{47}$$

Note the similarity to the conditions under which the ESS involves nonzero expression from both loci (expression (28)).

If $\beta_x > \beta_y$, then for any value of η_0 consistent with expression from both loci, a reduction in $\beta_x(\Delta\beta_x < 0)$ will be selectively favored, but will result in a decrease in fitness in the long term. If $\beta_x < \beta_y$, then there will exist a range of values of $\eta_0(-\alpha(1+2/\beta_x) < \eta_0 < -\alpha(1+2/\beta_y))$ for which canalization of gene expression will result in a long-term fitness increase, although even in these cases, the long-term fitness will be lower than the fitness immediately following the reduction in $\beta_x(\Delta w_{ST} > \Delta w_{LT} > 0)$. The short- and long-term fitness consequences of canalization modifiers are illustrated in Figure 4. The expression for the

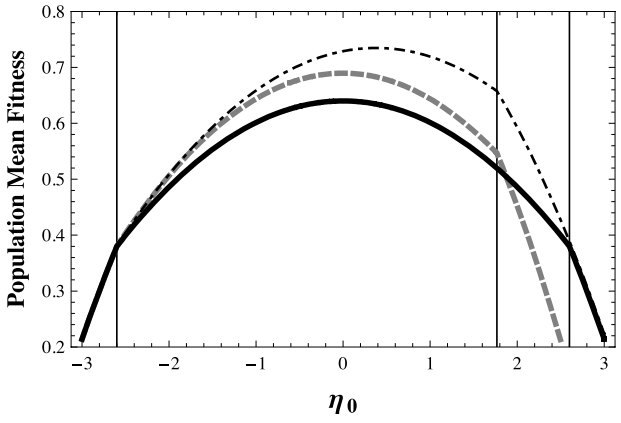

Figure 4. This graph illustrates the short-term and long-term fitness consequences of a canalization modifier acting on the maternally expressed *Y* locus. As in Figure 2, $\alpha = 0.1$, $\sigma_0^2 = 0.1$, and $\beta_X = 0.08$. The dashed gray line represents the equilibrium population mean fitness with $\beta_Y = 0.12$, identical to the dashed line in Figure 2a. The immediate fitness effect of a canalization modifier that reduces β_Y to 0.08 is indicated by the dot-dashed black line. The solid black line indicates the population mean fitness after the system has reached the equilibrium associated with the new value of β_Y. Note that for any system starting off with nonzero expression from both loci, the long-term consequence of canalization is a reduction in the population mean fitness. Prior to fixation of the canalization modifier, the both loci have nonzero expression for the range ($-2.6 < \eta_0 < 2.6$). After expression levels have reached their new equilibrium following fixation of the canalization modifier, this range is reduced to ($-2.6 < \eta_0 < 1.76667$). These three critical values of η_0 (-2.6, 1.76667, and 2.6) are indicated by vertical lines in the figure.

long-term fitness consequences of a change in β_y is given by

$$\Delta w_{\mathrm{LT}} = \frac{\Delta\beta_y \left(\alpha^2 (2+\beta_x)^2 - \beta_x^2 \eta_0^2\right)}{(\beta_x + \beta_y + \beta_x\beta_y)(\beta_x + \beta_y + \Delta\beta_y + \beta_x(\beta_y + \Delta\beta_y))}. \quad (48)$$

Decanalization and Common Disease States

I will outline briefly how the phenomenon of conflict-induced decanalization, resulting from the antagonistic coevolution of a pair of oppositely imprinted genes, might have contributed to the elevated frequencies of certain disease states, or extreme phenotypes, the clinical consequences of which suggest that they should be strongly disfavored by natural selection. I hope that this discussion will serve two purposes. First, it will illustrate the implications of the primary results of the model within a specific biological context, hopefully making those results more intuitively accessible. Second, it will point toward specific arenas where the concepts developed here might contribute to our understanding of complex diseases, and eventually aid in the interpretation of data bearing on the genetic contributors to those diseases.

FETAL GROWTH AND GESTATION TIME

Most of our understanding of imprinted genes comes from their effects on prenatal growth. As noted in the introductory sections of this article, the general pattern is that alleles favor placing a greater demand on maternal resources when paternally derived than when maternally derived. This conflict can manifest in the form of conflict over the fetal growth rate and/or duration of gestation. For closely related reasons, imprinted gene effects have been linked to traits that modulate resource acquisition postnatally, including suckling and weaning behaviors (Haig and Wharton 2003; Úbeda 2008) and metabolism (Frontera et al. 2008).

One specific phenotype to which this decanalization process might prove relevant is timing of parturition. The classical models of imprinting, which focus on the conflict over the distribution of maternal resources, would predict that paternally derived alleles should favor a slightly longer gestation period, whereas maternally derived alleles should favor a slightly shorter one. Decanalization arising from this conflict could help to explain, for example, preterm birth in humans, which occurs at high frequencies (~0.7% of U.S. births occurring at less than 28 weeks), and is associated with a substantial increase in the probability of neonatal death (Goldenberg et al. 2008).

