Horizontal Transmission Rapidly Erodes Disequilibria Between Organelle and Symbiont Genomes

Yaniv Brandvain,*,†,1 Charles Goodnight,‡ and Michael J. Wade*
*Department of Biology, Indiana University, Bloomington, Indiana 47405, †Center for Population Biology, University of California, Davis, California 95616, and ‡Department of Biology, University of Vermont, Burlington, Vermont 05401

ABSTRACT We investigate the generation and decay of interspecific disequilibrium (ID) between organelle and symbiont genomes as a function of the rate of horizontal transmission. We show that rare horizontal transmission greatly diminishes the covariance between organelle and symbiont genomes. This result has two important implications. First, a low level of ID does not indicate low levels of vertical transmission. Second, even with low levels of horizontal transmission, the additive effects of host and symbiont loci will determine the response to selection, while epistatic effects will not be selectable.

THE evolutionary merger of genomes from different species or kingdoms has facilitated major innovations in eukaryotic life (Maynard Smith and Szathmáry 1998; Michod 2000). However, the split control of a phenotype by two genomes generates conflicts and inefficiencies. The mode of symbiont transmission can mediate these tensions (Frank 1994, 1997; Wade 2007). With horizontal transmission, beneficial intergenomic combinations are broken apart, and selection favors traits that maximize the fitness of one partner without regard to the fitness of the other. With vertical transmission, favorable intergenomic combinations are sustained, increasing the response to selection on host–symbiont gene combinations and reducing conflicts between parties.

Although many host–symbiont relationships exhibit complex adaptations to ensure horizontal or vertical transmission, some hosts employ a mixture of both systems of symbiont acquisition (Bright and Bulgheresi 2010). Building on previous studies of multilocus disequilibria for nuclear (Cockerham and Weir 1977; Tachida and Cockerham 1989; Vitalis and Couvet 2001) and cytonuclear (Asmussen and Orive 2000; Orive and Barton 2002; Wade and Goodnight 2006) gene pairs, we quantify the influence of partial horizontal transmission on the covariance between symbiont and organelle genomes. This interspecific disequilibrium (ID) (Sanchez *et al.* 2000) is analogous to linkage disequilibrium (LD) between alleles in one genome.

ID has two important interpretations. First, ID can be used to interpret empirical associations between organelle and symbiont genomes. For example, Hurtado *et al.* (2003) found a strong association between mitochondrial loci of the vent clam, *Calyptogena magnica*, and loci in its sulfur-oxidizing γ-subdivision proteobacterium symbiont. They inferred that this association was evidence for high levels of maternal symbiont transmission. However, the quantitative relationship between genomes as a function of the degree of vertical transmission is currently unclear. If, for example, Hurtado *et al.* (2003) found no association between mitochondrial and symbiont genes, would it be safe to assume that horizontal transmission was predominant or could a model of mixed horizontal and vertical transmission be plausible? A theoretical expectation of the association between host and symbiont genomes will become increasingly important as studies of the population genomics of symbiotic associations become more common (see Wu *et al.* 2006, for a genomic approach to the symbiosis between the sharpshooter, *Homalodisca coagulata*, and two of its bacterial symbionts). Specifically, since the acquisition of genomic data from complex symbioses will likely outpace the experiments and observations traditionally used to characterized these associations, it is worthwhile to understand what can be inferred from sequence data.

doi: 10.1534/genetics.111.130906
Manuscript received May 20, 2011; accepted for publication June 30, 2011
Supporting information is available online at http://www.genetics.org/content/suppl/2011/07/13/genetics.111.130906.DC1.
[1]Corresponding author: Center for Population Biology, One Shields Ave., Davis, CA 95616. E-mail: ybrandvain@gmail.com

Second, when fitness is a function of the interaction between genomes, the response to selection on interspecific genetic combinations is proportional to the maintenance of covariance between genomes, as only heritable combinations can respond to selection (Goodnight 1987, 1988, 1995; Barton and Turelli 2004; Hill *et al.* 2006). For example, Brandvain *et al.* (2007) found that selfing and clonal plants transfer more mitochondrial genes to their nucleus than outcrossing plants. This pattern is consistent with the interpretation that by increasing the heritability of mitochondrial–nuclear gene pairs, both selfing and clonal reproduction facilitate the adaptive movement of mitochondrial genes to the nucleus (Brandvain and Wade 2009). Thus, high levels of interspecific disequilibrium can facilitate complex adaptations of host–symbiont genomes.

We show that without horizontal symbiont acquisition, drift and maternal symbiont transmission can create and maintain a significant association between organelle and symbiont genomes; however, rare horizontal symbiont transmission rapidly erodes most of the host–symbiont association generated by forces such as admixture and selection.

Analytical Results

We describe a dioecious host population of size N, in which N_{φ} individuals are females. In our idealized population, all hosts are singly infected by a single symbiont (however, we relax this assumption in the simulation study, below). Here, we present a series of recursive equations that describe the expected identity coefficients necessary to estimate the covariance between organelle and symbiont loci in the next generation. A single-generation implementation of these recursive equations requires no assumption about the selective effects of host or symbiont loci, so long as identity coefficients in the parental generation are measured after selection and identity measures in the offspring generation precede selection.

