Supporting Information

Hohenlohe and Arnold 10.1073/pnas.1003537107

SI Text

Mate-choice model. The univariate model of Arnold et al. (1) envisions mating probability as a function of a single male trait and a single female preference. In this model, the probability of mating between a female randomly chosen from population i and a male from population j is (Eq. **S1**)

$$\pi_{ij} = c_{ij} \exp\left(-\left(\bar{z}_j - \bar{y}_i\right)^2 / 2\left(\tau_i^2 + \nu_i^2 + \sigma_j^2\right)\right). \qquad \textbf{[S1]}$$

Following the notation from ref. 1, we represent the combined variance term in this equation as $\sigma_{ij}^2 = \tau_i^2 + \nu_i^2 + \sigma_j^2$.

To extend this model to d dimensions, let $\mathbf{y}_i^*$ and $\mathbf{z}_j^*$ be column vectors of d female preference values and d corresponding male traits, respectively. Let mating probability again be given by an absolute preference function, now multivariate Gaussian with (co)variance N_i^2 (analogous to ν_i^2 in the univariate model, using uppercase letters to denote matrices). Thus, the probability of mating between a single male $\mathbf{z}^*$ and a single female $\mathbf{y}^*$ is (Eq. **S2**)

$$\psi(\mathbf{z}^* | \mathbf{y}^*) \propto \exp\left(-\frac{1}{2}(\mathbf{z}^* - \mathbf{y}^*)^T \left(\mathsf{N}^2\right)^{-1} (\mathbf{z}^* - \mathbf{y}^*)\right). \qquad \textbf{[S2]}$$

[see equation 2 in Arnold et al. (1)]. We assume that these traits are multivariate normally distributed within populations with means $\bar{\mathbf{y}}_i^*$ and $\bar{\mathbf{z}}_j^*$ and phenotypic variance–covariance matrices T_i^2 and Σ_j^2 (analogous to τ_i^2 and σ_j^2). Now, the probability of mating between a female randomly chosen from population i and a male from population j is (Eq. **S3**)

$$\pi_{ij} = c_{ij} \exp\left(-\frac{1}{2}\left(\bar{\mathbf{z}}_j^* - \bar{\mathbf{y}}_i^*\right)^T \left(\Sigma_{ij}^2\right)^{-1} \left(\bar{\mathbf{z}}_j^* - \bar{\mathbf{y}}_i^*\right)\right), \qquad \textbf{[S3]}$$

where $\Sigma_{ij}^2 = (\mathsf{T}_i^2 + \mathsf{N}_i^2 + \Sigma_j^2)$ [see equations 5 and 6 in Arnold et al. (1)].

Note that, under our model, male traits may covary with each other and female preferences may covary with each other; we do not require the covariance structure of male traits and female preferences to be similar in any way within populations. Furthermore, the covariance structure of the preference function in Eq. **S2** may be different from that of either male traits or female preferences.

However, the key assumption required for our analysis is that the combined (co)variance term $\Sigma_{ij}^2 = (\mathsf{T}_i^2 + \mathsf{N}_i^2 + \Sigma_j^2)$ is equal across all population comparisons (females from population i vs. males from population j). This is generally satisfied by assuming that T_i^2 and N_i^2 are each constant across all populations i and Σ_j^2 is constant across all populations j, and therefore, Σ_{ij}^2 is equal across all i and j. Because Σ_{ij}^2 is a variance/covariance matrix, it can be decomposed into a set of orthogonal eigenvectors with nonnegative eigenvalues. We use these eigenvectors and eigenvalues to scale and rotate $\mathbf{y}_i^*$ and $\mathbf{z}_j^*$ into a new set of d latent traits $\mathbf{y}_\mathrm{i}$ and $\mathbf{z}_\mathrm{j}$. In this rotated space, Eq. **S3** simplifies to Eq. **S4**

$$\pi_{ij} = c_{ij} \exp\left(\left(\bar{\mathbf{z}}_j - \bar{\mathbf{y}}_i\right)^T \left(\bar{\mathbf{z}}_j - \bar{\mathbf{y}}_i\right)\right). \qquad \textbf{[S4]}$$

All further analyses were done using these rescaled and rotated latent-trait axes.

Testing the assumption of constant (co)variance. Here, we relax the assumption that Σ_{ij}^2 is constant for all i and j to test whether it causes any bias in the estimate of dimensionality of sexual isolation, using the *Desmognathus* dataset as a test case.

Relaxing this assumption, the general mating probability for any population comparison reverts to Eq. **S3**. We do not have sufficient degrees of freedom to independently estimate each (co)variance term from the data in addition to the coordinates of male and female population means **Z** and **Y**. Instead, we simulated variation in the variance/covariance structure as follows and assessed the effect on estimates of dimensionality.

