Technical comment on 'Negative-assortative mating for color in wolves'

Hedrick *et al.* (2016) reported on "negative-assortative mating for color in wolves" from Yellowstone National Park, the "first documented case of significant negative-assortative mating in mammals." Based on the close correspondence of genotype and allele frequencies observed in the wild to that predicted by their population genetic model, they conclude that "negative-assortative mating could be entirely responsible for the maintenance of this well-known color polymorphism." While researching examples of nonrandom mating in the wild to teach in class I discovered a mistake in their model. The mathematical error does not substantially alter their inference because the equilibrium genotype and allele frequencies are similar in both. However, it is important that the mathematical biology literature provide correct and logically consistent analysis so that future researchers may benefit most from its insights.

Table 1: Glossary of mathematical symbols, variable string used in source code, and description.

Symbol	Variable string	Description
k	k	recessive beta-defensin variant
K	K	dominant beta-defensin variant
p	p	frequency of k allele
q	q	frequency of K allele
P	P	frequency of kk genotype
H	H	frequency of Kk genotype
Q	Q	frequency of KK genotype
A	A	proportion negative-assortatively mating

For consistency, I use the same symbols as Hedrick et al. (2016) (Table 1 lists all symbols and their definitions). I used Sympy version 1.7.1 (Meurer et al. 2017) for symbolic derivations through Python version 3.6 and the *R* package **reticulate** version 1.24 (Ushey et al. 2022). All other computations were performed in *R* version 4.1.2 (R Core Team 2021). I have provided my source code for review and will deposit my code in public GitHub and Zenodo repositories upon publication.

In Hedrick et al. (2016), the frequency of gray × black matings is incorrectly written as $2P(H + Q)$ (*cf* Table 1). However, this value does not account for all possible outcomes that result in Gray × black matings (Table 2). As a result, the genotype frequencies (Table 3) do not sum to 1 as they should if they account for all possible outcomes. Ironically, I found an analogous derivation in Hedrick and Ritland (2012) for positive-assortative mating where the genotype frequencies do sum to 1 when ignoring other evolutionary forces such as selection. The correct expression derived by summing all ways gray × black matings can occur

24 is provided in Table 3. The corrected genotype frequencies sum to 1 as expected (see code for analytical
25 derivation in Supporting Information).

Table 2: The probability of every mating outcome in the negative-
assortative mating model analyzed by Hedrick *et al.* (2016).

Parent 1	Pr[Parent 1]	Mating	Pr[Mating]	Parent 2	Pr[Parent 2]	Pr[Total]	Color
kk	P	assortative	A	kk	0	0	Gray $\times$ gray
kk	P	random	$(1-A)$	kk	P	$P^2(1-A)$	Gray $\times$ gray
kk	P	assortative	A	$K-$	1	AP	Gray $\times$ black
kk	P	random	$(1-A)$	$K-$	$(1-P)$	$P(A-1)(P-1)$	Gray $\times$ black
$K-$	$(1-P)$	assortative	A	kk	1	$A(1-P)$	Gray $\times$ black
$K-$	$(1-P)$	random	$(1-A)$	kk	P	$P(A-1)(P-1)$	Gray $\times$ black
$K-$	$(1-P)$	assortative	A	$K-$	0	0	Black $\times$ black
$K-$	$(1-P)$	random	$(1-A)$	$K-$	$(1-P)$	$(1-A)(P-1)^2$	Black $\times$ black

Table 3: Hedrick *et al.* (2016) incorrectly derive the frequency of
gray $\times$ black. The corrected expressions are provided here.

Color	Mating Genotypes	Frequency (Hedrick *et al.* 2016)	Frequency (this paper)
Gray $\times$ grey	$kk \times kk$	$P^2(1-A)$	$P^2(1-A)$
Gray $\times$ black	$kk \times K-$	$2AP(H+Q)$	$AP - A(H+Q) - 2P(1-A)(H+Q)$
Black $\times$ black	$K- \times K-$	$(H+Q)^2(1-A)$	$(H+Q)^2(1-A)$

26 Hedrick et al. (2016) account for the fact that genotype frequencies do not sum for 1 by regularizing the
27 frequencies (*cf* equation 1a-b), as normally done in models of selection. In effect, by not accounting for all
28 possible outcomes, they are accidentally assuming a type of selection. However, the equilibrium genotype
29 frequencies they derive are very similar to the correct equilibrium. In both models, $\hat{P} = 0.5$, implying
30 $0.5 = \hat{H} + \hat{Q}$. I find that $\hat{Q} = (-A/2 + \sqrt{2(A+1)} - 1.5)/(A-1)$, which is close to the equilibrium values
31 obtained in Hedrick et al. (2016) through recursion (Fig. 1).