To derive identity coefficients within and between genomes, we choose two hosts at random from the population. We find the probability that organelle and/or symbiont genes in these hosts are identical by descent, given their transmission mode. We weight results from each possible pair of transmission modes by their expected frequency (*i.e.*, the probability that both, one, or neither host inherits its symbiont maternally). We present derivations for each case in the *Appendix*.

Regardless of symbiont transmission mode, two individuals share a mother with probability $1/N_{\varphi}$. Thus, organellar genomes of two randomly chosen hosts are identical by descent (IBD) with probability

$$F'_c = \frac{1 + F_c(N_{\varphi} - 1)}{N_{\varphi}} \approx F_c, \tag{1}$$

where the approximation is taken in the case of a large host population.

The probability that symbiont genomes of two randomly chosen individuals are IBD, F_s, depends on transmission mode. Individuals inherit their symbiont maternally with probability v and horizontally with probability $h = 1 - v$. This assumes that the sum of horizontal and vertical transmission equals one (we relax this assumption in our simulations).

When two hosts are both directly infected by their mothers (which occurs with probability v^2), they are IBD at a symbiont locus with probability

$$F'_{s|vv} = \frac{1 + F_s(N_{\varphi} - 1)}{N_{\varphi}} \approx F_s. \tag{2a}$$

When one or both hosts inherit their symbiont horizontally [with probabilities $2v(1 - v)$ and $(1 - v)^2$, respectively], F'_s equals

$$F'_{s|vh} = F'_{s|hh} = \frac{1 + F_s(N - 1)}{N} \approx F_s. \tag{2b}$$

The probability that two randomly chosen individuals are IBD at both organelle and symbiont loci equals θ', which is $F'_c F'_s$ plus the covariance between loci, η' (Vitalis and Couvet 2001; Wade and Goodnight 2006). θ' between two individuals depends on transmission mode and equals

$$\theta'_i = F'_{c|i} F'_{s|i} + \eta'_{|i}, \tag{3a}$$

where i refers to a set of transmission modes for the two randomly selected hosts (*i.e.*, $i = vv$, vh, or hh).

The relevant identity coefficients of organelle and symbiont loci in this model are described above. The covariance between organelle and symbiont loci under each pair of transmission modes, η'_i, equals

$$\eta'_{|vv} = \frac{\eta(N_{\varphi} - 1)}{N_{\varphi}} + \frac{(N_{\varphi} - 1)(1 - F_s)(1 - F_c)}{N_{\varphi}^2} \approx \eta \tag{3b}$$

$$\eta'_{|vh} = \eta \frac{(N_{\varphi} - 1)}{N N_{\varphi}} \approx 0 \tag{3c}$$

$$\eta'_{|hh} = 2\eta \frac{(N_{\varphi} - 1)}{N^2 N_{\varphi}} \approx 0. \tag{3d}$$

Equation 3b demonstrates that vertical transmission can maintain and even generate [*e.g.*, when $F_c = F_s = \eta = 0$, $\eta'_{|vv} = (N_{\varphi} - 1)/N^2_{\varphi} > 0$] an association between organelle and symbiont genes. However, Equations 3c and 3d show that horizontal transmission erodes most of the preexisting intergenomic association.

We derive identity coefficients across the entire population by weighting each identity coefficient (that is, $F'_{c|i}$, $F'_{s|i}$, or $\theta'_{|i}$) associated with each pair of transmission modes, i, by their probabilities, p_i (*e.g.*, $\theta' = \sum p_i \theta'_i = v^2 \theta'_{|vv} + 2vh\theta'_{|vh} + h^2 \theta'_{|hh}$). With these results, we derive the covariance between organelle and symbiont loci, η', by subtracting $F'_c F'_s$ from θ':

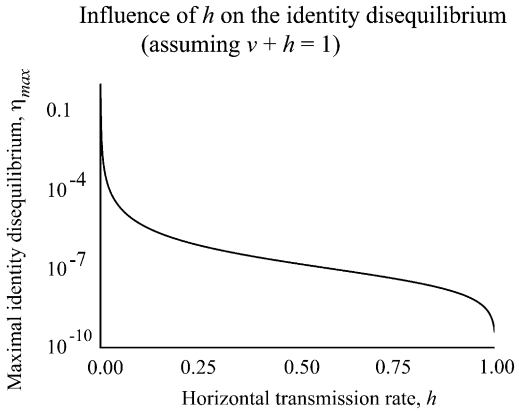

Figure 1 The maximum organelle–symbiont covariance established in a dioecious population of 1000 individuals with an equal sex ratio and no selection.

$$\eta' = \frac{N_{\varphi} - 1}{N_{\varphi}}\left(\frac{2\eta(1-v)(1+v(N-1))}{N^2} + \frac{v^2(\eta N_{\varphi} + (1-F_c)(1-F_s))}{N_{\varphi}}\right) \approx \eta v^2. \tag{4}$$

Thus in a large population, only v^2 of any previous association is maintained across generations.