First, we randomly assigned (co)variance matrices $(\mathsf{T}_i^2 + \mathsf{N}_i^2)$ to females of population i and Σ_j^2 to males of population j. Note that the phenotypic (co)variance T_i^2 and strength of preference N_i^2 for females never appear separately in the model, and therefore, we treat them here as a single term. We did this in d = 8 dimensions to allow comparison with the last line in Table 1, and we chose (co)variance matrices in three different ways. In the first case, we kept the eigenvectors of the (co)variance matrices equal across all populations and randomly assigned eigenvalues for each male and female term. Eigenvalues were chosen from a uniform distribution and normalized to have a mean of 1, creating equal covariance structure across all males and females but independent variances along each axis.

In the next two cases, we varied the eigenvectors of the covariance structure among populations as well. We created a random orthonormal basis of eigenvectors by successive Householder reflections about a vector with elements chosen from a uniform distribution and normalized to unit length, starting at one dimension and stepping up to eight dimensions. This is expected to produce a uniform distribution of orthogonal basis vectors about a unit hypersphere (2). We then assigned random eigenvalues to each eigenvector as above. In the second case, we normalized the eigenvalues of both the female and male terms to a mean of 1 as above. In the third case, we normalized the male eigenvalues to a mean of 1 and the female eigenvalues to a mean of 10. The justification for this is that the strength of female preference is expected to be low (hence, large values in N_i^2) relative to phenotypic variance in both males and females (3).

Next, we fit points to male and female population means **Z** and **Y** as described in *Methods*, except using Eq. **S3** as the mating probability for each population comparison (Eq. **1**). We calculated the effective number of dimensions n_D (4) for the maximum-likelihood distribution of population means. We did this for 100 replicates each for the three cases described above.

In all three cases, allowing the variance and covariance structure to vary across populations creates variance in the estimate of dimensionality, but it does not bias it in either direction (Fig. S2). With the assumption of equal (co)variance, n_D = 2.03 (Table 1). This point lies in the 45th, 46th, and 53rd percentiles of replicates across the three cases, respectively. In addition, the variance created in the estimate of dimensionality did not substantially exceed the range in estimates across the nine datasets. For none of the replicates did n_D lie below 1.5 or exceed 3.0.

1. Arnold SJ, Verrell PA, Tilley SG (1996) The evolution of asymmetry in sexual isolation: A model and a test case. *Evolution* 50:1024–1033.
2. Diaconis P, Shahshahani M (1987) The subgroup algorithm for generating uniform random variables. *Prob Eng Info Sci* 1:15–32.
3. Uyeda JC, Arnold SJ, Hohenlohe PA, Mead LS (2009) Drift promotes speciation by sexual selection. *Evolution* 63:583–594.
4. Kirkpatrick M (2009) Patterns of quantitative genetic variation in multiple dimensions. *Genetica* 136:271–284.

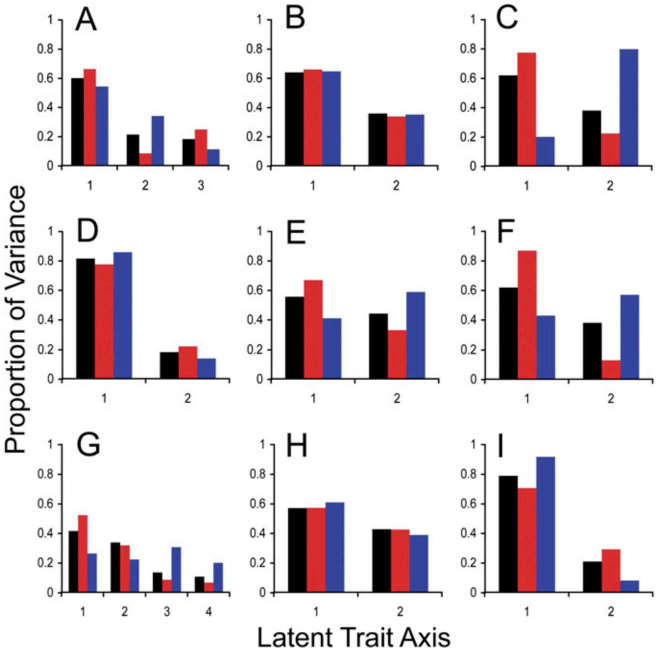

Fig. S1. Proportion of total variance (black), variance among females (red), and variance among males (blue) along each latent-trait axis for the best-fit number of dimensions in each of nine datasets. Latent-trait axes are rank-ordered by overall variance among all population means. *A*–*I* are as in Fig. 2.

