



## ORIGINAL RESEARCH

# A morphological trait involved in reproductive isolation between *Drosophila* sister species is sensitive to temperature

Alexandre E. Peluffo<sup>1</sup> | Mehdi Hamdani<sup>1</sup> | Alejandra Vargas-Valderrama<sup>1</sup>  |  
Jean R. David<sup>2,3</sup> | François Mallard<sup>4</sup>  | François Graner<sup>5</sup>  |  
Virginie Courtier-Orgogozo<sup>1</sup> 

<sup>1</sup>Institut Jacques Monod, CNRS, Univ. de Paris, Paris, France

<sup>2</sup>Institut Systématique Evolution Biodiversité (ISYEB), CNRS, MNHN, Sorbonne Université, EPHE, Paris, France

<sup>3</sup>Laboratoire Evolution, Génomes, Comportement, Biodiversité (EGCE), CNRS, IRD, Univ. Paris-sud, Université Paris-Saclay, Gif-sur-Yvette, France

<sup>4</sup>Institut de Biologie de l'École Normale Supérieure, CNRS UMR 8197, PSL Research University, Paris, France

<sup>5</sup>Matière et Systèmes Complexes, CNRS UMR 7057, Univ. de Paris, Paris, France

## Correspondence

Virginie Courtier-Orgogozo and Alexandre E. Peluffo, Institut Jacques Monod, CNRS, Univ. de Paris, 75013 Paris, France.

Email: virginie.courtier@normalesup.org (V. C.-O.); alexandre.peluffo@gmail.com (A. E. P.)

## Funding information

Agence Nationale de la Recherche, Grant/Award Number: ANR-11-IDEX-0005-02, ANR-11-LABX-0071 and ANR-18-IDEX-0001; FP7 Ideas: European Research Council, Grant/Award Number: 337579

## Abstract

Male genitalia are usually extremely divergent between closely related species, but relatively constant within one species. Here we examine the effect of temperature on the shape of the ventral branches, a male genital structure involved in reproductive isolation, in the sister species *Drosophila santomea* and *Drosophila yakuba*. We designed a semi-automatic measurement machine learning pipeline that can reliably identify curvatures and landmarks based on manually digitized contours of the ventral branches. With this method, we observed that temperature does not affect ventral branches in *D. yakuba* but that in *D. santomea* ventral branches tend to morph into a *D. yakuba*-like shape at lower temperature. We found that male genitalia structures involved in reproductive isolation can be relatively variable within one species and can resemble the shape of closely related species' genitalia through plasticity to temperature. Our results suggest that reproductive isolation mechanisms can be dependent on the environmental context.

## KEYWORDS

automatic detection, *Drosophila*, genitalia, machine learning, plasticity, reproductive isolation, shape analysis, speciation

## 1 | INTRODUCTION

Phenotypic plasticity, the capacity for one genotype to generate multiple phenotypes in response to environmental variation, is a pervasive feature of biological systems (Debat & David, 2001; Klingenberg, 2019). The connection between plasticity and speciation is multifaceted (Lafuente & Beldade, 2019). On the one hand, plasticity can be heritable and modified by selection. On the other hand, plasticity can favor adaptation and speciation. As animals colonize novel habitats or face changing climate conditions, the phenotypic

traits that are optimal for fitness are usually different from those experienced in the ancestral population. Waddington was among the first to suggest that organisms may solve this challenge by phenotypic plasticity first and later on by genetic fixation of what was previously an environmentally induced phenotypic trait (a process he called "genetic assimilation"; Waddington, 1942). According to several authors, the trait variations enabled by plasticity can initiate and accelerate the pace of adaptive evolution and promote morphological diversification. This central idea is at the basis of the "flexible stem hypothesis" (Schneider & Meyer, 2017; West-Eberhard, 2003)