If there is no change in identity coefficients across generations (*e.g.*, no selection or migration), the solution to the recursion (4) is the expected association between organelle and symbiont loci. Solving this recursion when all initial identity coefficients equal zero, we derive the covariance between organelle and symbiont loci at generation t,

$$\eta_t = \frac{Nv^2\left(\left(2(1-v)(1+v(N-1)) + N^2v^2\right)/N_{\varphi}N^2\right)^t - (1-F_c)(1-F_s)}{N^2v^2 - N(N_{\varphi}-2)(1-v)(1+N+v(N-1))}, \tag{5}$$

where F_c and F_s are the probabilities of IBD at organelle and symbiont loci in generation t. F_c and F_s equal $1-((N_{\varphi}-1)/N_{\varphi})^t$ and $1-((N-1)/N + v^2(1/N_{\varphi} - 1/N))^t$, respectively. In Figure 1, we plot the maximal covariance between organelle and symbiont loci η_{max} against the degree of horizontal transmission. This demonstrates that under neutrality low levels of horizontal transmission prevent the generation of any strong covariance between host and symbiont loci. For example, with no horizontal transmission, $\eta_{max} = 0.25$, but when 1 of 200 hosts are horizontally infected, η_{max} decreases to 0.10.

Simulation Results

Here, we summarize results from a series of stochastic simulations that allow us to both compare the fit between predictions and simulation results and to relax some assumptions of our analytical model. We include the R scripts used for these simulations in supporting information, File S1 and File S2; however, we briefly outline the implementation below.

We begin with N_{φ} clonal females who are genetically unique at both organelle and symbiont loci (note that this is a slight departure from our analytical model of a dioecious population). Each generation, we randomly sample N_{φ} organelle genomes with replacement. v of these females inherit their symbionts from their mothers, while $h_{|\text{uninfected}}$ of the $(1 - v)$ uninfected females inherit a randomly selected symbiont from the previous generation. The case in which $h_{|\text{uninfected}} = 1$ (Figure 2A) approximates our analytical model ($h + v = 1$). By allowing $h_{|\text{uninfected}} < 1$, we can relax our assumption of complete infection (Figure 2B). In a second script, we allow hosts to acquire a second symbiont lineage horizontally. In this set of simulations, offspring that are vertically infected by doubly infected mothers randomly acquire one of their mother's two symbiont lineages. This allows us to relax our assumption of singly infected mothers (Figure 2C).

The qualitative results of simulations under the assumptions of our analytical model (Figure 2A) match expectations: in a given instantiation of our simulation with strict vertical transmission (Figure 2A, blue dots), η approaches 0.25 and then decreases; while lower levels of vertical transmission ($v = 0.99$ and $v = 0.90$ in red and green, respectively) prevent the generation of appreciable levels of interspecific disequilibrium.

By relaxing the assumptions of our model, we show that long-term cotransmission of organelle and symbiont genomes is necessary for an appreciable intergenomic association. With no horizontal transmission, but imperfect vertical transmission, a strong association between organelle and symbiont is generated before the symbiont is lost from the population (Figure 2B, blue). This suggests that the absence of horizontal transmission, rather than the rate of vertical transmission, is the key driver of the organelle–symbiont association. In Figure 2C, we display the case in which individuals can be infected both vertically and horizontally. Here, even when $v = 1$, a symbiont that was horizontally acquired in one generation can be vertically transmitted in the next, breaking apart organelle–symbiont lineages.

In sum, these stochastic simulations support our major analytical results: a low level of horizontal transmission prevents the generation of strong associations between organelle and symbiont genomes.

Discussion

The relationship between host and symbiont is fragile and can range from mutualistic to parasitic. It is thought that by reducing conflict between host and symbiont, and maintaining adaptive intergenomic combinations, vertical transmission can result in an intimate and interdependent relationship between species.

Our results show that the absence of horizontal transmission is necessary for the transmission mode to maintain a meaningful association between genomes. With even low levels of horizontal transmission, only additive effects of host and symbiont loci will respond to selection. Additionally, rare horizontal transmission may require the evolution of

