



## Research

**Cite this article:** Vestergaard JS, Twomey E, Larsen R, Summers K, Nielsen R. 2015 Number of genes controlling a quantitative trait in a hybrid zone of the aposematic frog *Ranitomeya imitator*. *Proc. R. Soc. B* **282**: 20141950. <http://dx.doi.org/10.1098/rspb.2014.1950>

Received: 7 August 2014

Accepted: 26 March 2015

**Subject Areas:**

evolution, genomics

**Keywords:**

*Ranitomeya imitator*, hybridization, image analysis, quantitative phenotyping

**Author for correspondence:**

Jacob S. Vestergaard  
e-mail: [jsve@dtu.dk](mailto:jsve@dtu.dk)

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2014.1950> or via <http://rspb.royalsocietypublishing.org>.

# Number of genes controlling a quantitative trait in a hybrid zone of the aposematic frog *Ranitomeya imitator*

Jacob S. Vestergaard<sup>1</sup>, Evan Twomey<sup>2</sup>, Rasmus Larsen<sup>1</sup>, Kyle Summers<sup>2</sup> and Rasmus Nielsen<sup>3</sup>

<sup>1</sup>Department of Applied Mathematics and Computer Science, Technical University of Denmark, Building 324/130, Matematiktorvet, 2800 Kgs Lyngby, Denmark

<sup>2</sup>Department of Biology, East Carolina University, Howell Science Complex N314, Greenville, NC 27858, USA

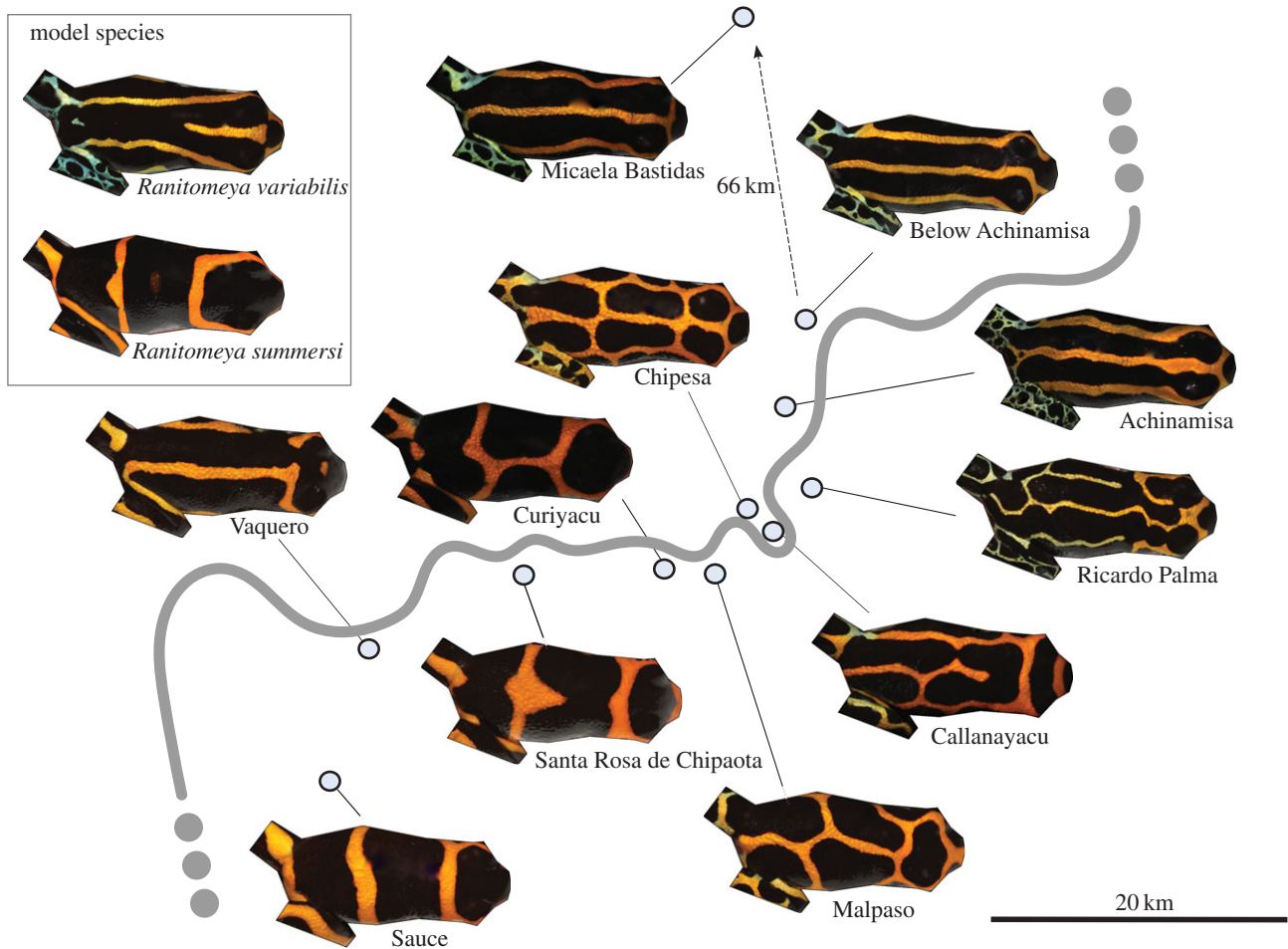
<sup>3</sup>Department of Integrative Biology, University of California Berkeley, 4098 VLSB, Berkeley, CA 94720, USA

JSV, 0000-0001-6299-3843

The number of genes controlling mimetic traits has been a topic of much research and discussion. In this paper, we examine a mimetic, dendrobatid frog *Ranitomeya imitator*, which harbours extensive phenotypic variation with multiple mimetic morphs, not unlike the celebrated *Heliconius* system. However, the genetic basis for this polymorphism is unknown, and not easy to determine using standard experimental approaches, for this hard-to-breed species. To circumvent this problem, we first develop a new protocol for automatic quantification of complex colour pattern phenotypes from images. Using this method, which has the potential to be applied in many other systems, we define a phenotype associated with differences in colour pattern between different mimetic morphs. We then proceed to develop a maximum-likelihood method for estimating the number of genes affecting a quantitative trait segregating in a hybrid zone. This method takes advantage of estimates of admixture proportions obtained using genetic data, such as microsatellite markers, and is applicable to any other system where a phenotype has been quantified in an admixture/introgression zone. We evaluate the method using extensive simulations and apply it to the *R. imitator* system. We show that probably one or two, or at most three genes, control the mimetic phenotype segregating in a *R. imitator* hybrid zone identified using image analyses.

## 1. Introduction

The analysis of phenotypic and genetic variation in geographical areas where two or more phenotypically distinguishable groups of organisms meet and exchange genes has been of substantial interest to evolutionary biologists [1–3]. The evolutionary dynamics in these zones, referred to as hybrid zones, introgression zones or admixture zones depending on context, provide us a basis to study processes relating to speciation and for understanding the genetic and ecological underpinnings of adaptive traits, including mimetic and aposematic traits. Substantial work has been done on such systems, including the now classical work on the *Bombina bombina* versus *Bombina variegata* hybrid zone [4,5] and the hybrid zones between various species of *Heliconius* butterflies [6–8]. Of primary interest in these studies is to understand the genetic basis of the phenotypic traits, how selection is affecting these traits, and to understand the relative role of population history, gene flow and natural selection in determining the evolutionary dynamics of the hybrid zone. Furthermore, there has recently been renewed interest in mapping the genetic variants associated with reproductive isolation or adaptive traits in the hybrid zone using so-called admixture mapping or mapping by admixture linkage disequilibrium [9–14]. Of special interest to us are admixture zones, exemplified by the previously mentioned examples in *Bombina* and



**Figure 1.** Sketch of sampling locations along the Huallaga River (grey). The two model species (*R. variabilis* and *R. summersi*) are shown in the upper left corner and examples of *R. imitator* are connected to their sampling localities. (Online version in colour.)