Pregnancy is associated with a range of other common disorders potentially related to imprinted genes. Imprinted genes play a role in placental development (Lambertini et al. 2008), dysregulation of which is associated with a large fraction of spontaneous abortions (Bressan et al. 2009). Altered levels of gene expression have also been associated with intrauterine growth restriction (Diplas et al. 2009) and preeclampsia (Sitras et al. 2009). Reduced birth weight (Small for Gestational Age) has also been associated with variation in expression of particular imprinted genes (Guo et al. 2008).

In considering the possibility that conflict-induced decanalization might contribute to the high frequency of one or more of these disorders, it is important to keep in mind that imprinted genes are not the only locus of genetic conflict in pregnancy. In addition to the conflict between maternally and paternally derived alleles within the fetus and placenta, there is an evolutionary conflict between mother and fetus, where the transfer of maternal resources is modulated through the exchange of various signaling molecules and manipulation of the maternal vasculature (Haig 1993). The arguments that have been developed here for a pair of imprinted genes apply equally to the maternal–fetal conflict, where genes expressed in the mother and the fetus (whether imprinted or not) with opposing effects on the fetal growth phenotype will become involved in an arms race, leading to possible decanalization through the mechanisms described here.

MAJOR PSYCHIATRIC DISORDERS, INCLUDING AUTISM AND SCHIZOPHRENIA

The major psychiatric disorders represent an evolutionary conundrum, because they are often associated with severe cognitive dysfunction and likely dramatically reduced fitness. Naïve calculations based on mutation–selection balance would suggest that heritable diseases should be rare in the population, occurring at frequencies of 10^{-4}–10^{-5}. Despite devastating clinical manifestations, however, many of these disorders are significantly more common than this simple calculation would predict, with frequencies as high as 10^{-2}.

The results presented here suggest an explanation for the high frequencies of these disorders that differs from other explanations in the literature. Some of those other explanations depend on compensatory selective benefits associated with extreme cognitive phenotypes (e.g., creativity or social cognition in the context of schizophrenia (Jonas and Jonas 1975; Nettle and Clegg 2006; Pearlson and Folley 2008), and mathematical or systematic reasoning in the context of autism (Jolliffe and Baron-Cohen 1997; Ijichi et al. 2008; Baron-Cohen et al. 2009)). The other major class of candidate explanation in the literature focuses on complex genetic underpinnings for the diseases. These models have recently been reviewed and synthesized (Slatkin 2008), and demonstrate that if a disease is associated with a sufficient number of loci, and the epistatic interactions are sufficiently strong, the observed patterns of heritability may be consistent with mutation–selection balance. At the moment, this model suggests an explanation that could either be an alternative to those other classes of explanation, or could be a complement to them. Quantitative analysis of the relative contribution of different processes to explaining the high frequencies of these heritable diseases will require more detailed modeling of particular diseases, as well as additional empirical research.

In addition to their effects on early growth and development, many imprinted loci exhibit parent-of-origin-specific expression in adult tissues, including the brain. The effects of imprinted gene expression in neural tissues are not well understood, but imprinted genes have been associated with a number of cognitive and behavioral phenotypes (reviewed in Wilkinson et al. 2007), including postnatal feeding behaviors (Haig and Wharton 2003; Curley et al. 2004; Plagge et al. 2004), maternal care for offspring (Lefebvre et al. 1998; Li et al. 1999; Wilkins and Haig 2003a), kin recognition (Isles et al. 2002), reactivity to novel environments (Plagge et al. 2005), and various aspects of social cognition (Goos and Silverman 2006; Goos and Ragsdale 2008). In addition to the behavioral phenotypes associated with knockouts of imprinted genes, a number of studies have identified neurological effects of imprinted genes. Specific imprinted genes have been implicated in memory consolidation (Brambilla et al. 1997), long-term potentiation of neurons (Jiang et al. 1998), learning reversal (Davies et al. 2005), attention (Relkovic et al. 2010), and attention switching (Woodcock et al. 2009).

In a recent review, Crespi has catalogued the extensive evidence linking various imprinted genes to both autistic spectrum disorder (ASD) and psychotic spectrum disorder (PSD) conditions (Crespi 2008). This evidence includes psychiatric consequences associated with known genetic disorders involving the loss or duplication of chromosomal regions that include imprinted genes, as well as statistical associations of individual imprinted genes with both autism and schizophrenia. The pattern that runs throughout these different forms of evidence is that a deficit in paternal gene expression (or an excess of maternal expression) creates a predisposition for the development of PSD (including schizophrenia). Conversely, excess paternal expression (or deficient maternal expression) increases the risk for ASD. This pattern implies the existence of a phenotypic axis in the cognitive-behavioral domain over which there exists an intragenomic conflict between maternally and paternally derived alleles.

Badcock and Crespi have suggested that it may reflect variation in the relative amounts of neural resources devoted to "mentalistic" versus "mechanistic" brain functions (Badcock and Crespi 2006; Crespi and Badcock 2008). In this scenario, maternally derived alleles favor greater investment in mentalistic functions (including theory of mind), whereas paternally derived alleles favor greater investment in mechanistic functions (including spatial reasoning). Other possibilities include a straightforward extension of the conflict over maternal resources, which is played out through other mechanisms prenatally. Under this model, the social deficits associated with ASD (including lack of empathy) are equated with selfish behavior, producing a greater demand for resources (Crespi 2008).