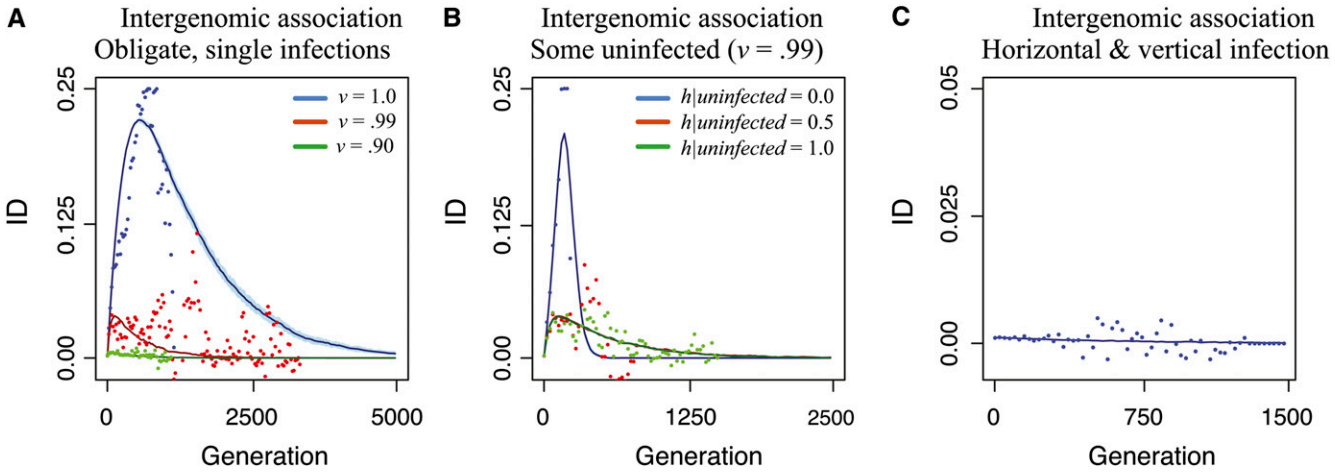

Figure 2 The generation of interspecific disequilibria over time from simulation models. In all cases, lines represent means of 1000 simulations, in a population of 1000 clonal females. Shading represents bootstrapped 95% confidence intervals of the mean, and dots represent a single simulation. (A) All uninfected hosts are horizontally infected: $v = 1$, $v = 0.99$, $v = 0.90$. (B) Not all hosts are infected: $v = 0.99$, $h_{\text{uninfected}} = 0$, $h_{\text{uninfected}} = 0.5$, $h_{\text{uninfected}} = 1.0$. (C) All individuals inherit one symbiont vertically and another horizontally ($h = 1$, $v = 1$).

mechanisms to reduce conflict [*e.g.*, host sanctions (Denison 2000; Kiers *et al.* 2003)] to prevent the exploitation of the host–symbiont relationship.

Even in the absence of horizontal transmission an opportunity for conflict between host and symbiont genomes remains because symbionts do not generally gain from host fitness via male function (Frank and Hurst 1996; Hurst *et al.* 1996; Gemmell *et al.* 2004; but see Unckless and Herren 2009; Wade and Brandvain 2009). This can lead to the evolution of symbiotic manipulators of host reproduction. The expected association between mitochondria and cytoplasmic incompatibility-inducing *Wolbachia* in a biallelic system have been derived by Turelli *et al.* (1992). Like the case detailed herein, associations between *Wolbachia* and a plasmid with an arbitrary degree of maternal transmission rapidly decrease without strict cotransmission (Turelli and Hoffmann 1999).

Our results aid in interpretation of natural patterns of disequilibrium between mitochondrial and symbiont loci. A measurable interspecific disequilibrium is strong evidence for the absence of horizontal symbiont transmission; however, no association between organelle and symbiont does not indicate that vertical transmission is infrequent.

As only $\approx(1 - h)^2$ of the identity disequilibrium of the previous generation is transmitted to the next generation, our results demonstrate that the response to selection on host–symbiont interactions will be weak and that the response to selection will be dominated by the additive effects of each locus. Nonetheless, cases of highly specialized host–symbiont relationships are observed in the absence of vertical transmission (Bright and Bulgheresi 2010), suggesting that selection on additive effects of host and symbiont loci is sufficient for the evolution of many specialized host–symbiont relationships.

Future work is needed to address the ability of strong selection to maintain host–symbiont associations over many generations in the face of horizontal transmission breaking apart these associations. Given that imperfect vertical transmission without horizontal infection can generate high levels of interspecific disequilibrium, it seems likely that under a model where selection maintains an infected class of hosts, incomplete vertical transmission without horizontal transmission will generate high ID. Another promising avenue of future research is to include more of the biological complexity of symbiont transmission.

Although we do not incorporate the entirety of complications associated with host–symbiont relationships, our result is clear: with even modest levels of horizontal symbiont acquisition, neither vertical transmission nor genetic drift can either generate or maintain a meaningful association between organelle and symbiont genomes. Our result also motivates a few related conjectures: since nuclear alleles in an outbreeding population are not strictly transmitted with maternally inherited symbionts, we predict no association between nuclear and symbiont genes in outbreeding, panmictic populations, unless selection is exceedingly strong. Since rare horizontal transmission breaks apart intraspecific genetic associations, we speculate that even low levels of interspecific symbiont transmission may erode interspecific associations between mitochondria and symbionts (*e.g.*, Stewart *et al.* 2008).

Acknowledgments

The manuscript was improved by comments from J. D. Van Dyken, G. Coop, M. Turelli, P. Ralph, A. Sedghifar, M. May, H. Spencer, Maria Orive, and two anonymous reviewers. This work was supported by National Institutes of Health grant R01-GM084238 (to M.J.W.).

Literature Cited

Asmussen, M. A., and M. E. Orive, 2000 The effects of pollen and seed migration on nuclear-dicytoplasmic systems. I. Nonrandom