*Heliconius*, in which complex morphological traits such as colour patterns are segregating and are probably of adaptive significance [15,16].

Mimetic traits in admixture zones, or otherwise, have often been hypothesized to be associated with so-called supergenes [17]. Supergenes are tightly linked clusters of genes that control a suite of traits that will allow Mendelian, or close to, Mendelian behaviour of the mimicry trait. The existence of such supergenes could help us explain the strong phenotypic correlation between many different phenotypes required to produce pure mimetic forms. If many traits are needed to produce an adaptive mimetic phenotype, then one would expect selection to favour genetic variants that increase the correlation between these traits. Much discussion has ensued on the existence of supergenes, particularly in relation to mimetic phenotypes in butterflies [18]. A recent paper by Kunte *et al.* [19] shows that a proposed supergene underlying mimetic phenotypes in *Papilio* butterflies in fact is a single Mendelian gene that serves as a genetic switch for the mimetic type. For both *Papilio* and *Heliconius*, it appears that the mimetic phenotypes are often controlled by one or a few genes or supergenes that behave in a largely Mendelian fashion. However, the degree to which mimetic phenotypes have a similar genetic basis in other systems is uncertain.

The dendrobatid frog *Ranitomeya imitator* [20,21] provides us with a new vertebrate model system that shares many features with the well-known *Heliconius* system. In Peru, there are four distinct colour pattern morphs of *R. imitator* that occupy different geographical regions [22]. In each of these

regions, the colour pattern of *R. imitator* clearly resembles that of a co-occurring species of dendrobatid frog [20,21]. Phylogenetic analyses indicate that these co-occurring species generally diverged prior to the divergence between the divergent populations of *R. imitator* [23,24]. Evidence for rapid divergence under selection [21,22], and the similarity of each *R. imitator* colour pattern morph to the more anciently diverged co-occurring species, indicates that *R. imitator* has undergone a mimetic radiation, in which different populations have evolved to resemble distinct colour patterns displayed by the local 'model' species [21,22]. Recent analyses of colour pattern variation, genetic structure and gene flow have identified multiple zones of admixture where distinct colour pattern morphs of *R. imitator* come into contact and interbreed [20]. These regions vary in terms of the width of the zone of admixture and the degree of genetic divergence (in neutral markers) found across the zone, making this system useful for comparative analyses of divergence. In one region, the zone of admixture is fairly broad (7 km), and populations in the zone show high variability that appears to include elements of both distinct colour pattern morphs. Note in figure 1 that frogs at one end of the admixture zone, where they are mimetic with *Ranitomeya summersi*, tend to be banded with black and orange legs, whereas frogs, on the other end, where they are mimetic with *Ranitomeya variabilis*, tend to be striped with a reticulated green and black pattern on the legs. The genetic basis of this polymorphism is of primary interest, but given the large genome sizes of dendrobatid frogs (e.g. up to 9 Gb, Camper *et al.* [25]), lack of genetic resources and difficulty of

captive breeding, direct mapping of the genes involved is a non-trivial task. An objective of this paper is instead to obtain more information about the genetic basis of this polymorphism, solely using image analyses and limited microsatellite typing. In particular, we will be interested in examining whether the polymorphism is controlled by a single Mendelian gene, perhaps a supergene, or by multiple genes.

In order to address this problem, we will first develop automated methods for describing complex colour pattern phenotypes based on images that can be applied in this system and other systems. The advantage of such methods is that they are not subject to the same biases that may occur when a researcher chooses which traits to measure after having observed the images. In addition, such methods may have the potential for identifying important biological features that were otherwise not readily identifiable.

We will then proceed to develop a method for estimating the number of genes affecting a phenotype in an admixture/hybrid zone. For natural populations, in which controlled crosses are difficult or expensive to carry out, and for which parent–offspring pairs cannot easily be sampled, there are no appropriate methods for determining how many genes affect a trait. In other settings, there has been substantial previous work on this problem. The well-known Castle–Wright estimator [26,27] is based on the amount of segregating variation observed in the offspring of controlled crosses of inbred lines. The objective is to estimate the effective number of loci controlling a quantitative trait, i.e. the number of loci required to explain the variance in the trait if all loci have the same effect. There have been numerous extensions of the method, including the incorporation of linkage and variation in effects size [28,29]. Lande [30] showed that the assumption of complete homozygosity in the parental lines is not necessary and provided an estimator applicable to natural populations, rather than to controlled crosses. Building on the idea, dating back to Pearson [31], that the relationship between the variance in the offspring phenotypic values and midparent value depend on the number of genes controlling the trait, Slatkin [32] provided another estimator applicable to outbred populations.

We are interested in estimating the number of genes affecting a trait in a hybrid/admixture zone. This is a problem that has been considered by Szymura & Barton [4] who, based on theory developed in Barton [33] and Barton & Bengtsson [34], estimated the number of genes contributing to selection against gene flow in the *B. bombina* versus *B. variegata* hybrid zone using comparisons of the amount of linkage disequilibrium at the centre of a hybrid zone to the width of the cline. The method we will develop is in the spirit of the of the Castle–Wright–Lande estimators, but is based on using a genetically inferred admixture proportion in each individual. This method does not require data on controlled crosses. It also does not rely on any assumptions regarding selection models and processes shaping linkage disequilibrium. It is less ambitious in that it does not attempt to determine the number of genes affecting fitness, but the number of genes affecting an observable phenotype. There is substantially less information regarding the number of loci when controlled crosses have not been performed. However, as we will show, there is still sufficient information to distinguish between hypotheses regarding one, two or several genes affecting the trait.

We will apply the methods developed in this paper to images and genetic data from the aforementioned dendrobatid frog *R. imitator*. Using these methods, we can estimate the

number of genes affecting the mimetic phenotype without the use of experimental crosses or mapping approaches.

## 2. Image analysis/quantitative phenotyping

A common way to quantify variation in image analysis is to extract a number of so-called descriptors, combine these into a vector of measurements for each individual and use statistical decomposition methods to condense the collected information. Prior to analysis, all individuals have been warped to a mean shape determined by Procrustes analysis [35]. Manual annotation of 22 anatomical landmarks was used to establish point correspondences.

Descriptors are typically designed to capture elementary characteristics of an image, such as colour or shape. Individually, descriptors are usually too specific, but a well-chosen suite of descriptors can provide us with a rich basis for further analysis.

In our study, we use three different phenotypic descriptors: colour/non-colour ratio, gradient orientation histograms and shape index histograms [36], each of which is defined on the pixel-level and described in detail in the electronic supplementary material, S1. These descriptors collect local zeroth-, first- and second-order information about the image. In the current setting, these three standard descriptors can loosely be thought of as measuring features relating to the proportion of coloured area, the degree to which changes in colour occur along the anteroposterior axis or along the left–right axis (banded patterns versus striped patterns), and the degree to which the pattern consists of stripes/bands as opposed to reticulation, respectively. The quantified information is visualized in figure 2 and in the electronic supplementary material, S1.

All descriptors are extracted on a per-pixel basis and pooled together at four distinct interest points, namely each of the frog's legs, lower back (dorsum) and on the back of the head. An interest point is defined in terms of its coordinates  $\mathbf{x}_k = [x_k, y_k]$  and a radius  $r_k > 0$ . Thus, an average of each of the descriptors is accumulated for the four regions shown in figure 2c (see also the electronic supplementary material, S1).