For the purposes of the present analysis, however, the labeling of the axis is unimportant. The argument presented here does not rely on any particular interpretation of the phenotype that is the subject of the conflict. Nor does the argument rely on our having an understanding the evolutionary basis of the conflict. What matters here is that the molecular evidence implies a difference between the matrilineal and patrilineal fitness functions with respect to some cognitive phenotype, and that overshoot in the direction favored by maternally derived alleles results in risk for PSD, whereas overshoot in the direction favored by paternally derived alleles is associated with ASD.

The implications of this conflict in light of the model developed here are illustrated in Figure 5. ASD and PSD are treated as opposite extremes of a quantitative trait that is influenced by the expression levels of a number of loci, some of which are imprinted. The matrilineal and patrilineal fitness functions have different optima with respect to this trait.

In the absence of imprinting, the phenotypic variance in the population is limited by the deleterious consequences of extreme

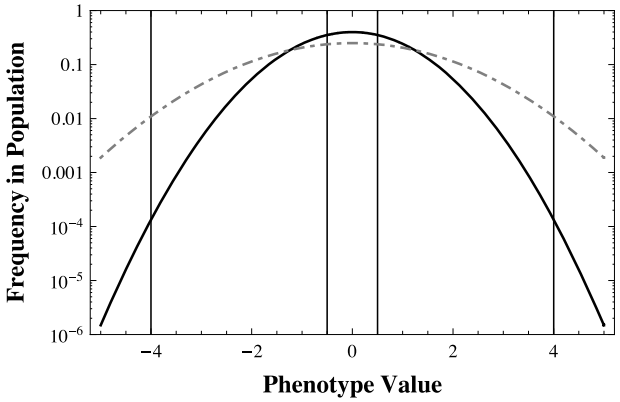

Figure 5. **This graph illustrates how the results developed in this article might help to explain the high frequencies of heritable major psychiatric disorders. The horizontal axis represents a quantitative trait. Deleterious effects are associated with extreme phenotype values, with autistic-spectrum disorders represented by the extreme right side of the graph (phenotype values greater than four), and psychotic-spectrum disorders represented by the extreme left side (phenotype values less than negative four). The vertical axis indicates the probability density function of the distribution of phenotypes in the population. The optimal phenotypes from the perspective of maternally and paternally derived alleles are at $-\alpha$ and α, respectively. In this figure, $\alpha = 0.5$, and $-\alpha$ and α, are indicated by vertical lines at these locations. In the absence of imprinting (solid black line), the width of the phenotype distribution is limited by mutation–selection balance, and the fraction of the population with autism or schizophrenia is limited to the 0.01% most extreme phenotypes. Genomic imprinting generates conflict-induced decanalization, resulting in a broader distribution of phenotypes (dot-dashed gray line), such that a much larger fraction of the population has the clinical manifestations associated with extreme phenotype values (in this case, ∼1%).**

phenotypes at either end of the spectrum. Roughly speaking, we might expect each of these extreme tails to represent on the order of 10^{-4} of the population. With genomic imprinting, however, genetic conflict among loci results in increased expression levels, and an increase in phenotypic variance. This results in a larger fraction of the phenotypic distribution lying beyond the points corresponding to clinical determinations of ASD or PSD. The arms race among loci ends when the deleterious consequences of the accumulation of extreme phenotypes outweigh the selective benefits of further escalation. In this case, the escalation terminates when the frequencies of these major psychiatric disorders approach 10^{-2}.

Conclusion

Models of genomic imprinting suggest that oppositely imprinted loci affecting the same trait will become involved in an interlocus, intragenomic arms race, with each locus under selection to increase expression from the transcriptionally active allele. This article develops a model for the termination of this arms race, based on the accumulation of transcriptional (and phenotypic) variance associated with increasing the level of gene expression. An evolutionarily stable outcome is attained when the fitness costs associated with a further increase in expression (due to the increase in phenotypic variance) counterbalance the fitness benefits of the increase.

In the first sections of the article, a general form of the model provides a few intuitively straightforward conclusions. The antagonistic two-locus system can resolve either into a state where one of the two loci is completely silenced, or a state in which both loci are expressed. Two-locus expression is favored when the magnitude of the conflict (the difference between the matrilineal and patrilineal fitness functions, 2α) is large, and when the baseline phenotype (the phenotype in the absence of expression from either locus, η_0) is close to the optimal phenotypes. Two-locus expression is also favored if the loci have low noise suseptivities (rates at which expression variance increases with the mean expression level, ν_X and ν_Y). The general model also revealed the phenotype at the ESS to be closer to the optimum of the locus with the lower noise susceptivity. That is, the locus that is more capable of increasing its expression level without inducing a corresponding increase in variance in some sense is the "winner" of the conflict.