associations and equilibrium structure with both maternal and paternal cytoplasmic inheritance. Genetics 155: 813–831.
Barton, N. H., and M. Turelli, 2004 Effects of genetic drift on variance components under a general model of epistasis. Evolution 58: 2111–2132.
Brandvain, Y., and M. J. Wade, 2009 The functional transfer of genes from the mitochondria to the nucleus: the effects of selection, mutation, population size and rate of self-fertilization. Genetics 182: 1129–1139.
Brandvain, Y., M. S. Barker, and M. J. Wade, 2007 Gene coinheritance and gene transfer. Science 315: 1685.
Bright, M., and S. Bulgheresi, 2010 A complex journey: transmission of microbial symbionts. Nat. Rev. Microbiol. 8(3): 218–230.
Cockerham, C. C., and B. Weir, 1977 *Two-Locus Theory in Quantitative Genetics*. Iowa University Press, Ames, IA, pp. 247–269.
Denison, R. F., 2000 Legume sanctions and the evolution of symbiotic cooperation by rhizobia. Am. Nat. 156: 567–576.
Frank, S. A., 1994 Genetics of mutualism: the evolution of altruism between species. J. Theor. Biol. 170: 393–400.
Frank, S. A., 1997 Models of symbiosis. Am. Nat. 150: S80–S99.
Frank, S. A., and L. D. Hurst, 1996 Mitochondria and male disease. Nature 383: 224.
Gemmell, N. J., V. J. Metcalf, and F. W. Allendorf, 2004 Mother's curse: the effect of mtDNA on individual fitness and population viability. Trends Ecol. Evol. 19: 238–244.
Goodnight, C. J., 1987 On the effect of founder events on epistatic genetic variance. Evolution 41: 80–91.
Goodnight, C. J., 1988 Epistasis and the effect of founder events on the additive genetic variance. Evolution 42: 441–454.
Goodnight, C. J., 1995 Epistasis and the increase in additive genetic variance: implications for phase 1 of Wright's shifting balance process. Evolution 49: 502–511.
Hill, W. G., N. H. Barton, and M. Turelli, 2006 Prediction of effects of genetic drift on variance components under a general model of epistasis. Theor. Popul. Biol. 70: 56–62.
Hurst, L. D., A. Atlan, and B. O. Bengtsson, 1996 Genetic conflicts. Q. Rev. Biol. 71: 317–364.
Hurtado, L. A., M. Mateos, R. A. Lutz, and R. C. Vrijenhoek, 2003 Coupling of bacterial endosymbiont and host mitochondrial genomes in the hydrothermal vent clam Calyptogena magnifica. Appl. Environ. Microbiol. 69: 2058–2064.
Kiers, E. T., R. A. Rousseau, S. A. West, and R. F. Denison, 2003 Host sanctions and the legume-rhizobium mutualism. Nature 425: 78–81.
Maynard Smith, J., and E. Szathmáry, 1998 *The Major Transitions in Evolution*. Oxford University Press, New York.
Michod, R., 2000 *Darwinian Dynamics*. Princeton University Press, Princeton, NJ.
Orive, M. E., and N. H. Barton, 2002 Associations between cytoplasmic and nuclear loci in hybridizing populations. Genetics 162: 1469–1485.
Sanchez, M. S., J. Arnold, and M. A. Asmussen, 2000 Symbiont survival and host-symbiont disequilibria under differential vertical transmission. Genetics 154: 1347–1365.
Stewart, F. J., C. R. Young, and C. M. Cavanaugh, 2008 Lateral symbiont acquisition in a maternally transmitted chemosynthetic clam endosymbiosis. Mol. Biol. Evol. 25: 673–687.
Tachida, H., and C. C. Cockerham, 1989 Effects of identity disequilibrium and linkage on quantitative variation in finite populations. Genet. Res. 53: 63–70.
Turelli, M., and A. A. Hoffmann, 1999 Microbe-induced cytoplasmic incompatibility as a mechanism for introducing transgenes into arthropod populations. Insect Mol. Biol. 8: 243–255.
Turelli, M., A. A. Hoffmann, and S. W. McKechnie, 1992 Dynamics of cytoplasmic incompatibility and Mtdna variation in natural *Drosophila simulans* populations. Genetics 132: 713–723.
Unckless, R. L., and J. K. Herren, 2009 Population genetics of sexually antagonistic mitochondrial mutants under inbreeding. J. Theor. Biol. 260: 132–136.
Vitalis, R., and D. Couvet, 2001 Two-locus identity probabilities and identity disequilibrium in a partially selfing subdivided population. Genet. Res. 77: 67–81.
Wade, M. J., 2007 The co-evolutionary genetics of ecological communities. Nat. Rev. Genet. 8: 185–195.
Wade, M. J., and Y. Brandvain, 2009 Reversing mother's curse: selection on male mitochondrial fitness effects. Evolution 63: 1084–1089.
Wade, M. J., and C. J. Goodnight, 2006 Cyto-nuclear epistasis: two-locus random genetic drift in hermaphroditic and dioecious species. Evolution 60: 643–659.
Wu, D., S. C. Daugherty, S. E. Van Aken, G. H. Pai, K. L. Watkins *et al.*, 2006 Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. PLoS Biol. 4: e188.

Communicating editor: H. G. Spencer

Appendix

Consider a dioecious host species with an obligately maternally transmitted organelle and some combination of maternal and horizontal transmission (from both males and females) of an associated symbiont. We assume that all hosts are singly infected by this symbiont. A population of N reproducing hosts consists of $N_{♀}$ females and $N_{♂} = N - N_{♀}$ males. We assume that both host and symbiont loci are neutral and that no segregating loci influence the probability of vertical transmission. Let v be the frequency of vertical transmission of microbe directly from host mother to offspring and let the frequency of horizontal transmission equal $h = 1 - v$. Since the infecting individual is drawn at random under horizontal transmission, the source of offspring contagious infection could be its own mother with probability $1/N$.

Two hosts, A and B, are sampled at random from a population. Each host is genotyped at both organelle and symbiont loci and assigned a two-locus genotype. With respect to the host organelle gene, hosts A and B are IBD with probability 1, if they have the same mother ($M_A = M_B$, which occurs with probability $1/N_{♀}$), and with probability F_c (the probability of being IBD at the organelle locus in the previous generation), if they have different mothers ($M_A \neq M_B$, which occurs with probability $(N_{♀}-1)/N_{♀}$). The symbionts of these two hosts are IBD with probability 1, if they were transmitted by the same host ($I_A = I_B$), and with probability F_s if they were transmitted by different hosts ($I_A \neq I_B$).