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

and the “plasticity-first” model (Levis & Pfennig, 2016). A key feature of all these views is that the phenotypic change triggered by the plastic response, which allows the colonization of the new niches, is a phenocopy, that is, that the phenotypic change can be developmentally triggered by environmental variation or genetic variation interchangeably (Lafuente & Beldade, 2019). As we learn more about the genes mediating phenotypic plasticity (Gibert, 2017), it appears that similar phenotypic changes, either environmentally or genetically induced, can sometimes involve the same genetic loci. For example, the same enhancer of the gene *tan* contributes to both phenotypic plasticity in *Drosophila melanogaster* (Gibert et al., 2016) and interspecific evolution between sister species *Drosophila santomea* and *Drosophila yakuba* with respect to abdomen pigmentation (Jeong et al., 2008).

Depending on the setting, plasticity can either accelerate, slow down, or have little effect on evolution and species divergence (Price et al., 2003). Speciation, the process through which lineages diverge and become reproductively isolated, involves the accumulation over time of barriers limiting interbreeding, including divergence in ecological niches, behavioral isolation, and genomic incompatibilities (Coyne & Orr, 2004). As early as 1844, anatomical differences in genitalia between closely related species were proposed to be an essential mechanism maintaining reproductive isolation, as the so-called “lock-and-key” hypothesis (Dufour, 1844; Masly, 2011). In animals with internal fertilization, genitalia are the most rapidly evolving organs in terms of morphology (Eberhard, 1988), suggesting that a significant part of the speciation process involves anatomical divergence in genitalia. Alternatively, genital evolution can be a by-product of other evolutionary processes occurring within single lineages, independently of speciation (such as sexual selection), and lead to reproductive isolation as a by-product, when individuals attempt to hybridize with other lineages (Masly, 2011).

The lock-and-key hypothesis, even in species where it seems applicable, has been challenged by a variety of observations, including the facts that (1) genitalia in females do not differ as much as in males, (2) closely related species with conspicuous genital differences can still often produce hybrids, (3) males with laser-ablated genital organs can still copulate with no observed defect, and (4) genitalia morphology can be sensitive to temperature or nutrition (Andrade et al., 2005; Arnqvist & Thornhill, 1998; LeVasseur-Viens et al., 2015; Masly, 2011; Shapiro & Porter, 1989; Simmons, 2014 and references therein). It is thus possible that in some taxonomic groups interspecific differences in genital morphology do not contribute much to reproductive isolation.

To better comprehend the link between plasticity and speciation, careful examinations of particular cases are essential, and genital traits involved in reproductive isolation represent highly relevant model systems. How plastic are genitalia in general? Surprisingly, few studies have examined genitalia after raising organisms in various conditions. In the water strider *Aquarius remigis*, the mosquito *Aedes aegypti* and the fly *D. melanogaster*, changes in larval crowding, nutrition conditions or temperature were found to affect adult body size but had little effect on the size of the

external genitalia (Fairbairn, 2005; Shingleton et al., 2009; Wheeler et al., 1993). However, in two other species, the mosquito *Anopheles albimanus* and the fly *Drosophila mediopunctata*, the size and shape of the male intromittent organ were found to vary with rearing temperature (Andrade et al., 2005; Hribar, 1996). Overall, analysis of individuals sampled from the wild show that for a given arthropod or mammal species, the genitalia are usually more or less the same size whereas adult body size varies extensively (Dreyer & Shingleton, 2011; Eberhard et al., 1998 and references therein). These observations are concordant with the “lock-and-key hypothesis,” where male genitalia have to be of a particular size and shape to physically fit with the female genitalia. They are also explained by the “one-size-fits-all” hypothesis, where females appear to prefer males with genitalia of intermediate size (Eberhard et al., 1998).