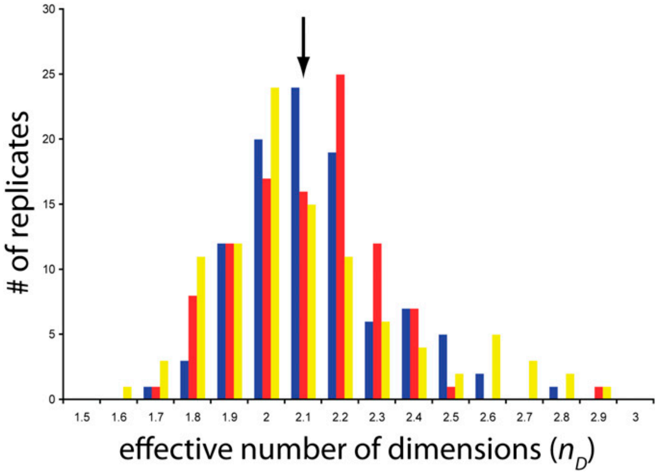

Fig. S2. Dimensionality in the *Desmognathus* dataset when (co)variance matrices vary among populations. Shown are the number of replicates (of 100) with the indicated effective number of dimensions n_D for each of three cases described above. Blue, equal covariance structure (eigenvectors) but random variance terms (eigenvalues); red, random covariance structure with mean eigenvalue = 1 for both sexes; yellow, random covariance structure with mean eigenvalue = 10 for females and mean eigenvalue = 1 for males. The arrow shows n_D under the assumption of equal variance and covariance structure from Table 1.

Table S1. Mating-trial results used to estimate dimensionality (*D. ochrophaeus* complex; nine populations) (1)

	MR	MM	RB	SI	HP	UN	JK	WA	WR
MR	106/180	17/30	1/30	0/30	3/30	6/30	0/0	5/30	0/0
MM	5/30	124/210	1/30	2/30	2/30	10/30	0/30	7/30	0/0
RB	1/30	12/30	150/240	15/30	45/90	13/30	0/0	13/30	0/0
SI	1/30	1/30	4/30	163/270	5/60	2/30	20/30	14/30	10/30
HP	0/30	7/30	17/90	9/60	162/300	9/30	3/30	8/30	0/0
UN	2/30	17/30	8/30	3/30	6/30	147/240	0/30	10/30	7/30
JK	0/0	0/30	0/0	18/30	1/30	6/30	68/120	0/0	0/0
WA	1/30	8/30	7/30	9/30	4/30	7/30	0/0	113/210	0/30
WR	0/0	0/0	0/0	12/30	0/0	6/30	0/0	4/30	52/90

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials. MR, Mt. Rogers; MM, Mt. Mitchell; RB, Rough Butt Bald; SI, Standing Indian Mountain; HP, Highlands Plateau; UN, Unaka Mountain; JK, John's Knob; WA, Wayah Bald; WR, Waterrock Knob.

1. Arnold SJ, Verrell PA, Tilley SG (1996) The evolution of asymmetry in sexual isolation: A model and a test case. *Evolution* 50:1024–1033.

Table S2. Mating-trial results used to estimate dimensionality [*Drosophila paulistorum* (Amazonian race); four populations] (1)

	PanamaA	Trinidad	Icana	Belem
PanamaA	88/127	27/34	47/138	15/59
Trinidad	19/27	96/144	2/12	22/51
Icana	24/68	15/36	92/201	38/52
Belem	10/30	41/74	19/48	124/183

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials.

1. Carmody G, et al. (1962) Mating preferences and sexual isolation within and between the incipient species of *Drosophila paulistorum*. *Am Midl Nat* 68:67–82.

Table S3. Mating-trial results used to estimate dimensionality [*D. paulistorum* (Andean–South Brazilian race); eight populations] (1)

	LlanosB	SD	Tarapoto	TM	Coroico	MG	Angra	Urubamba
LlanosB	146/233	0/0	0/0	20/42	21/49	0/0	0/0	0/0
SD	21/48	72/150	29/48	0/0	0/0	26/46	1/47	7/44
Tarapoto	38/57	0/0	95/177	55/89	36/47	0/0	17/49	0/0
TM	0/0	13/47	9/18	146/268	0/0	13/42	0/0	40/91
Coroico	0/0	23/42	0/0	29/46	114/192	28/49	0/0	0/0
MG	22/42	0/0	18/50	0/0	0/0	113/147	30/48	26/47
Angra	17/44	17/40	0/0	16/48	17/45	0/0	101/191	0/0
Urubamba	22/38	0/0	21/47	17/36	16/48	0/0	12/47	115/193

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials. SD, Santo Domingo; TM, Tingo Maria; MG, Minas Gerais.

1. Carmody G, et al. (1962) Mating preferences and sexual isolation within and between the incipient species of *Drosophila paulistorum*. *Am Midl Nat* 68:67–82.