Occasionally, two randomly sampled hosts will be identical by descent at both organelle and symbiont loci. This is the probability of double identity by descent (DIBD), which we summarize with the symbol θ. The expectation for θ when

the identity at cytoplasmic and symbiotic loci is independent is simply the product of F_c and F_s. The deviation from this expectation (*i.e.*, $\theta - F_cF_s$) is the covariance between organelle and symbiont loci, represented by the symbol, η.

Consider the transmission of symbionts to two individuals. Both could have inherited their symbionts vertically (*i.e.*, maternally), with probability v^2, one may be vertically infected while the other is horizontally infected with probability $2v(1 - v)$, or both could be horizontally infected with probability, $(1 - v)^2$. Below we consider each of these three situations (cases i–iii) separately. We then turn to the general case, in which we derive F_c, F_s, and θ, by summing the expected values of each case weighted by the likelihood of each case. With those results, we can derive η by subtracting the product of F_c and F_s from θ. As our analysis addresses the evolution of identity measures in the absence of selection and mutation, F_c, F_s, and θ approach 1 as $t \to \infty$. Since $\eta = \theta - F_cF_s = \theta$, $\eta \to 0$, as $t \to \infty$.

Case i: Both hosts acquired their symbiont vertically (vv)

When two randomly sampled hosts, A and B, are both vertically infected, the probability of DIBD depends upon whether they share the same mother. A and B share a mother (with probability $1/N_{\varphi}$), or they do not (with probability $1 - 1/N_{\varphi}$). With strict maternal transmission, if two hosts share a mother, they are identical at both organelle and symbiont loci and are therefore DIBD. In this case, the contribution to the identity measures in the next generation, F'_c, F'_s, and $\theta' = 1$. By contrast, if A and B do not share a mother, they are not immediately IBD for either organelle or symbiont gene. However, even if they do not share a mother, hosts A and B can be IBD at organelle or symbiont loci if their mothers were IBD at their organelle or symbiont loci (with probabilities F_c and F_s, respectively). Similarly, as individuals A and B each inherited both symbiont and organelle genes as a unit, the probability of DIBD = θ, the probability of DIDB for two randomly drawn, infected hosts in the previous generation. From these relations, we derive the recursive equations for the identity coefficients:

$$F'_{c|vv} = \frac{(1 + F_c(N_{\varphi} - 1))}{N_{\varphi}} \tag{A1a}$$

$$F'_{s|vv} = \frac{(1 + F_s(N_{\varphi} - 1))}{N_{\varphi}} \tag{A1b}$$

$$\theta'_{|vv} = \frac{(1 + \theta(N_{\varphi} - 1))}{N_{\varphi}}. \tag{A1c}$$

Note that when $F_c = F_s$, Equations A1a and A1b are also equal because both have been vertically transmitted by females to offspring.

The covariance in identity by descent, η', equals

$$\eta' = \mathrm{Cov}\left(F'_{c|vv}, F'_{s|vv}\right) = \theta' - F'_cF'_s. \tag{A2a}$$

Note that η' is always ≥ 0, as the probability of DIBD (θ) is greater than or equal to the product, $F'_cF'_s$—the expected probability of DIBD when haplotypes are randomly associated. When transmission is strictly vertical (*i.e.*, $v = 1$), the recursion for η' equals

$$\eta'_{|vv} = \frac{N_{\varphi} - 1}{N_{\varphi}}\left(\eta + \frac{(1 - F_s)(1 - F_c)}{N_{\varphi}}\right). \tag{A2b}$$

Equation A2b shows that when both hosts inherit their symbionts maternally, η' increases from zero at a rate proportional to $1/N_{\varphi}$.

The contribution of case i to the change in η is

$$\Delta\eta_{|vv} = \frac{(N_{\varphi} - 1)(1 - F_c)(1 - F_s)}{N_{\varphi}^2} - \frac{\eta}{N_{\varphi}}. \tag{A2c}$$

Equation A2c shows that with strict vertical transmission, $\Delta\eta$ can be positive. As a result, vertical transmission can generate an interspecific genetic association between host organelle and symbiont alleles. The contrast to cases with some degree of horizontal transmission is striking, as we show below.

Case ii: One host's symbiont is vertically transmitted while the other's is horizontally transmitted (vh)

If one host has been infected vertically but the other has been infected horizontally, the probability of DIBD is different from that derived above. We (arbitrarily) label host A as the individual who inherits its symbiont maternally. The probability of the five outcomes (below) sums to one.

1. Hosts A and B may have the same mother as hosts A and B may have been infected horizontally by contact with A's mother with probability $(1/NN_{\varphi})$. The contributions of this case to the identity coefficients, F'_c, F'_s, and θ', are 1, 1, and 1, respectively.
2. Hosts A and B may be maternal siblings but B contracted the symbiont from a host other than A's mother with probability $(N-1)/NN_{\varphi}$. The contributions of this case to the identity coefficients, F'_c, F'_s, and θ', are 1, F_s, and F_s, respectively.
3. Hosts A and B may be offspring of different mothers, but B was infected by its mother, with probability $(N_{\varphi}-1)/NN_{\varphi}$. The contributions of this case to the identity coefficients, F'_c, F'_s, and θ', are F_c, F_s, and θ, respectively.
4. Hosts A and B may be offspring of different mothers but B was infected by A's mother, also with probability $(N_{\varphi}-1)/NN_{\varphi}$. The contributions of this case to the identity coefficients, F'_c, F'_s, and θ', are F_c, 1, and F_c, respectively.
5. A and B have different mothers and B was infected by an individual other than either of their mothers with probability $(N-2)(N_{\varphi}-1)/NN_{\varphi}$. The contributions of this case to the identity coefficients, F'_c, F'_s, and θ', are F_c, F_s, and F_cF_s, respectively.