Later sections of the article develop more specific results based on a simple, linear model of phenotype construction, and a noise susceptivity model that assumes a constant coefficient of variation for gene expression. Despite its simplicity, this model yields some useful and perhaps surprising conclusions. First, if both loci are expressed at the ESS, phenotypic variance is increased relative to the ESS in the absence of genomic imprinting. This conflict-induced decanalization is associated with a reduced population mean fitness, which can be quantified in terms of an imprinting load.

The second conclusion from the linear model is that canalization of gene expression often leads to decanalization of the phenotype. A modifier that reduces expression variance is always selectively favored in the model. However, under most parameter combinations for which both loci are expressed at the ESS, this reduction in variance induces a further escalation between the two loci, such that the new ESS actually has a higher phenotypic variance and reduced fitness.

Genomic imprinting therefore imposes two decanalization-related costs on the organism. The first cost is the increase in phenotypic variance associated with having two antagonistically coevolving imprinted genes that interact to generate the phenotype. The second cost relates to the evolutionary dynamics of decanalization made possible in this system, where gene expression increases, becoming successively more and more canalized,

whereas the phenotype becomes progressively decanalized. The magnitude of this second cost in real systems depends critically on empirical questions regarding the extent to which it is possible to substantially alter the relationship between expression mean and variance for individual loci. How large these two costs are relative to other sources of phenotypic variability (e.g., stochastic variation in gene expression at other loci, effects of deleterious mutations, or environmental variation, indicated here by the term σ_0^2) will require more detailed and quantitative modeling that lies beyond the scope of this article, but will be the focus of future work.

The results and conclusions presented here have been developed in the context of an extremely simple model of phenotype construction, where the phenotype value is a strictly linear function of the two expression levels. In any real system, interactions among genes will exhibit nonlinearities that will be need to be incorporated to provide an accurate quantitative assessment of the decanalization phenomenon described here. Mechanistically inspired models that produce nonlinearities lie beyond the scope of this article, however, and will be the subject of future work. It is worth noting, however, that, at least for small changes in the gene expression levels, these more realistic models could be approximated by the linear model developed above, suggesting that certain of the conclusions presented here may well hold qualitatively.

Finally, the results of the model are briefly discussed in two contexts in which imprinted genes are known to exhibit a significant influence, and where disease states with deleterious outcomes occur at high frequency. The first context is early developmental and growth effects, where imprinting is associated with a genetic conflict over the level of demand placed on maternal resources. The second context focuses on a particular coupled pair of common, yet highly deleterious heritable psychiatric disorders: autism and psychosis, where loss of maternal expression is associated with autism and ASDs, whereas loss of paternal expression is associated with schizophrenia and PSDs. The results presented here suggest a potential mechanism through which an intragenomic conflict over growth or cognitive phenotypes could have increased the phenotypic variance and produced elevated frequencies of major psychiatric disorders as well as disorders related to pregnancy.

ACKNOWLEDGMENTS

I would like to thank T. Bhattacharya, A. Clauset, N. Collins, B. Crespi, W. Kets, J. Masel, T. Taylor, and J. Van Cleve for helpful discussions and comments on the manuscript.

LITERATURE CITED

Badcock, C., and B. Crespi. 2006. Imbalanced genomic imprinting in brain development: an evolutionary basis for the aetiology of autism. J. Evol. Biol. 19:1007–1032.

Baron-Cohen, S., E. Ashwin, C. Ashwin, T. Tavassoli, and B. Chakrabarti. 2009. Talent in autism: hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. Philos. Trans. R. Soc. Lond. B 364:1377–1383.

Bergman, A., and M. L. Siegal. 2003. Evolutionary capacitance as a general feature of complex gene networks. Nature 424:549–552.

Brambilla, R., N. Gnesutta, L. Minichiello, G. White, A. J. Roylance, C. E. Herron, M. Ramsey, D. P. Wolfer, V. Cestari, C. Rossi-Arnaud, et al. 1997. A role for the Ras signalling pathway in synaptic transmission and long-term memory. Nature 390:281–286.

Brandvain, Y. 2010. Matrisibs, patrisibs and the evolution of genomic imprinting on autosomes and sex chromosomes. Am. Nat. 176:511–521.

Bressan, F. F., T. H. De Bem, F. Perecin, F. L. Lopes, C. E. Ambrosio, F. V. Meirelles, and M. A. Miglino. 2009. Unearthing the roles of imprinted genes in the placenta. Placenta 30:823–834.

Crespi, B. 2008. Genomic imprinting in the development and evolution of psychotic spectrum conditions. Biol. Rev. Camb. Philos. Soc. 83:441–493.

Crespi, B., and C. Badcock. 2008. Psychosis and autism as diametrical disorders of the social brain. Behav. Brain Sci. 31:241–261; discussion 261–320.

Curley, J., S. Barton, A. Surani, and E. Keverne. 2004. Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. Proc. Biol. Sci. 271:1303–1309.

Davies, W., A. Isles, R. Smith, D. Karunadasa, D. Burrmann, T. Humby, O. Ojarikre, C. Biggin, D. Skuse, P. Burgoyne, et al. 2005. Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice. Nat. Genet. 37:625–629.