In order to analyze and quantify the possible link between plasticity, reproductive isolation, and interspecific divergence, we chose to examine the effect of temperature on a male primary sexual trait likely involved in reproductive isolation between two *Drosophila* sister species, *D. santomea* and *D. yakuba*. These two species form an attractive system because their natural environment is relatively well characterized, they are known to hybridize, and one of their most remarkable morphological differences is a primary sexual trait that seems to be involved in a “lock-and-key” mechanism. *D. santomea* and *D. yakuba* diverged approximately 0.5–1 million years ago (Turissini & Matute, 2017). They can be crossed to generate fertile F1 females (Lachaise et al., 2000). *D. santomea* is endemic to the island of São Tomé, a volcanic island off the coast of Gabon (Lachaise et al., 2000), while *D. yakuba* is found in São Tomé and throughout sub-Saharan Africa (Lachaise et al., 1988, 2000). In São Tomé, *D. santomea* lives in the mist forests at high elevations while *D. yakuba* is found in open habitats associated with human presence, mostly at low elevations (Llopart et al., 2005a, 2005b). Both species co-occur at mid-elevation, around 1,150 m, and hybrids have been found consistently in this hybrid zone since its discovery in 1999 (Comeault et al., 2016; Cooper et al., 2018; Lachaise et al., 2000; Llopart et al., 2005a). *D. santomea* being insular, it is thought that this species originated from a common ancestor with *D. yakuba*, which colonized the island about 0.5–1 million years ago (Cariou et al., 2001; Llopart et al., 2002; Turissini & Matute, 2017) and that the present co-occurrence of *D. santomea* and *D. yakuba* in São Tomé reflects secondary colonization by *D. yakuba* from the African mainland, maybe during the last 500 years when Portuguese colonized the island (Cariou et al., 2001). Analysis of genomic and mitochondrial DNA sequences indicate that gene flow occurred between the *D. santomea* and *D. yakuba* more than 1,000 generations ago (Cooper et al., 2019; Turissini & Matute, 2017).

Multiple potential reproductive isolating mechanisms have been identified between the two species, such as genetic incompatibilities (Coyne et al., 2004; Moehring et al., 2006), ecological niche divergence (Matute et al., 2009), mate discrimination (Coyne et al., 2002; Lachaise et al., 2000), behavioral (Cande et al., 2012), physiological (Matute, 2010), and morphological differences (Jeong et al., 2008; Lachaise et al., 2000; Liu et al., 2019; Nagy et al., 2018).