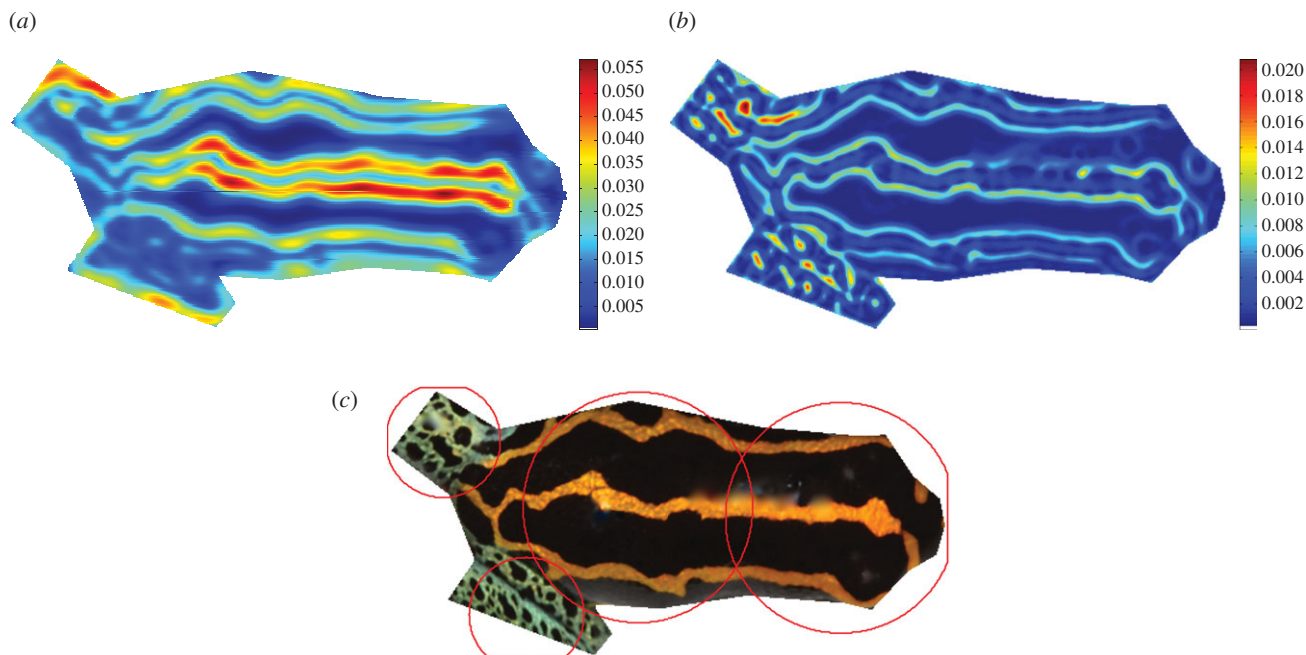
### (a) Revealing a mimicry-related phenotype with sparse discriminant analysis

The collected phenotypic descriptors are here condensed into a single mimicry-related phenotype. This amounts to determining the low-dimensional manifold, in the high-dimensional feature space, describing the phenotype. We have chosen to use sparse discriminant analysis (SDA) by Clemmensen *et al.* [37] for this task. More detail on this procedure can be found in the electronic supplementary material, S1, where alternative methods are also considered.

The composite phenotype is constructed as the linear combination  $\beta$  of the  $p$  descriptors  $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_p]$  that best describes the mimicry across the hybridization transect, i.e. the direction in the  $p$ -dimensional space that maximizes the ratio of the between-group variance to the within-group variance under elastic net regularization [38].

We define the mimicry-related phenotype for the  $i$ th individual as the projection onto the one-dimensional subspace spanned by  $\beta$ :

$$z_i = \sum_{j=1}^p \mathbf{D}_{ij} \beta_j, \quad (2.1)$$



**Figure 2.** Examples of phenotypic descriptors and illustration of the spatial pooling scheme with four interest points. (a) Gradient at  $b = \pm (\pi/2)$ . (b) Shape index at  $b = -0.8$ . (c) Pooling of phenotypic descriptors. (Online version in colour.)

where  $D_{ij}$  is the  $j$ th descriptor value for the  $i$ th individual. For all  $n$  individuals, this is equivalent to  $\mathbf{z} = \mathbf{D}\boldsymbol{\beta}$ .

### 3. A likelihood method for identifying the effective number of genes

We are interested in estimating the effective number of genes,  $K$ , affecting a trait, i.e. the number of genes required to explain the observed phenotypic variation assuming all genes have the same effect. We assume we have a sample of  $n$  individuals from an admixture zone, each with some associated genetic data (e.g. microsatellite data). We will take advantage of the fact that even limited genetic data can be used to infer an admixture fraction for each individual,  $\mathbf{f} = \{f_i\}_1^n$ , under the assumption that pure forms exist at each end of the transect in the admixture zone.  $f$  and  $1 - f$  then represent the proportion of an individual's genome that is identical to individuals in the right and left end of the transect, respectively. The method we use for estimating the admixture fractions is described in the electronic supplementary material, S1, and is based on the kernel discriminant analysis (KDA) of Mika & Ratsch [39] with the kernel suggested by Martin [40]. The estimated admixture proportions are shown in the electronic supplementary material, S6.

We will assume that the phenotypic values,  $\mathbf{z} = \{z_i\}_1^n$ , are normally distributed, given the underlying genotype, and that each locus contributing to the phenotype has the same effect and dominance factors, and that the effects are additive among loci. We will also assume that each locus is di-allelic and that the allele favouring the phenotype in the right end of the transect has frequency 1 in the right extreme of the transect and frequency 0 in the left end of the transect. We will also, without loss of generality, denote the alleles favouring the phenotype in the right and left ends of the transect by  $a$  and  $A$ , respectively. An individual with admixture proportion  $f$ , assuming independence among the parental

contributions, then has genotype  $AA$  in any locus with probability  $(1 - f)^2$ .

We consider the phenotype,  $z$ , of an individual to be a realization of the stochastic variable  $Z$  with the conditional distribution

$$Z | \mathbf{g} \sim \mathcal{N}(\mathbf{h}^T \boldsymbol{\mu}, \sigma_e^2), \quad (3.1)$$

where  $\sigma_e^2$  is the environmental variance and  $\mathbf{g} = \{G_k\}_1^K$  is a vector of the  $K$  genotypes

$$G_k = \begin{cases} 0 & \text{if } AA, & p(G_k = 0|f) = (1 - f)^2 \\ 1 & \text{if } Aa, & p(G_k = 1|f) = 2f(1 - f) \\ 2 & \text{if } aa, & p(G_k = 2|f) = f^2. \end{cases} \quad (3.2)$$

Three averages are used for the conditional Gaussians  $\boldsymbol{\mu} = [\mu_0, \mu_1, \mu_2]^T$  and

$$\mathbf{h}_k = [h_0, h_1, h_2]^T, \quad \text{where } h_q = \frac{1}{K} \sum_{k=1}^K I(G_k = q),$$

i.e. a vector containing fractions of the  $K$  genes having the genotypes  $AA$ ,  $Aa$  and  $aa$  respectively.

So, for example, if  $K = 3$ , an individual with genotypes  $AA$ ,  $Aa$  and  $aa$  in the three loci, respectively, will have mean phenotype  $2\mu_0 + \mu_2$ .