32 In conclusion, the mathematical errors in Hedrick et al. (2016) do not undermine their primary conclusion
33 that negative-assortative mating by color may explain the distribution of genotype frequencies at the beta
34 definsin locus in the Yellowstone population of wolves (*Canis lupus*). The derivation here may prove useful to
35 future research on detecting evidence for negative-assortative mating.

Literature cited

37 Hedrick, P. W., and K. Ritland. 2012. Population genetics of the white-phased "Spirit" bear of British
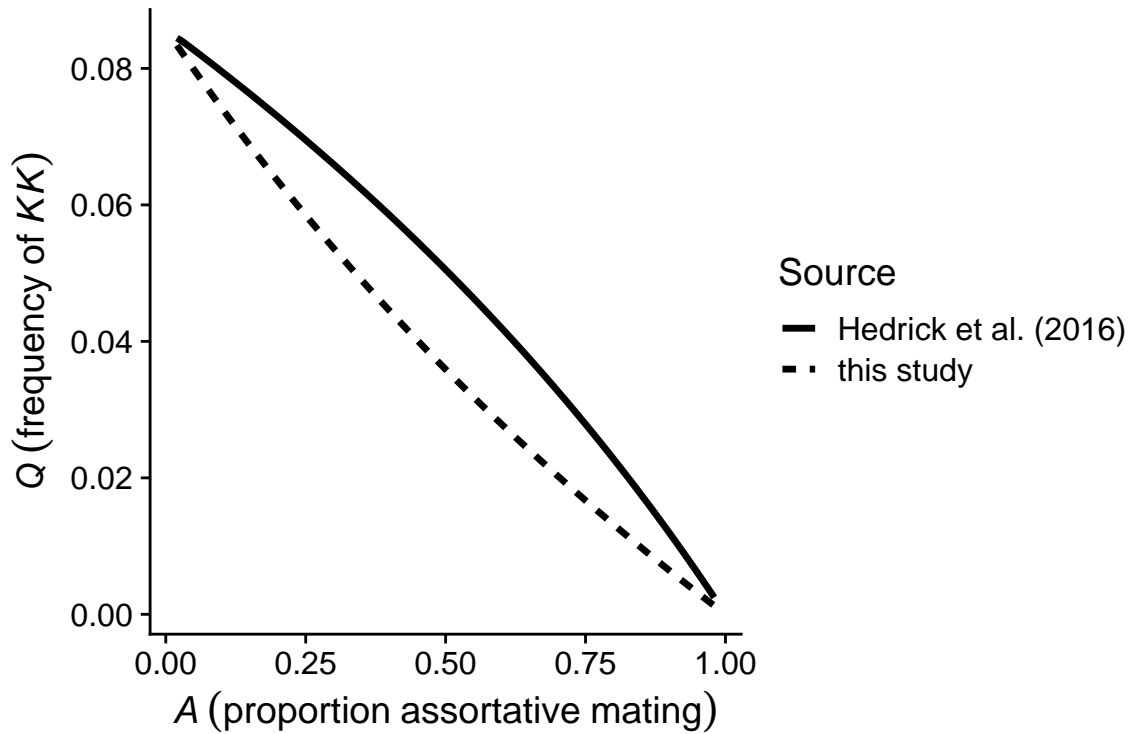38   Columbia. Evolution 66:305–313.

Figure 1: The equilibirum frequency of Q, the KK homozygote in this study (dashed line) and Hedrick *et al.* (2016) (solid line) for possible values of A, the proportion of wolves mating assortatively by color.

39 Hedrick, P. W., D. W. Smith, and D. R. Stahler. 2016. Negative-assortative mating for color in wolves.
40   Evolution 70:757–766.

41 Meurer, A., C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K.
42   Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats,
43   F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, Š. Roučka, A. Saboo, I. Fernando, S. Kulal, R.
44   Cimrman, and A. Scopatz. 2017. SymPy: Symbolic computing in Python. PeerJ Computer Science
45   3:e103.

46 R Core Team. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical
47   Computing, Vienna, Austria.

48 Ushey, K., J. J. Allaire, and Y. Tang. 2022. Reticulate: Interface to 'Python'.