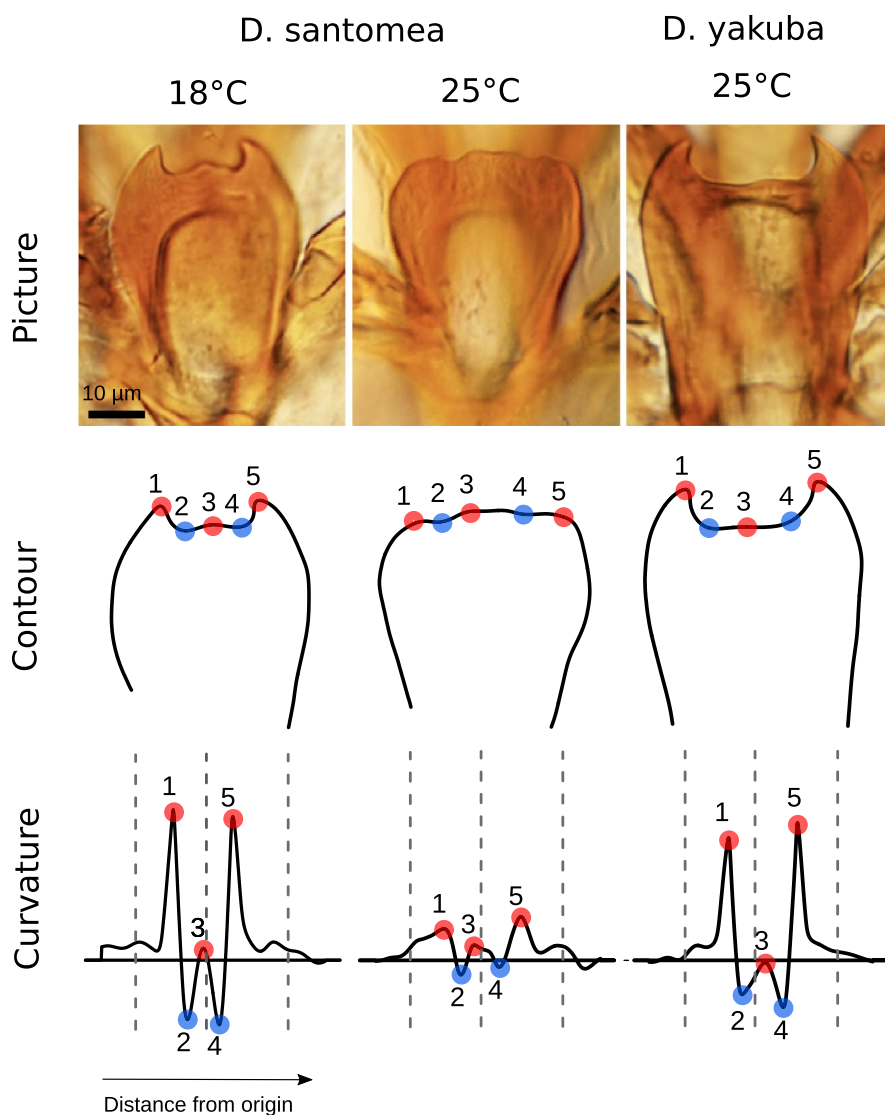
One reproductive isolating mechanism between *D. yakuba* and *D. santomea* involves a difference in ventral branches shape in the male genitalia and is the most conspicuous difference in male genitalia shape between the two species (Kamimura & Mitsumoto, 2012b; Yassin & Orgogozo, 2013; Figure 1). Ventral branches are located below the aedeagus (i.e., the insect phallus; Rice et al., 2019) and are only found in the *D. yakuba* complex, which comprises *Drosophila teissieri*, *D. yakuba*, and *D. santomea* (Yassin & Orgogozo, 2013).

In *D. yakuba*, spiny ventral branches insert inside female protective pouches during mating. In *D. santomea*, the male spines and female pouches are absent. These structures appear to play important roles during copulation. When mating with *D. yakuba* males, *D. santomea* females are wounded by the spines of the male ventral branches and they live shorter than females mating with conspecific males (Kamimura, 2012; Kamimura & Mitsumoto, 2012b; Matute & Coyne, 2010). Compared to *D. teissieri* females, *D. santomea* females also survive less to interspecific copulation with *D. mauritiana* (Yassin & David, 2016). Moreover, Kamimura and Mitsumoto (2012b) reported that “copulating pairs of *D. santomea* males × *D. yakuba* females dislodge readily when disturbed,” suggesting that

the spines may fasten genital coupling (Masly, 2011). We previously found that a major QTL on chromosome 3L contributes to the ventral branches shape difference between *D. santomea* and *D. yakuba* (Peluffo et al., 2015).

In São Tomé, the climate is very stable throughout the year, with only a 2.5°C-difference between the average daily temperature of the warmest month (March) and of the coldest one (July), and daily oscillations of about 5°C only (<https://en.climate-data.org/>, [www.worldclim.org/bioclimate](http://www.worldclim.org/bioclimate)). Based on temperature measurements at Monte Café (<https://en.climate-data.org/>), we estimate that the average temperature in the hybrid zone of Bom Sucesso (1153 m) varies between 15.5°C and 18°C throughout the year. In the wild, *D. santomea* flies are thus likely developing mainly at temperatures around 18°C or lower.