Thus, in a noise-free scenario, a single gene would be able to explain a trait as a piecewise constant function (of the admixture proportion) with three steps.  $K$  genes would be able to explain a trait attaining  $\binom{K+2}{2}$  different values.

Here, a noise-free scenario would mean no environmental variance in the phenotype *and* no noise caused by the quantification of the phenotype.

To calculate the likelihood, all possible combinations of genotypes must be considered. The set of all possible combinations will be denoted  $\mathbf{G}(K) = \{0, 1, 2\}^K$ , i.e. the  $K$ th Cartesian power of possible genotypes, where a single tuple from this set will be denoted  $\mathbf{g}_j = [G_{j1}, G_{j2}, \dots, G_{jK}]$ . This set consists of all possible combinations, with

replacement, where the order is significant. A total of  $3^K$  such combinations exists.

The probability of a certain combination of genotypes  $\mathbf{g}_j$  given the mixture proportion  $f$  is

$$p(\mathbf{g}_j|f) = \prod_{k=1}^K p(G_{jk}|f).$$

The likelihood of observing the phenotypic trait over the entire population, allowing  $K$  genes to contribute to the expression of the trait, is modelled as

$$p_K(\mathbf{z}|f) = \prod_{i=1}^n \left[ \sum_{\mathbf{g}_j \in G(K)} p(z_i|\mathbf{g}_j) p(\mathbf{g}_j|f_i) \right]. \quad (3.3)$$

However, the estimates of  $f_i$  may be associated with statistical uncertainty. Ignoring this uncertainty could lead to biased estimates. We therefore provide an alternative formulation that incorporates uncertainty in the estimates of  $f_i$  using a bootstrap approach, i.e. we assume that marker loci used for estimation of  $f_i$  have been bootstrapped to provide a bootstrap distribution  $\{f_i^b\}_{b=1}^B$ . The likelihood of observing the phenotypic trait over the entire population, allowing  $K$  genes to contribute to the expression of the trait, is then modelled as

$$p_K(\mathbf{z}|f) = \prod_{i=1}^n \left[ \sum_{\mathbf{g}_j \in G(K)} p(z_i|\mathbf{g}_j) \frac{1}{B} \sum_{b=1}^B p(\mathbf{g}_j|f_i^b) \right]. \quad (3.4)$$

For a fixed value of  $K$ , we maximize this function for  $\mu_0$ ,  $\mu_1$ ,  $\mu_2$  and  $\sigma_e^2$  using the Broyden-Fletcher-Goldfarb-Shanno algorithm [41]. We then repeat this procedure for multiple values of  $K$  and choose the value of  $K$  that maximizes this profile likelihood function as our maximum-likelihood estimate of  $K$ . To increase the probability of converging to a global maximum, we use a scheme with multiple starting points, see the electronic supplementary material, S2 for details.

We evaluate the performance of the method using simulations allowing for varying heritability and uncertainty in the estimates of  $f$ . The heritability is defined as the fraction of the total phenotypic variance  $V_P$  that can be attributed to genetic variance:

$$H^2 = \frac{V_G}{V_P} = \frac{V_G}{V_G + \sigma_e^2}. \quad (3.5)$$

The average phenotypic value is  $\bar{z} = \sum_{j=1}^{N_G} z_j p_j$ , where  $z_j$  is the phenotypic value determined by the genotype and  $p_j$  is the proportion of individuals with the  $j$ th genotype.

The genetic variance is determined as

$$V_G = \sum_{j=1}^{N_G} (z_j - \bar{z})^2 p_j. \quad (3.6)$$

To simulate data for a phenotype determined by  $K$  genes,  $n$  mixture proportions  $\mathbf{f} = \{f_i\}_1^n$  are drawn, e.g. from a uniform distribution on the interval  $[0, 1]$ . The genotype for each of the  $K$  loci is then drawn from a multinomial distribution with probabilities as in equation (3.2). Phenotypes are then assigned by simulating from a normal distribution as in equation (3.1). In simulations with noise in the estimate of  $f$ , we simulate  $B$  samples from a normal distribution with standard deviation  $\sigma_f$  around  $f_i$ , such that mixture proportions used for inference  $\hat{f}_i^b \sim \mathcal{N}(f_i, \sigma_f^2)$ .

## 4. Image and microsatellite data

We used published microsatellite data from three sources: Twomey *et al.* [20] (92 samples), Twomey *et al.* [42] (36 samples) and Twomey [43]. The final dataset consisted of 285 *R. imitator* individuals from 16 localities in Peru: the 11 localities shown in figure 1 and five localities between Santa Rosa de Chipaota and Achinamisa (i.e. within the banded-stripped transition area). For the unpublished microsatellite data, amplification methods follow Twomey *et al.* [20].

We used JPEG compressed images of six *R. summersi*, seven *R. variabilis* and 304 *R. imitator* individuals from the 11 localities shown in figure 1. The images are  $3888 \times 2592$  pixels of size captured with a Canon EOS Rebel XS SLR. Both microsatellite data and image data were available for 179 of the *R. imitator* individuals. See the electronic supplementary material, S5 for a full overview of the image material.