Table S4. Mating-trial results used to estimate dimensionality [*D. paulistorum* (multiple races); seven populations] (1)

	Honduras	CR	PanamaA	LlanosA	LlanosB	SM	GuianaB
Honduras	165/221	12/28	0/0	14/45	0/0	0/0	0/20
CR	17/36	48/68	0/20	1/19	0/0	16/44	0/0
PanamaA	0/0	0/20	150/198	0/44	2/47	19/49	4/56
LlanosA	7/66	2/20	9/76	148/202	1/38	30/97	4/39
LlanosB	0/0	0/0	3/70	0/39	114/134	0/0	1/44
SM	29/101	0/0	0/0	0/0	0/0	140/203	4/32
GuianaB	0/17	0/0	3/46	2/41	4/49	0/0	155/197

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials. CR, Costa Rica; SM, Santa Marta.

1. Carmody G, et al. (1962) Mating preferences and sexual isolation within and between the incipient species of *Drosophila paulistorum*. *Am Midl Nat* 68:67–82.

Table S5. Mating-trial results used to estimate dimensionality [*D. sturtevanti*; five populations] (1)

	Tamaz	Quirigua	Belem	Rio	Bertioga
Tamaz	317/524	26/107	34/82	64/130	26/108
Quirigua	71/192	204/408	24/99	40/113	29/109
Belem	35/109	25/97	247/441	42/106	28/88
Rio	52/112	42/96	77/157	253/444	51/106
Bertioga	37/104	33/103	39/100	53/106	215/406

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials.

1. Dobzhansky T (1944) Experiments on sexual isolation in *Drosophila*: III. Geographic strains of *Drosophila sturtevanti*. *Proc Natl Acad Sci USA* 30:335–339.

Table S6. Mating-trial results used to estimate dimensionality (*D. willistoni*; five populations) (1)

	PG	Bertioga	Rio	Belem	Quirigua
PG	137/198	33/57	49/71	0/0	57/92
Bertioga	48/73	134/207	0/0	44/78	23/65
Rio	36/64	0/0	107/190	44/71	26/63
Belem	0/0	41/78	35/60	122/210	29/60
Quirigua	37/66	14/69	24/60	19/74	92/280

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials. PG, Praia Grande.

1. Dobzhansky T, Mayr E (1944) Experiments on sexual isolation in *Drosophila*: I. Geographic strains of *Drosophila willistoni*. *Proc Natl Acad Sci USA* 30:238–244.

Table S7. Mating-trial results used to estimate dimensionality (*D. prosaltans*; seven populations) (1)

	Chilpan	Zopilote	Huichi	Guat	Belem	Bertioga	Iporanga
Chilpan	381/408	51/90	41/47	76/106	76/102	73/111	60/89
Zopilote	39/65	298/400	79/114	44/75	69/113	73/108	56/80
Huichi	23/70	25/72	281/481	39/80	30/100	49/96	43/95
Guat	18/63	11/59	35/92	206/443	17/106	33/80	60/155
Belem	23/86	11/42	11/82	11/102	309/644	20/94	33/80
Bertioga	3/48	15/76	13/74	6/51	42/116	204/547	52/108
Iporanga	8/57	17/53	36/75	6/52	47/94	26/67	282/615

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials.

1. Dobzhansky T, Streisinger G (1944) Experiments on sexual isolation in *Drosophila*: II. Geographic strains of *Drosophila prosaltans*. *Proc Natl Acad Sci USA* 30:340–345.

Table S8. Mating-trial results used to estimate dimensionality (*D. auraria* complex; four species) (1)

	A	B	T	Q
A	581/734	2/187	9/229	11/288
B	38/269	412/600	18/228	52/266
T	27/199	2/189	462/725	149/319
Q	175/260	38/229	227/273	782/868

Shown are the mating-trial results for each male–female comparison (female population is given at left and male population is along the top). Numbers given are *k*/*n*, where *k* is the number of successful matings in *n* trials.

1. Kurokawa H, Oguma Y, Tachibana N (1982) Sexual isolation among four species of *D. auraria* complex. *Drosoph Inf Serv* 58:98–99.

Table S9. Mating-trial results used to estimate dimensionality (*Pseudotropheus zebra* complex; five populations) (1)

	CH	MR	NB	EM	RW
CH	19	1	11	1	5
MR	1	9	2	2	2
NB	8	0	9	2	1
EM	1	0	1	26	4
RW	0	0	1	1	8

Shown are the mating-trial results for each male—female comparison (female population is given at left and male population is along the top). The number is the count of successful matings over a defined experimental period. CH, Chisumulu; MR, Mara Rocks; NB, Nkhata Bay; EM, Mphanga Rocks *emmiltos*; RW, Ruarwe.

1. Knight ME, Turner GF (2004) Laboratory mating trials indicate incipient speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra* from Lake Malawi. *Proc Biol Sci* 271:675–680.