In previous studies of ventral branch shape, flies were raised either at 21°C (Yassin & Orgogozo, 2013) or 25°C (Kamimura, 2012; Kamimura & Mitsumoto, 2012b; Peluffo et al., 2015). Here, we report that *D. santomea* males raised at 18°C develop spiny ventral branches comparable to those of *D. yakuba* raised at 25°C. This is a surprising example where organs potentially directly linked with



**FIGURE 1** Landmark detection for ventral branches at 18°C and 25°C. For each individual, a picture of the ventral branches is taken (top panel). The contour is digitized by hand and smoothed (middle panel). The curvature along the contour is obtained by finite differences, which are iterated for refining; the resulting values of curvature are smoothed too (bottom panel). Smoothed curvature (vertical axis), measured in inverse micrometers, is plotted along the contour, starting from the leftmost point. The horizontal axis is the distance along the contour, called the curvilinear abscissa, measured in micrometers. Here both axes are normalized by size and represented in arbitrary units. In each plot, the left dashed vertical line is the automatic detection lower bound, the middle dashed line is the imputed global midpoint, and the right dashed line is the automatic detection upper bound (see methods). Red points represent peaks and therefore curvature maxima whereas blue points represent cavities and therefore curvature minima. Since *Drosophila yakuba* is not sensitive to temperature, we only show one characteristic shape. To understand where this genital structure is positioned within the male genitalia, see Figure 1 of Kamimura and Mitsumoto (2012b)

reproductive isolation undergo a plastic modification similar to the difference between two sister species. To better characterize the morphological change in ventral branches shape, we developed a user-friendly method to quantify contour curvatures and automatically detect spines using machine learning. We used it to examine the plastic response of ventral branches development at 18°C and 25°C both in newly collected wild strains and in strains kept in the laboratory for many years.

## 2 | MATERIAL AND METHODS

### 2.1 | Fly rearing and imaging

Fly strains (Table 1) were kept at 22°C on standard yeast-cornmeal-agar medium in uncrowded conditions before the beginning of the experiments. For each strain, roughly 20 individuals were transferred from the 22°C stock to either 18°C or 25°C, kept for a minimum of two nonoverlapping adult generations. Adult males were 5 to 7 days old when frozen at -80°C for subsequent dissection. Dissection of genitalia was performed in 1X PBS at room

temperature. Each genitalia was mounted on standard glass slides in DMHF (Dimethyl Hydantoin Formaldehyde, Entomopraxis) medium and kept overnight before imaging on an Olympus IX83 inverted station at 40x.

### 2.2 | Raw contour acquisition

All contours were digitized by the same person. Pictures were anonymized for manual contour acquisition so that the digitizer did not know the genotype. Digitization was skipped when the quality of the mounting was judged to be poor. For each picture, a custom ImageJ plugin was used to extract x, y coordinates (in pixels) of the contour. The plugin is designed to open all the pictures contained in a directory, allowing the user to manually draw a contour of the object of interest using the freehand tool of ImageJ. The raw contour is a series of points  $p_1, p_2, \dots, p_n$  in a two-dimensional space  $x, y$  where  $n$  is the number of points over which the contour passes (usually  $500 < n < 1,000$ ). The contour is open, and its endpoints are unimportant (Figure 1). It is analyzed (and twice smoothed) as follows.