## 5. Results

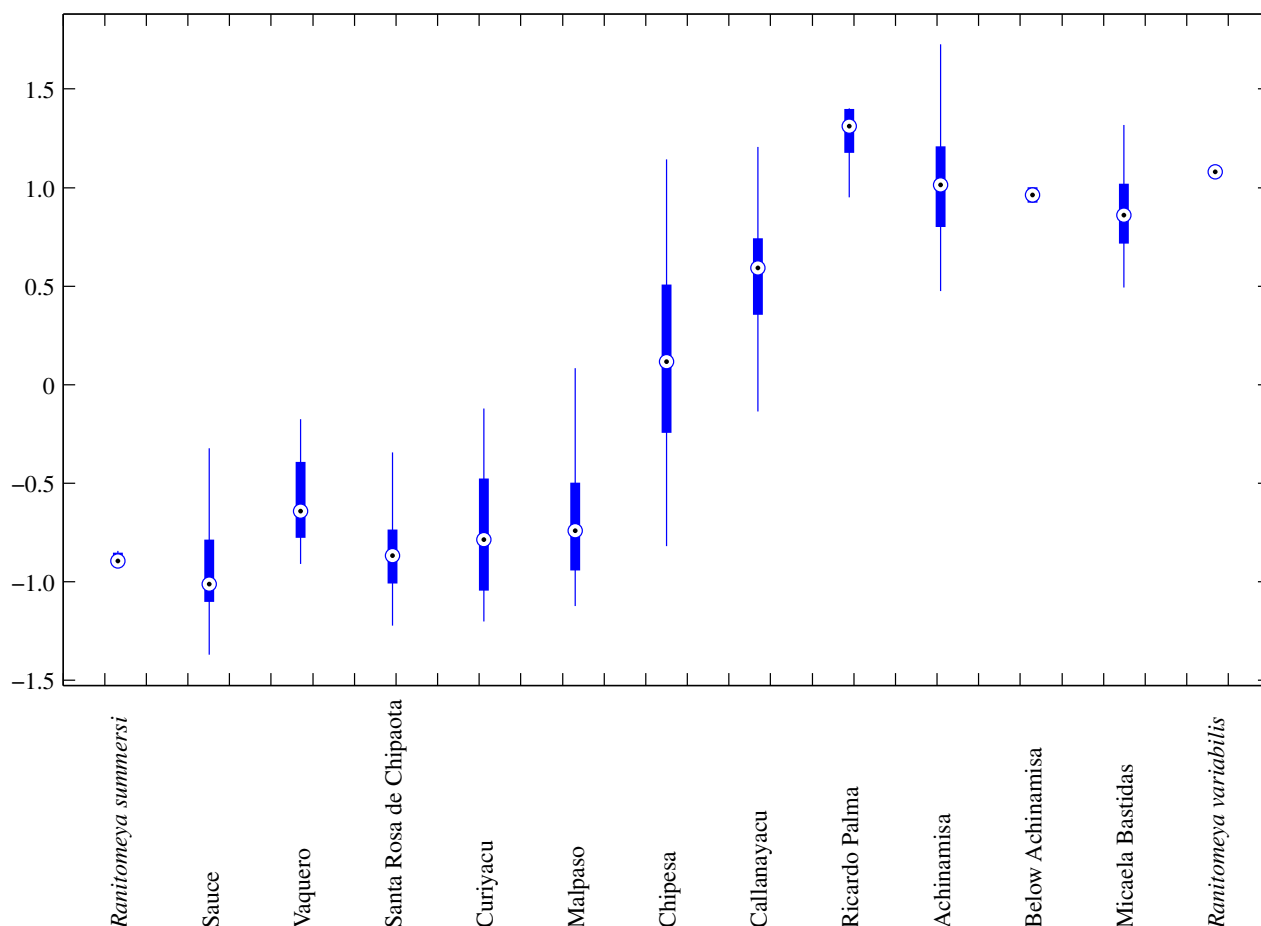
### (a) Phenotypic descriptors

The phenotypic descriptors described in §2 were automatically extracted from all 317 images. Different aspects of the patterns in the population are captured by this collection of descriptors, the most dominant being the stripe directionality; for more detail on the phenotypic variance captured by these descriptors, see the electronic supplementary material, S1. For every individual, the suite of descriptors extracted for the four interest points (left leg, right leg, lower back, upper back) are colour/non-colour ratios for each point of interest, gradient orientation histograms binned in two bins at scales  $\sigma = [2, 7]$  with tonal range  $\beta = 1$  and shape index histograms in five bins at scales  $\sigma = [4, 8]$  with tonal range  $\beta = 1$ . This adds up to a total of  $p = 4 \cdot (1 + 2 \cdot 2 + 5 \cdot 2) = 60$  extracted phenotypic descriptors collected in  $\mathbf{D} \in \mathbb{R}^{n \times p}$ . The columns of this matrix are centred and normalized to unit variance prior to further analysis.

### (b) Mimicry-related phenotype

We use SDA to identify the linear combination of phenotypic descriptors that best captures the variation in mimetic phenotypes. Under the assumption that the mimetic phenotype has been under selection to resemble the phenotypes of either *R. variabilis* in one end of the transect, or *R. summersi* in the other, we use images of seven *R. variabilis* individuals to represent one group and six imaged *R. summersi* the other group, as the training set. The *R. imitator* individuals only enter the analysis to influence the choice of regularization parameter; see details in the electronic supplementary material, S1.

In the supplementary material, we also provide results when instead using the most extreme *R. imitator* populations to represent the end populations. There are disadvantages and advantages of both of these approaches. Using the model species amounts to defining the mimicry-related phenotype in terms of similarity to those species. This is desirable when the mimicry-related phenotype is of prime interest. However, it has the disadvantage that the two model species may differ in traits not mimicked by *R. imitator*. Using the most extreme *R. imitator* populations has the disadvantage that some of the individuals may not be pure mimetic forms. We obtained similar results using either of these



**Figure 3.** Composite mimicry-related phenotype. Locations are ordered left to right from south to north along the Huallaga River. The dot on each box indicates the median, the edges of the box the 25th and 75th percentile and the whiskers extend 1.5 times the inter-quartile range beyond these percentiles. (Online version in colour.)

approaches, or if we pool both the most extreme *R. imitator* populations and the model species individuals (see the electronic supplementary material). In the following, we will refer to the extreme groups as the mimicry-defining groups, independently of how they were defined. Combinations of these different ways of specifying the mimicry-defining groups and the three manifold learning methods used to quantify the phenotype are included as the electronic supplementary material, S1.

The mimicry-related phenotypic value for each individual is obtained by projecting onto the direction  $\beta$  according to equation (2.1) and will be denoted  $z_i$  for the  $i$ th individual. The values are scaled linearly such that the average value for each of the model species is  $-1$  and  $1$ , respectively.

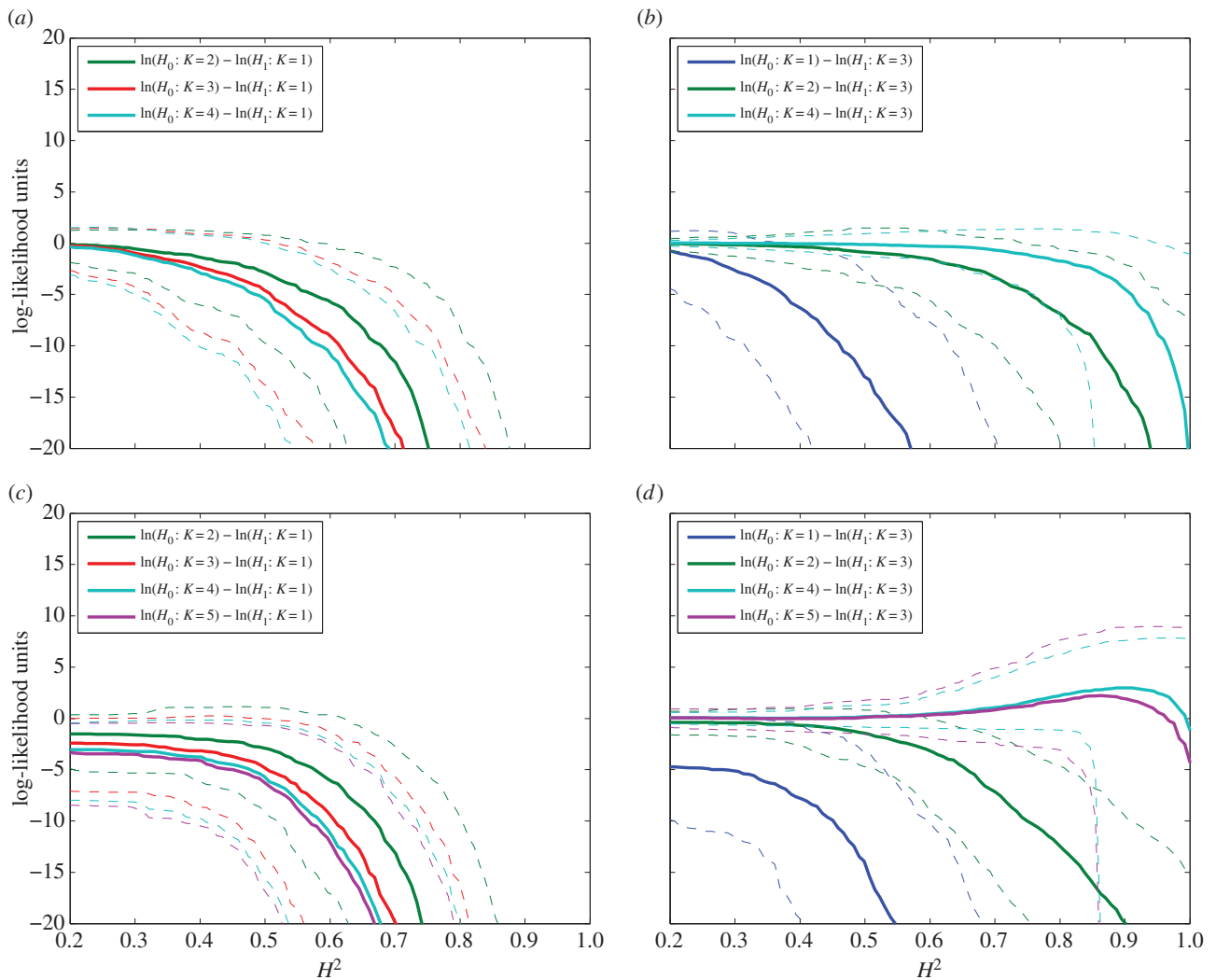
Grouping the individuals by location and ordering them along the transect from south to north (figure 1), a boxplot summarizing the mimicry-related phenotypic values as a function of location can be seen in figure 3. Note that the first half of the locations tend to have a value similar to *R. summersi*, whereas the other half have values closer to *R. variabilis*. Chipesa and Callanayacu have phenotypic values that are more intermediate and with relative high variances. Note that the ordering of the locations on the  $x$ -axis does not correspond to the actual geographical distances.