Table A1 The number of ways and likelihoods that A and B inherit organelle genes and contract the pathogen when both contract it horizontally (case iii, above)

Description	Probability	F'_c	F'_s	θ'
$M_A = M_B = S_A = S_B$	$\frac{1}{N^2 N_♀}$	1	1	1
$M_A = M_B = S_A \neq S_B$	$\frac{N-1}{N^2 N_♀}$	1	F_s	F_s
$M_A = M_B = S_B \neq S_A$	$\frac{N-1}{N^2 N_♀}$	1	F_s	F_s
$M_A = M_B \neq S_A = S_B$	$\frac{N-1}{N^2 N_♀}$	1	1	1
$M_A \neq M_B = S_A = S_B$	$\frac{(N-1)(N-2)}{N^2 N_♀}$	1	F_s	F_s
$M_A = S_A = S_B \neq M_B$	$\frac{(N_♀-1)}{N^2 N_♀}$	F_c	1	F_c
$M_A \neq M_B = S_A = S_B$	$\frac{(N_♀-1)}{N^2 N_♀}$	F_c	1	F_c
$M_A = S_A \neq M_B = S_B$	$\frac{(N_♀-1)}{N^2 N_♀}$	F_c	F_s	θ
$M_A = S_B \neq M_B = S_A$	$\frac{(N_♀-1)}{N^2 N_♀}$	F_c	F_s	θ
$M_A = S_A \neq M_B \neq S_B$	$\frac{(N_♀-1)(N-2)}{N^2 N_♀}$	F_c	F_s	$F_c F_s$
$M_A \neq S_A \neq M_B = S_B$	$\frac{(N_♀-1)(N-2)}{N^2 N_♀}$	F_c	F_s	$F_c F_s$
$M_A \neq M_B \neq S_A = S_B$	$\frac{(N_♀-1)(N-2)}{N^2 N_♀}$	F_c	1	F_c
$M_A \neq S_B \neq M_B = S_A$	$\frac{(N_♀-1)(N-2)}{N^2 N_♀}$	F_c	F_s	$F_c F_s$
$M_A = S_B \neq M_B \neq S_A$	$\frac{(N_♀-1)(N-2)}{N^2 N_♀}$	F_c	F_s	$F_c F_s$
$M_A \neq M_B \neq S_A \neq S_B$	$\frac{(N_♀-1)(N-2)(N-3)}{N^2 N_♀}$	F_c	F_s	$F_c F_s$

Note that the probabilities of outcomes with $S_A = S_B$ sum to $1/N$, the probabilities of outcomes with $M_A = M_B$ sum to $1/N_♀$, the probabilities of outcomes with $M_A \neq M_B$ sum to $(N_♀-1)/N_♀$, and the probabilities with $S_A \neq S_B$ sum to $(N-1)/N$. $M_{A(B)}$ refers to the mitochondrial donor of individual A (or B). $S_{A(B)}$ refers to the symbiont donor of individual A (or B).

Note that outcome 3, the case in which B indirectly inherits its mother's symbiont, is the only one in which the contribution to θ' differs from the product of F_c and F_s and is therefore the sole contributor to $\eta_{|vh}$. By taking the weighted sum of each outcome's contribution to F'_c, we derive $F'_{c|vh}$, which is equivalent to equation (A1a). That is, as expected, the probability of IBD for the organelle gene is completely independent of the symbiont's mode of transmission.

$F'_{s|vh}$ is the weighted sum of the symbiont gene's probability of IBD, (F'_s) across all cases, and equals

$$F'_{s|vh} = \frac{1 + F_s(N-1)}{N}. \tag{A3}$$

Note that this result differs from case i. Here, F'_s is dependent on the total population size, N, rather than on the number of infected mothers, $N_♀$. Finally, the probability that both host organelle and symbiont loci are DIBD, $\theta'_{|vh}$, is

$$\theta'_{s|vh} = \left(1 + F_c\frac{N_♀ - 1}{N_♀}\right)\left(1 + F_s\frac{N-1}{N}\right) + \eta\frac{N_♀ - 1}{NN_♀}. \tag{A4}$$

Note that Equation A4 equals the product $F'_{c|vh}$ and $F'_{s|vh}$ plus a small portion $(N_♀-1)/NN_♀$ of η from the previous generation. The change in η across generations equals $\Delta\eta_{|vh} = -\eta(1-(N_♀-1)/N_♀)$. For any reasonably sized population, $-((N_♀-1)/NN_♀) \approx -1/N$, and $-\eta(1-1/N) \approx -\eta$. This clearly demonstrates that in this case, most of η is lost in one generation. Thus, although vertical transmission can increase DIBD, horizontal transmission effectively destroys it. Note that this result is more severe than the semiconservative process of free recombination between two nuclear alleles, in which disequilibria between loci are decreased by only $\frac{1}{2}$.