Davies, W., A. R. Isles, T. Humby, and L. S. Wilkinson. 2008. What are imprinted genes doing in the brain? Adv. Exp. Med. Biol. 626:62–70.

Diplas, A. I., L. Lambertini, M. J. Lee, R. Sperling, Y. L. Lee, J. Wetmur, and J. Chen. 2009. Differential expression of imprinted genes in normal and IUGR human placentas. Epigenetics 4:235–240.

Feinberg, A. P., and R. A. Irizarry. 2010. Evolution in health and medicine Sackler colloquium: stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. Proc. Natl. Acad. Sci. USA 107(Suppl 1):1757–1764.

Flatt, T. 2005. The evolutionary genetics of canalization. Q. Rev. Biol. 80:287–316.

Frontera, M., B. Dickins, A. Plagge, and G. Kelsey. 2008. Imprinted genes, postnatal adaptations and enduring effects on energy homeostasis. Adv. Exp. Med. Biol. 626:41–61.

Gibson, G. 2009. Decanalization and the origin of complex disease. Nat. Rev. Genet. 10:134–140.

Goldenberg, R. L., J. F. Culhane, J. D. Iams, and R. Romero. 2008. Epidemiology and causes of preterm birth. Lancet 371:75–84.

Goos, L. M., and G. Ragsdale. 2008. Genomic imprinting and human psychology: cognition, behavior and pathology. Adv. Exp. Med. Biol. 626:71–88.

Goos, L. M., and I. Silverman. 2006. The inheritance of cognitive skills: does genomic imprinting play a role? J. Neurogenet. 20:19–40.

Guo, L., S. Choufani, J. Ferreira, A. Smith, D. Chitayat, C. Shuman, R. Uxa, S. Keating, J. Kingdom, and R. Weksberg. 2008. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. Dev. Biol. 320:79–91.

Haig, D. 1993. Genetic conflicts in human pregnancy. Q. Rev. Biol. 68:495–532.

———. 1996. Placental hormones, genomic imprinting, and maternal-fetal communication. J. Evol. Biol. 9:357–380.

———. 2000. The kinship theory of genomic imprinting. Annu. Rev. Ecol. Syst. 31:9–32.

Haig, D., and R. Wharton. 2003. Prader-Willi syndrome and the evolution of human childhood. Am. J. Hum. Biol. 15:320–329.

Ijichi, S., N. Ijichi, Y. Ijichi, Y. Kawamura, T. Hashiguchi, and H. Morioka. 2008. For others: epistasis and the evolutionary survival of an extreme tail of the quantitative distribution of autistic assets. Med. Hypotheses 70:515–521.

Isles, A. R., M. J. Baum, D. Ma, A. Szeto, E. Keverne, and N. D. Allen. 2002. A possible role for imprinted genes in inbreeding avoidance and dispersal from the natal area in mice. Proc. Biol. Sci. 269:665–670.

Jiang, Y. H., D. Armstrong, U. Albrecht, C. M. Atkins, J. L. Noebels, G. Eichele, J. D. Sweatt, and A. L. Beaudet. 1998. Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation. Neuron 21:799–811.

Jolliffe, T., and S. Baron-Cohen. 1997. Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? J. Child Psychol. Psyc. Allied Disciplines 38:527–534.

Jonas, A. D., and D. F. Jonas. 1975. An evolutionary context for schizophrenia. Schizophr. Bull. Spring:33–41.

Kondoh, M., and M. Higashi. 2000. Reproductive isolation mechanism resulting from resolution of intragenomic conflict. Am. Nat. 156:511–518.

Lambertini, L., A. I. Diplas, M. J. Lee, R. Sperling, J. Chen, and J. Wetmur. 2008. A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta. Epigenetics 3:261–269.

Lefebvre, L., S. Viville, S. C. Barton, F. Ishino, E. B. Keverne, and M. A. Surani. 1998. Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest. Nat. Genet. 20:163–169.

Li, L., E. B. Keverne, S. A. Aparicio, F. Ishino, S. C. Barton, and M. A. Surani. 1999. Regulation of maternal behavior and offspring growth by paternally expressed Peg3. Science 284:330–333.

Masel, J., and M. L. Siegal. 2009. Robustness: mechanisms and consequences. Trends Genet. 25:395–403.

Meiklejohn, C. D., and D. L. Hartl. 2002. A single mode of canalization. Trends Ecol. Evol.17: 468–473.

Mochizuki, A., Y. Takeda, and Y. Iwasa. 1996. The evolution of genomic imprinting. Genetics 144:1283–1295.

Nettle, D., and H. Clegg. 2006. Schizotypy, creativity and mating success in humans. Proc. Biol. Sci. 273:611–615.

Newman, J. R., S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman. 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. Nature 441:840–846.

Ozbudak, E. M., M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. 2002. Regulation of noise in the expression of a single gene. Nat. Genet. 31:69–73.

Pearlson, G. D., and B. S. Folley. 2008. Schizophrenia, psychiatric genetics, and Darwinian psychiatry: an evolutionary framework. Schizophr. Bull. 34:722–733.