### (c) Simulation studies

We evaluated the accuracy of the method for determining the number of genes underlying a quantitative phenotype

presented in §3 on simulated data for different values of the heritability (see Methods section in the electronic supplementary material). The accuracy is evaluated under a variety of scenarios constructed by: (i) varying the true number of genes  $K$ , (ii) sampling the admixture proportions from a uniform or a bimodal distribution, and (iii) adding white noise to the admixture proportions. The heritability was varied by simulating data for 1000 different values of  $\sigma_e^2$ . The graphs in figure 4 show kernel density estimates of the median-likelihood ratios, 5th and 95th percentiles of the hypothesis of  $K = 1, 2, 3$  or 4 genes versus the alternative hypothesis of the true number of genes under which the data are simulated. Thus, a negative likelihood ratio means that the correct number of genes is selected. Simulated data and likelihood ratios for a few of these thousands of simulations can be found in the electronic supplementary material, S3.

Generally, the chance of accurate estimation is reduced when: (i) the true number of genes is high, (ii) the heritability decreases, or (iii) the sample size decreases. A measure of confidence in the inference can be obtained by bootstrapping individuals, using the likelihood ratios comparing different hypotheses as statistics. However, if the estimates of  $f$  are very noisy, there tends to be a systematic bias towards a higher number of genes for intermediate heritabilities. The effect of this can be seen in figure 4d. Using a bootstrap test, we find  $-6.63$  and  $0.30$  as the 5th and 95th percentiles of the likelihood ratio associated with the null hypothesis of  $H_0: K = 3$  versus  $H_1: K = 4$ , for the scenario with a heritability of approx. 0.85, despite the fact that  $K = 3$  is the true



**Figure 4.** Likelihood ratios as a function of the heritability  $H^2$  for simulated data. The graphs show median-likelihood ratios (solid lines), 5th and 95th percentiles (dashed lines) for  $K = \{1, \dots, 5\}$  versus the true  $K$ . The captions show the true parameters used to simulate the data for each scenario. One thousand estimations were performed for each of the scenarios. All simulations show that there is never significant support for choosing the wrong value of  $K$ , except when the estimation noise on  $f$  is high ( $d$ ) where the model is biased towards a higher number of genes for intermediate values of  $H^2$ . ( $a$ )  $K = 1, f \sim \mathcal{U}(0, 1)$ , ( $b$ )  $K = 3, f \sim \mathcal{U}(0, 1)$ , ( $c$ )  $K = 1, f \sim \mathcal{U}(0, 1) + \mathcal{N}(0, 0.01^2)$ , ( $d$ )  $K = 3, f \sim \mathcal{U}(0, 1) + \mathcal{N}(0, 0.1^2)$ . (Online version in colour.)

number of underlying genes. Thus, sensitivity to estimation variance in the admixture proportion must be kept in mind when applying this likelihood model.

#### (d) Number of genes underlying the mimicry phenotype

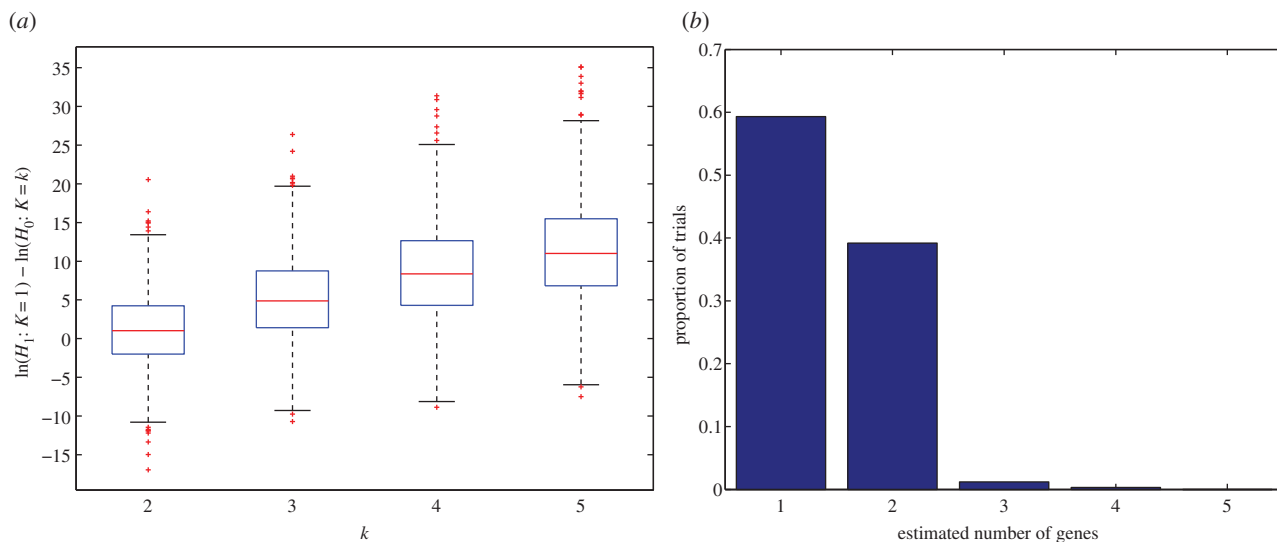
The above-described likelihood model was used to estimate the number of genes underlying the quantified phenotype in *R. imitator*. We use 1000 bootstrap replicates to obtain a distribution of likelihood ratios between different alternative models. The bootstrap is performed by sampling individuals with replacement. First, a bootstrap distribution of the mixture proportions for each individual is obtained using the available 285 samples. We take into account uncertainty in the estimation of  $f$ , by, for each simulation, re-estimating  $f$  (see the electronic supplementary material, S2) by also bootstrapping microsatellite loci within each individual.

The maximum-likelihood values of  $K$ , for  $K = \{1, \dots, 5\}$  were then determined for each replicate in a separate bootstrap experiment using the 179 samples with the genetic and

phenotypic data available. Figure 5a shows a boxplot of the distribution of likelihood ratios associated with the hypothesis  $H_0: K = k$  for  $k = 1, 2, 3, 4, 5$ , against the alternative hypothesis of  $H_A: K = 1$  and figure 5b the proportion of bootstrap replicates in which each model obtained the highest likelihood value. This proportion can be interpreted as a measure of statistical confidence. In the electronic supplementary material, S4 the full distribution of likelihood ratios associated with the test of  $H_0: K = 2$  against  $H_A: K = 1$  can be seen.

The maximum log-likelihood values are numerically highest for  $K = 1$ , and in the vast majority of the runs, this model is selected as the most likely. The point estimates of the parameters for the hypothesis of  $K = 1$  are  $[\mu_0, \mu_1, \mu_2, \sigma_e] = [0.882, 0.071, -0.855, 0.274]$ . The  $p$ -value associated with different model comparisons is shown in table 1.