Case iii: Both hosts infected horizontally (hh)

If individuals A and B have both been infected horizontally, there are 15 possible relationships between them (Table A1). The identity coefficients, $F_{c|hh}$ and $F_{s|hh}$, for case iii are equivalent to those derived for case ii (*i.e.*, $F_{c|hh}$ = Equation 1, and $F_{c|hh}$ = Equation A3). However, in this case, θ and η differ from those above.

We derive $\theta_{|hh}$ by summing across all possibilities with each weighted by its contribution to $\theta_{|hh}$ (*i.e.*, $\Sigma p_i\theta_i$). Overall, we find that the probability of DIBD is

$$\theta'_{s|hh} = \left(1 + F_c\frac{N_♀ - 1}{N_♀}\right)\left(1 + F_s\frac{N-1}{N}\right) + 2\eta\frac{N_♀ - 1}{N^2N_♀}. \tag{A5}$$

Note that, like Equation A4, Equation A5 equals the product of the recursive equations for $F_{c|hh}$ and $F_{s|hh}$, plus a small fraction of η from the previous generation [in this case, $2\eta((N_♀-1)/N^2N_♀)$]. Again, this deviation from the product of $F_{c|hh}$ and $F_{s|hh}$ is the interspecific disequilibrium, $\eta_{hh,}$ and again $\Delta\eta \approx -\eta$ when N is large. Again, in this case, any covariance in identity-by-descent established by processes such as migration and selection will vanish instantaneously.

The general case

We now use the weighted sum of probabilities for cases i–iii, above [$vv = v^2$, $vh = 2v(1-v)$, and $hh = (1-v)^2$] to calculate the recursions for F'_c, F'_s, θ', and η'· As F_c it equals Equation A1a (Equation 1 in the main text). When F_c at $t = 0$ equals $F_{c(0)}$, the recursive solution for F_c at time t equals

$$F_{c(t)} = 1 - \left(1 - F_{c(0)}\right)\left(1 - \frac{1}{N_♀}\right)^t. \tag{A6}$$

This is the classic recursion equation for a maternally inherited organelle in a dioecious population. Its inheritance is unaffected by the presence of a symbiont or by the symbiont's mode of transmission.

The probability of identity by descent at a symbiont locus, F'_s, depends on its transmission mode:

$$F_s' = F_s + (1 - F_s)\left(\frac{v^2}{N_\varphi} + \frac{1 - v^2}{N}\right) \tag{A7a}$$

$$\Delta F_s = (1 - F_s)\left(\frac{v^2}{N_\varphi} + \frac{1 - v^2}{N}\right). \tag{A7b}$$

When F_s at $t = 0$ equals $F_{s(0)}$, the recursive solution for Equation A7a equals

$$F_{s(t)} = 1 - F_{s(0)}\left(1 - \frac{1}{N} - \frac{v^2}{N_\varphi} - \frac{v^2}{N}\right)^t. \tag{A7c}$$

When there is an equal sex ratio (*i.e.*, $N_\varphi = N/2$), Equations A7b and A7c, respectively, become

$$\Delta F_{s|\{N_\varphi = N/2\}} = (1 - F_s)\left(\frac{1 + v^2}{N}\right) \tag{A7d}$$

$$F_{s(t)|\{N_\varphi = N/2\}} = 1 - \left(1 - F_{s(0)}\right)\left(1 - \frac{1 + v^2}{N}\right)^t. \tag{A7e}$$

With entirely horizontal transmission ($v = 0$), the coefficient of Equation A7e reduces to $1/N$, which is equivalent to the change in IBD of mitochondrial or chloroplast genes in a population of N hermaphrodites. With entirely vertical transmission ($v = 1$), the coefficient of Equation A7e reduces to $1/N_\varphi$, the recursion for a maternally inherited gene in a dioecious population (see Equation A6). Because $N > N_\varphi$, any degree of vertical transmission increases the probability of IBD of symbionts (F_s) in hosts randomly drawn from the population. Nevertheless, large values of v or small values of N are required for vertical transmission to have a significant effect on ΔF_s.

We now turn to the probability of DIBD θ. In a population in which symbionts are maternally transmitted with probability v, we find that

$$\begin{aligned}\theta' = {} & (v^2)\frac{1 + \theta(N_\varphi - 1)}{N_\varphi} + 2\eta\frac{(N_\varphi - 1)(1 - v)(1 + v(N - 1))}{N^2 N_\varphi} \\ & + (1 - v^2)\frac{(1 + F_s(N - 1))(1 + F_c(N_\varphi - 1))}{N_\varphi N}.\end{aligned} \tag{A8}$$

With the results above, we derive η' as a function of the vertical transmission rate, v:

$$\eta' = \frac{N_\varphi - 1}{N_\varphi}\left(2\eta\frac{(1 - v)(1 + v(N - 1))}{N^2} + v^2\,\eta + \frac{(1 - F_c)(1 - F_s)}{N_\varphi}\right). \tag{A9}$$

The recursive solution for η is presented in Equation 5 of the main text.

Note that vertical transmission increases θ in two ways. First, maternal transmission increases F_s by reducing the number of potential host donors of symbionts from N to N_φ as v approaches 1. For this reason, increased vertical transmission increases the product of F_c and F_s, thereby increasing θ. Second, by coupling the transmission of organelle and symbiont genes, vertical transmission increases θ by generating a disequilibrium, η between genotypes.