Plagge, A., E. Gordon, W. Dean, R. Boiani, S. Cinti, J. Peters, and G. Kelsey. 2004. The imprinted signaling protein XL alpha s is required for postnatal adaptation to feeding. Nat. Genet. 36:818–826.

Plagge, A., A. R. Isles, E. Gordon, T. Humby, W. Dean, S. Gritsch, R. Fischer-Colbrie, L. S. Wilkinson, and G. Kelsey. 2005. Imprinted Nesp55 influences behavioral reactivity to novel environments. Mol. Cell. Biol. 25:3019–3026.

Queitsch, C., T. A. Sangster, and S. Lindquist. 2002. Hsp90 as a capacitor of phenotypic variation. Nature 417:618–624.

Relkovic, D., C. M. Doe, T. Humby, K. A. Johnstone, J. L. Resnick, A. J. Holland, J. J. Hagan, L. S. Wilkinson, and A. R. Isles. 2010. Behavioural and cognitive abnormalities in an imprinting centre deletion mouse model for Prader-Willi syndrome. Eur. J. Neurosci. 31:156–164.

Sangster, T. A., N. Salathia, H. N. Lee, E. Watanabe, K. Schellenberg, K. Morneau, H. Wang, S. Undurraga, C. Queitsch, and S. Lindquist. 2008a. HSP90-buffered genetic variation is common in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 105:2969–2974.

Sangster, T. A., N. Salathia, S. Undurraga, R. Milo, K. Schellenberg, S. Lindquist, and C. Queitsch. 2008b. HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc. Natl. Acad. Sci. USA 105:2963–2968.

Siegal, M. L., and A. Bergman. 2002. Waddington's canalization revisited: developmental stability and evolution. Proc. Natl. Acad. Sci. USA 99:10528–10532.

Sitras, V., R. Paulssen, J. Leirvik, A. Vårtun, and G. Acharya. 2009. Placental gene expression profile in intrauterine growth restriction due to placental insufficiency. Reprod. Sci. 16:701–711.

Slatkin, M. 2008. Exchangeable models of complex inherited diseases. Genetics 179:2253–2261.

Spielman, R. S., L. A. Bastone, J. T. Burdick, M. Morley, W. J. Ewens, and V. G. Cheung. 2007. Common genetic variants account for differences in gene expression among ethnic groups. Nat. Genet. 39:226–231.

Úbeda, F. 2008. Evolution of genomic imprinting with biparental care: implications for Prader-Willi and Angelman syndromes. PLoS Biol. 6:e208.

Úbeda, F., and A. Gardner. 2010. A model for genomic imprinting in the social brain: juveniles. Evolution 64:2587–2600.

Van Cleve, J., L. Lehmann, and M. W. Feldman. 2010. How demography, life history, and kinship shape the evolution of genomic imprinting. Am. Nat. 176:440–455.

Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. Nature 150:563–565.

———. 1959. Canalization of development and genetic assimilation of acquired characters. Nature 183:1654–1655.

Warton, D. I., I. J. Wright, D. S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. Biol. Rev. Camb. Philos. Soc. 81:259–291.

Wild, G., and S. A. West. 2009. Genomic imprinting and sex allocation. Am. Nat. 173:E1–E14.

Wilkins, J. F. 2009. Antagonistic coevolution of two imprinted loci with pleiotropic effects. Evolution 64:142–151.

Wilkins, J. F., and D. Haig. 2003a. Inbreeding, maternal care and genomic imprinting. J. Theor. Biol. 221:559–564.

———. 2003b. What good is genomic imprinting: the function of parent-specific gene expression. Nat. Rev. Genet. 4:359–368.

Wilkins, J. F., ed. 2008. Genomic imprinting. Springer, New York.

Wilkins, J. F., and D. Haig. 2001. Genomic imprinting of two antagonistic loci. Proc. Biol. Sci. 268:1861–1867.

Wilkinson, L. S., W. Davies, and A. R. Isles. 2007. Genomic imprinting effects on brain development and function. Nat. Rev. Neurosci. 8:832–843.

Woodcock, K. A., C. Oliver, and G. W. Humphreys. 2009. Task-switching deficits and repetitive behaviour in genetic neurodevelopmental disorders: data from children with Prader-Willi syndrome chromosome 15 q11–q13 deletion and boys with Fragile X syndrome. Cogn. Neuropsychol. 26:172–194.

Associate Editor: J. Kelly

Appendix

The major results of the analysis have been presented in the main text. The derivations of certain mathematical results have been provided here for completeness.

ESS CONDITIONS FOR THE MODEL

The model developed here focuses on the evolutionarily stable pattern of gene expression for a pair of loci with opposite effects on a single phenotypic value ϕ. The model assumes that increasing expression from the X locus increases this value, and that increasing expression from the Y locus decreases the value (Equation (1), main text). The model also assumes that the value of ϕ that maximizes the patrilineal inclusive fitness is greater than the value that maximizes the matrilineal inclusive fitness.