Overall, a model with one ( $K = 1$ ) or two genes ( $K = 2$ ) seems to fit the data best, whereas three genes ( $K = 3$ ) cannot be rejected. The mimetic phenotype, as measured here, is probably mostly influenced by one or two genes of major effect. No combinations of the three different ways of defining the end populations suggest more than three genes. Using



**Figure 5.** Maximum log-likelihood, model selection proportions and distribution of differences from bootstrapping 1000 samples. The boxplot summarizes the distribution of maximum log-likelihood values for a model assuming  $K = \{1, \dots, 5\}$  genes. The bar plot shows the proportion of runs where each  $K$  has the maximum log-likelihood. (a) Log-likelihood ratios  $\ln((H_1:K=1)/H_0:K=k)$ . (b) Model selection proportions. (Online version in colour.)

**Table 1.**  $p$ -values for hypotheses of the number of genes, where different mimicry-defining groups are chosen. (SDA is used with regularization parameter  $\delta = 0.01$  and  $n_+$  non-zero loadings.  $p$ -values of less than 0.05 are typeset in bold and  $p$ -values more than 0.95 are typeset in italic. See the electronic supplementary material S1 for more details.)

groups	$n_+$	$P_{(k=2)/(k=1)}$	$P_{(k=3)/(k=2)}$	$P_{(k=4)/(k=3)}$	$P_{(k=5)/(k=4)}$
models	13	0.401	<b>0.030</b>	<b>0.011</b>	<b>0.000</b>
imitator	33	<i>0.989</i>	0.136	<b>0.000</b>	<b>0.025</b>
both	28	<i>0.964</i>	0.158	<b>0.014</b>	<b>0.015</b>

alternative, nonlinear, manifold learning algorithms to quantify the mimicry-related phenotype (see the electronic supplementary material, S1), only a single combination (KDA with the model species defining the end groups) cannot reject  $H_0: K=4$  versus the alternative of  $H_1: K=3$  with a  $p$ -value of less than 0.05.

## 6. Discussion

We have here developed an automated procedure for characterizing complex phenotypes from images. We believe that this method, or related methods, could be of use in many systems where images are available for complex phenotypes. Automated extraction of phenotypic descriptors reduces the subjective biases that may occur when measurements are taken manually and allows for reproducibility of results. While other biases may be introduced through the choice of an image capturing system, lighting conditions and/or choice of descriptors, we believe these to be easier to identify and overcome. We note that such image analyses open up the possibility for a variety of statistical analyses of phenotypes, and their correlations, not pursued here. In this paper, we use the image analyses to define a quantitative measure of the mimetic phenotype in a transition zone between morphs of *R. imitator*. Using a new method for estimating the effective

number of genes affecting this phenotype, we show that the phenotype we measure is likely to be controlled by one or two, or at most three, genes of major effect, and is very unlikely to be affected by many major effect genes. However, there could be substantial phenotypic variation controlled by other genes, but not captured by our quantitative measure of mimetic phenotype.

The fact that we have identified a measure of mimetic phenotype that is controlled by a few genes suggests that future studies aimed at mapping this phenotype have a relatively high probability of succeeding. It is substantially easier to map the genes underlying a phenotype controlled by just one or a few genes, than a phenotype controlled by many genes. The phenotype defined here would be useful for such mapping studies.

We can compare our results with *Heliconius* butterflies where the genetic basis of Müllerian mimicry is better understood. In *Heliconius erato*, the transition between the 'postman' and the 'rayed' morphs in the well-studied hybrid zone near Tarapoto, Peru is controlled by three loci of major effect, whereas in the co-mimetic *Heliconius melpomene*, the same mimetic shift (postman to rayed) is controlled by five loci [44]. In another example, the polymorphism in *Heliconius cydno alithea* in western Ecuador is controlled by two unlinked loci, one that controls colour (white/yellow) and one that controls pattern (presence/absence of melanin in a specific region of the forewing) [45]. In poison frogs, the genetic basis of colour variation is less well understood. Early crossing studies in *Oophaga pumilio* [46] suggested that pattern is probably controlled by a single locus with a dominant melanin-producing allele, whereas colour may be polygenic or controlled by a single locus with incomplete dominance. However, unlike *O. pumilio*, in which a major axis of variation in pattern is presence/absence of melanin, all known populations of *R. imitator* possess melanin on the dorsum, legs and venter. Thus, a more relevant task in the *R. imitator* system would be identifying the gene or genes that influence the spatial distribution of melanin rather than its presence or absence. Finally, in a field pedigree study [47], it was suggested that the red/yellow polymorphism in a population of *O. pumilio* was controlled by a single locus



where red coloration was completely dominant over yellow. Thus, our estimates of one to three genes controlling the mimetic phenotype of *R. imitator* are fairly comparable to other systems.

The method we have developed for identifying the number of genes controlling a phenotype obtains its information from the degree of clustering of phenotypes and from the dependence of variance in the phenotype on the admixture fraction. It can, as illustrated here, be used to distinguish between one, two or several genes, but is not expected to perform well in estimating the exact number of genes, when many genes are involved. In the presence of many genes, the information regarding clustering of phenotypes is lost. We note that the method can be sensitive to the precision in the estimate of the admixture fraction, and results of the method should be

interpreted accordingly. Implementations of the presented methods are publicly available at <https://github.com/schackv>.

**Ethics statement.** All research was conducted following an animal use protocol approved by the Animal Care and Use Committee of East Carolina University.

**Acknowledgements.** Research permits were obtained from the Ministry of Natural Resources (DGFSS) in Lima, Peru (authorizations no. 050-2006-INRENA-IFFS-DCB, no. 067-2007-INRENA-IFFS-DCB, no. 005-2008-INRENA-IFFS-DCB). Tissue exports were authorized under Contrato de Acceso Marco a Recursos Genéticos no. 0009-2013-MINAGRI-DGFSS/DGEFFS, with CITES permit number 003302.

**Funding statement.** Field research and collections were supported by an NSF DDIG (1210313) grant awarded to E.T. and K.S., and a National Geographic Society grant awarded to K.S. (8751-10).

## References

1. Endler JA. 1977 *Geographic variation, speciation and clines*, volume 10. Princeton, NJ: Princeton University Press.
2. Barton NH, Hewitt G. 1985 Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* **16**, 113–148. (doi:10.1146/annurev.es.16.110185.000553)
3. Coyne JA, Orr HA. 2004 *Speciation*, volume 37. Sunderland, MA: Sinauer Associates.
4. Szymura JM, Barton NH. 1991 The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution* **45**, 237–261. (doi:10.2307/2409660)
5. Szymura JM, Barton NH. 1986 Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution* **40**, 1141–1159. (doi:10.2307/2408943)
6. Turner JRG. 1971 Two thousand generations of hybridisation in a *Heliconius* butterfly. *Evolution* **38**, 233–243. (doi:10.2307/2407345)
7. Mallet J. 1986 Hybrid zones of *Heliconius* butterflies in Panama and the stability and movement of warning colour clines. *Heredity* **56**, 191–202. (doi:10.1038/hdy.1986.31)
8. Jiggins CD, Naisbit RE, Coe RL, Mallet J. 2001 Reproductive isolation caused by colour pattern mimicry. *Nature* **411**, 302–305. (doi:10.1038/35077075)
9. Chakraborty R, Weiss KM. 1988 Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc. Natl Acad. Sci. USA* **85**, 9119–9123. (doi:10.1073/pnas.85.23.9119)
10. Briscoe D, Stephens J, O'Brien S. 1994 Linkage disequilibrium in admixed populations: applications in gene mapping. *J. Hered.* **85**, 59–63.
11. Patterson N *et al.* 2004 Methods for high-density admixture mapping of disease genes. *Am. J. Hum. Genet.* **74**, 979–1000. (doi:10.1086/420871)
12. Gompert Z, Buerkle C. 2009 A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Mol. Ecol.* **18**, 1207–1224. (doi:10.1111/j.1365-294X.2009.04098.x)
13. Winkler CA, Nelson GW, Smith MW. 2010 Admixture mapping comes of age. *Annu. Rev. Genomics Hum. Genet.* **11**, 65–89. (doi:10.1146/annurev-genom-082509-141523)
14. Crawford JE, Nielsen R. 2013 Detecting adaptive trait loci in nonmodel systems: divergence or admixture mapping? *Mol. Ecol.* **22**, 6131–6148. (doi:10.1111/mec.12562)
15. Sheppard PM, Turner JRG, Brown KS, Benson WW, Singer MC. 1985 Genetics and the evolution of Müllerian mimicry in *Heliconius* butterflies. *Phil. Trans. R. Soc. Lond. B* **308**, 433–610. (doi:10.1098/rstb.1985.0066)
16. Counterman BA *et al.* 2010 Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genet.* **6**, e1000796. (doi:10.1371/journal.pgen.1000796)
17. Clarke CA, Sheppard PM, Thornton IW. 1968 The genetics of the mimetic butterfly *Papilio memnon* L. *Phil. Trans. R. Soc. Lond. B* **254**, 37–89. (doi:10.1098/rstb.1968.0013)
18. Joron M, Mallet JL. 1998 Diversity in mimicry: paradox or paradigm? *Trends Ecol. Evol.* **13**, 461–466. (doi:10.1016/S0169-5347(98)01483-9)
19. Kunte K, Zhang W, Tenger-Trolander A, Palmer D, Martin A, Reed R, Mullen S, Kronforst M. 2014 Doublesex is a mimicry supergene. *Nature* **507**, 229–232. (doi:10.1038/nature13112)
20. Twomey E, Yeager J, Brown JL, Morales V, Cummings M, Summers K. 2013 Phenotypic and genetic divergence among poison frog populations in a mimetic radiation. *PLoS ONE* **8**, e55443. (doi:10.1371/journal.pone.0055443)
21. Symula R, Schulte R, Summers K. 2001 Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proc. R. Soc. Lond. B* **268**, 2415–2421. (doi:10.1098/rspb.2001.1812)
22. Yeager J, Brown JL, Morales V, Cummings M, Summers K. 2012 Testing for selection on color and pattern in a mimetic radiation. *Curr. Zool.* **58**, 668–676.
23. Symula R, Schulte R, Summers K. 2003 Molecular systematics and phylogeography of Amazonian poison frogs of the genus *Dendrobates*. *Mol. Phylogenet. Evol.* **26**, 452–475. (doi:10.1016/S1055-7903(02)00367-6)
24. Brown JL *et al.* 2011 *A taxonomic revision of the neotropical poison frog genus Ranitomeya (Amphibia: Dendrobatidae)*. Auckland, New Zealand: Magnolia Press.
25. Camper J, Ruedas L, Bickham J, Dixon JR. 1993 The relationship of genome size with developmental rates and reproductive strategies in five families of neotropical bufonoid frogs. *Genetics* **12**, 79–87.
26. Castle W. 1921 An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. *Science* **54**, 223. (doi:10.1126/science.54.1393.223)
27. Wright S. 1968 *Evolution and the genetics of populations*. Chicago, IL: University of Chicago Press.
28. Zeng Z-B. 1992 Correcting the bias of Wright's estimates of the number of genes affecting a quantitative character: a further improved method. *Genetics* **131**, 987–1001.
29. Otto SP, Jones CD. 2000 Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics* **156**, 2093–2107.
30. Lande R. 1981 The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* **99**, 541–553.
31. Pearson K. 1904 *Mathematical contributions to the theory of evolution. XII. On a generalised theory of alternative inheritance, with special reference to Mendel's laws*. *Phil. Trans. R. Soc. Lond. A* **203**, 53–86. (doi:10.1098/rsta.1904.0015)
32. Slatkin M. 2013 A method for estimating the effective number of loci affecting a quantitative character. *Theor. Popul. Biol.* **89**, 44–54. (doi:10.1016/j.tpb.2013.08.002)
33. Barton N. 1983 Multilocus clines. *Evolution* **37**, 454–471. (doi:10.2307/2408260)
34. Barton N, Bengtsson BO. 1986 The barrier to genetic exchange between hybridising populations. *Heredity* **57**, 357–376. (doi:10.1038/hdy.1986.135)
35. Goodall C. 1991 Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. B* **3**, 285–339.

36. Koenderink JJ, van Doorn AJ. 1992 Surface shape and curvature scales. *Image Vision Comput.* **10**, 557–564. (doi:10.1016/0262-8856(92)90076-F)
37. Clemmensen LB, Hastie T, Witten D, Ersbøll B. 2011 Sparse discriminant analysis. *Technometrics* **53**, 37–41. (doi:10.1198/TECH.2011.08118)
38. Zou H, Hastie T. 2005 Regularization and variable selection via the elastic net. *J. R. Stat. Soc.* **67**, 301–320. (doi:10.1111/j.1467-9868.2005.00503.x)
39. Mika S, Ratsch G, Weston J, Scholkopf B, Muller K. 1999 Fisher discriminant analysis with kernels. Neural networks for signal processing IX, 1999. In *Proc. 1999 IEEE Signal Processing Society Workshop, August 1999* (eds Y-H Hu, J Larson, E Wilson, S Douglas), pp. 41–48. New York, NY: IEEE Inc.
40. Martin F. 2011 An application of kernel methods to variety identification based on SSR markers genetic fingerprinting. *BMC Bioinformatics* **12**, 177. (doi:10.1186/1471-2105-12-177)
41. Fletcher R. 1970 A new approach to variable metric algorithms. *Comput. J.* **13**, 317–322. (doi:10.1093/comjnl/13.3.317)
42. Twomey E, Vestergaard JS, Summers K. 2014 Reproductive isolation related to mimetic divergence in the poison frog *Ranitomeya imitator*. *Nat. Commun.* **5**, 4749. (doi:10.1038/ncomms5749)
43. Twomey E. 2014 Reproductive isolation and mimetic divergence in the poison frog *Ranitomeya imitator*. PhD thesis, East Carolina University, Greenville, NC, USA.
44. Mallet J, Barton N, Lamas G, Santisteban J, Muedas M, Eeley H. 1990 Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics* **124**, 921–936.
45. Chamberlain NL, Hill RI, Kapan DD, Gilbert LE, Kronforst MR. 2009 Polymorphic butterfly reveals the missing link in ecological speciation. *Science* **326**, 847–850. (doi:10.1126/science.1179141)
46. Summers K, Cronin T, Kennedy T. 2004 Cross-breeding of distinct color morphs of the strawberry poison frog (*Dendrobates pumilio*) from the Bocas del Toro Archipelago, Panama. *J. Herpetol.* **38**, 1–8. (doi:10.1670/51-03A)
47. Richards-Zawacki CL, Wang IJ, Summers K. 2012 Mate choice and the genetic basis for colour variation in a polymorphic dart frog: inferences from a wild pedigree. *Mol. Ecol.* **21**, 3879–3892. (doi:10.1111/j.1365-294X.2012.05644.x)