An ESS of the system is defined as a pattern of expression where neither locus can be invaded by an allele that differs from the fixed allele in the population by a small change in its level of expression, either when maternally or paternally derived. The state of the system is defined by a vector of expression levels [x_m, x_p, y_m, and y_p], encoded by the fixed alleles at the X and Y loci. The state is evolutionarily stable if a small change in either x_m or y_m does not increase the matrilineal inclusive fitness, at the same time that a small change in either x_p or y_p will not increase the patrilineal inclusive fitness. For example, for a paternally derived allele at the X locus, this implies one of two sets of conditions: either

$$\frac{\partial\langle w_p\rangle}{\partial x_p} = 0 \quad \text{and} \quad \frac{\partial^2\langle w_p\rangle}{\partial x_p^2} < 0 \quad \text{or} \tag{A1}$$

$$x_p = 0 \quad \text{and} \quad \frac{\partial\langle w_p\rangle}{\partial x_p} < 0. \tag{A2}$$

Analogous conditions must hold simultaneously for each of the other three allelic positions. The relevant derivatives of the two fitness functions are obtained by differentiating equations (6) and (7) from the main text. Using z as a dummy to indicate either x or y (and Z for X or Y), these derivatives are

$$\frac{\partial\langle w_p\rangle}{\partial z_p} = -\left(\frac{\partial\sigma_\phi^2}{\partial z} + 2(\phi_0 - \alpha)\frac{\partial\phi_0}{\partial z}\right) \tag{A3}$$

$$\frac{\partial\langle w_m\rangle}{\partial z_m} = -\left(\frac{\partial\sigma_\phi^2}{\partial z} + 2(\phi_0 + \alpha)\frac{\partial\phi_0}{\partial z}\right). \tag{A4}$$

$$\frac{\partial^2\langle w_p\rangle}{\partial z_p^2} = -\left(\frac{\partial^2\sigma_\phi^2}{\partial Z^2} + 2\left(\frac{\partial\phi_0}{\partial Z}\right)^2 + 2(\phi_0 - \alpha)\frac{\partial^2\phi_0}{\partial Z^2}\right) \tag{A5}$$

$$\frac{\partial^2\langle w_m\rangle}{\partial z_m^2} = -\left(\frac{\partial^2\sigma_\phi^2}{\partial Z^2} + 2\left(\frac{\partial\phi_0}{\partial Z}\right)^2 + 2(\phi_0 - \alpha)\frac{\partial^2\phi_0}{\partial Z^2}\right). \tag{A6}$$

CONCLUSIONS FROM THE GENERAL MODEL

The results presented in the main text limit consideration to the case in which the two gene products contribute independently (additively) to the phenotype value of the individual, and focus on a particular relationship between the mean and variance of the expression level from a locus. It is also useful to investigate what conclusions can be drawn from an analysis of the general model that makes no assumptions regarding the relationship between gene expression level and phenotypic variance. Equations (A3) and (A4) can be combined to yield

$$\frac{\partial\langle w_p\rangle}{\partial z_p} - \frac{\partial\langle w_m\rangle}{\partial z_m} = 4\alpha\frac{\partial\phi_0}{\partial Z}. \tag{A7}$$

Recall that according to the assumptions and construction of the model, $\alpha > 0$, $\partial\phi_0/\partial X > 0$, and $\partial\phi_0/\partial Y < 0$. Therefore, for any values of X, Y, and ϕ_0,

$$\frac{\partial\langle w_p\rangle}{\partial z_p} > \frac{\partial\langle w_m\rangle}{\partial z_m} \quad \text{and} \tag{A8}$$

$$\frac{\partial\langle w_p\rangle}{\partial y_p} < \frac{\partial\langle w_m\rangle}{\partial y_m}. \tag{A9}$$

These relationships restrict the possible combinations of evolutionary stability conditions relevant to the system. Any ESS will have $x_\mathrm{m} = y_\mathrm{p} = 0$; x_p and y_m must each either be zero or at a local fitness maximum.

The case where $x_\mathrm{p} = y_\mathrm{m} = X = Y = 0$, corresponding to complete loss of expression from both loci, will be evolutionarily stable only under very specific combinations of parameter values, and will not receive further consideration here. This leaves three possible ESS combinations

$$\frac{\partial\langle w_p\rangle}{\partial x_p} = 0;\ \frac{\partial^2\langle w_p\rangle}{\partial x_p^2} < 0;\ \frac{\partial\langle w_m\rangle}{\partial y_m} = 0;\ \frac{\partial^2\langle w_m\rangle}{\partial y_m^2} < 0 \tag{A10}$$

$$x_p = 0;\ \frac{\partial\langle w_p\rangle}{\partial x_p} < 0;\ \frac{\partial\langle w_m\rangle}{\partial y_m} = 0;\ \frac{\partial^2\langle w_m\rangle}{\partial y_m^2} < 0 \tag{A11}$$

$$\frac{\partial\langle w_p\rangle}{\partial x_p} = 0;\ \frac{\partial^2\langle w_p\rangle}{\partial x_p^2} < 0;\ y_m = 0;\ \frac{\partial\langle w_m\rangle}{\partial y_m} < 0. \tag{A12}$$

Note that (A10) describes the case in which there is nonzero expression from both loci, whereas (A11) corresponds to the complete silencing of the X locus, and (A12) corresponds to complete silencing of the Y locus.

Considering (A11) in more detail, it is possible to define the conditions under which complete loss of expression from one of the two loci should be expected. Setting equation (A3) equal to zero for $z = y$ yields

$$\frac{\partial\sigma_\phi^2/\partial Y}{\partial\phi_0/\partial Y} = -2(\alpha + \phi_0). \tag{A13}$$

Similarly, the condition on x_p (recalling that $\partial\phi_0/\partial X > 0$) yields

$$\frac{\partial\sigma_\phi^2/\partial X}{\partial\phi_0/\partial X} > 2(\alpha + \phi_0). \tag{A14}$$

Combining these two equations produces the relationship

$$\frac{\partial\sigma_\phi^2/\partial X}{\partial\phi_0/\partial X} > \frac{\partial\sigma_\phi^2/\partial Y}{\partial\phi_0/\partial Y} + 4\alpha. \quad \text{(A15)}$$

An analogous set of calculations for the conditions in (A12) produces the identical relationship. In terms of the noise susceptivities defined in the main text, the condition for stability of a state in which one of the loci is silenced can then be rewritten as

$$\nu_X + \nu_Y > 4\alpha. \quad \text{(A16)}$$

EVOLUTIONARY STABILITY OF COMPLETE SILENCING OF THE *X* OR *Y* LOCUS

As noted in the main text, there are three possible classes of ESS for the model developed here. Which type of ESS is reached by the system depends on the relationship among the values of α, η_0, β_x, and β_y. Specifically, complete silencing of Y will be evolutionarily stable if the following conditions hold:

$$\left\langle \frac{\partial w_p}{\partial x_0} \right\rangle = 2\eta_x(\alpha - \eta_0 - \eta_x x_0 - \eta_x\beta_x x_0) = 0 \quad \text{and} \quad \text{(A17)}$$

$$\left\langle \frac{\partial w_m}{\partial y_0} \right\rangle = 2\eta_y(\alpha + \eta_0 + \eta_x x_0) < 0. \quad \text{(A18)}$$

Combining equations (A17) and (A18) yields the following condition in terms of the underlying parameters of the system

$$\eta_0 < -\alpha\left(1 + \frac{2}{\beta_x}\right). \quad \text{(A19)}$$

That is, silencing of the Y locus is stable only of the baseline phenotype is sufficiently large and negative. Under these conditions, the equilibrium expression level and the phenotypic mean and variance are

$$x_0 = \frac{\alpha - \eta_0}{\eta_x(1+\beta_x)}, \quad \text{(A20)}$$

$$\phi_0 = \frac{\eta_0\beta_x + \alpha}{1+\beta_x}, \quad \text{and} \quad \text{(A21)}$$

$$\sigma_\phi^2 = \sigma_0^2 + \beta_x\left(\frac{\alpha - \eta_0}{1+\beta_x}\right)^2. \quad \text{(A22)}$$

Combining (A19) and (A21) reveals that the equilibrium mean phenotype value $\phi_0 < -\alpha$. Note that the solution has no dependence on features of the Y locus, and that the conditions for silencing of Y and equilibrium phenotype value do not depend on η_x.

The fitness reduction at the imprinted ESS is

$$\xi_O = \frac{\alpha^2 + \eta_0\beta_x}{1+\beta_x} > \alpha^2\left(1 + \frac{4}{\beta_x}\right), \quad \text{(A23)}$$

where the inequality indicates the lower bound on this cost set by the conditions on η_0 set by the assumption that Y is silenced. To characterize the imprinting load, it is next necessary to determine the ESS expected to evolve in the absence of imprinting. Analysis of the unimprinted system will focus on the case in which $\eta_0 < 0$, for which expression from the Y locus will be silenced. It is straightforward to derive the analogous results for the $\eta_0 > 0$ case. In the absence of imprinting, the ESS takes on the following values:

$$\phi_0 = \frac{\beta_x\eta_0}{1+\beta_x}, \quad \text{(A24)}$$

$$\sigma_\phi^2 = \sigma_0^2 + \beta_x\left(\frac{\eta_0}{1+\beta_x}\right)^2, \quad \text{(A25)}$$

$$\xi_U = \frac{\beta_x\eta_0^2}{1+\beta_x}. \quad \text{(A26)}$$

Subtracting equation (A26) from equation (A23) produces the imprinting load:

$$\xi_l = \alpha^2\left(\frac{1}{1+\beta_x}\right). \quad \text{(A27)}$$

Note that the imprinting load is a monotonically decreasing function of β_x. That is, the greater the noise susceptivity of the locus, the smaller the cost associated with genomic imprinting.

Analogous calculations produce similar results and interpretation for complete silencing of X. Specifically, silencing of X is evolutionarily stable if

$$\eta_0 > \alpha\left(1 + \frac{2}{\beta_y}\right), \quad \text{(A28)}$$

and the equilibrium expression, phenotype, and fitness cost values are analogous to those presented in equations (A20)–(A24), and the exact expressions can be retrieved by substituting β_y for β_x.