**TABLE 1** List of Isofemale lines used in this study

Species	Name	Location	Year	Reference
<i>Drosophila yakuba</i>	Ivory Coast	Ivory Coast	1955	Cornell National Drosophila Species Stock Center, Strain #14021-0261.00 (given by D. Stern)
<i>D. yakuba</i>	BM2015	São Tomé, Bom Sucesso Botanical Garden, 1,150 m	February 2015	This study
<i>D. yakuba</i>	Oku	Cameroun, Mt. Oku, 2,000 m	April 2016	This study
<i>D. yakuba</i>	Raphia	Cameroun, Mt. Oku, 1,800 m	April 2016	This study
<i>Drosophila santomea</i>	STO.4	São Tomé, Obo Natural Reserve, submontane forest 1,300–1,450 m	1998	Lachaise et al. (2000). Cornell National Drosophila Species Stock Center, Strain #14021-0271.00 (given by D. Stern)
<i>D. santomea</i>	STO Cago 1482	São Tomé 1482 m	2001	Llopart et al. (2005b). This strain's original name is STO-LAGO 1,482 (given by D. Stern)
<i>D. santomea</i>	Quija22	São Tomé, Quija River, 650 m	2009	Gavin-Smyth and Matute (2013). This strain name is also Quija650.22 (given by D. Matute)
<i>D. santomea</i>	BM152	São Tomé Bom Sucesso Botanical Garden, 1,150 m	February 2015	This study
<i>D. santomea</i>	BM153	São Tomé Bom Sucesso Botanical Garden, 1,150 m	February 2015	This study
<i>D. santomea</i>	BM161	São Tomé Bom Sucesso Botanical Garden, 1,150 m	September 2016	This study
<i>D. santomea</i>	BM167	São Tomé Bom Sucesso Botanical Garden, 1,150 m	September 2016	This study
<i>D. santomea</i>	1563	EYFP laboratory strain derived from STO CAGO 1482, Insertion @ 3L:11.843.137	2001	Stern et al. (2017) (given by D. Stern)

Note: For each species, the most common name in the literature, the location, year of capture, and reference to origin of the strain are given. All lines are indicated in the same order as in Figure 2.

## 2.3 | Smoothed contour

The first layer of transformation is a rectangular smoothing filter over the raw contour to obtain the smoothed contour. At each point

$p_j$  with coordinates  $(x_j, y_j)$ , we derive  $p'_j$  with coordinates  $(x'_j, y'_j)$  where

$$x'_j = \frac{1}{2an} \sum_{i=j-an}^{j+an} x_i \text{ and } y'_j = \frac{1}{2an} \sum_{i=j-an}^{j+an} y_i.$$

Here the contour smoothing parameter  $\alpha$ , to be adjusted via learning, describes the proportion of points (relative to the total number of points forming the contour) to include in the smoothing. This implies that the smoothed contour is  $2\alpha n$  points shorter ( $\alpha n$  on each side) than the raw contour.

## 2.4 | Raw curvature of the smoothed contour

For each smoothed contour, the raw curvature  $k$  is computed with a sliding window of three points. For any set of three points  $M, N, P$  forming a triangle, the diameter of the circumscribed circle to this triangle,  $2r = MP/\sin(MN, NP)$  can be computed as the product of the Euclidean distances divided by the cross product of the two sides  $MN$  and  $NP$   $r$  is the curvature radius in  $N$ , and the curvature  $k$  in  $N$  is the inverse of  $r$ :

$$k = \frac{1}{r} = 2 \frac{|MN \times NP|}{MN \cdot NP \cdot PM}$$

The flatter the contour, the wider the circumscribed circle, the larger the radius  $r$ , and the smaller the curvature  $k$ . For each contour, the curvature profile is the curvature  $k_j$  computed over  $p'_j$  in  $[p'_1, p'_n]$  using its neighboring points ( $M, N, P = p'_{j-1}, p'_j, p'_{j+1}$ ) versus the curvilinear abscissa  $s_j$  of  $p'_j$  which is the sum of Euclidean distances from origin,  $s_j = p'_1 p'_2 + p'_2 p'_3 + \dots + p'_{j-1} p'_j$ .

## 2.5 | Refined curvature

We then use this first raw curvature estimation as information to refine the curvature in a second pass. In this second measure, the refined curvature  $k'_j$  is computed over an adaptive window of size  $a = \frac{1}{|k|}$  for  $k < 0.1$  and  $a = 10$  otherwise:  $M, N, P = p'_{j-a}, p'_j, p'_{j+a}$ . This means that the curvature is computed over a larger distance where it is small (and curvature radius is large), which requires more smoothing, without losing the sharpness of curvature peak determination where the curvature is large.

## 2.6 | Smoothed curvature

To improve curvature signal to noise ratio, for each point  $p'_j$  with coordinates  $(x'_j, y'_j)$  and refined curvature  $k'_j$ , we compute the

smoothed curvature  $k''_j$  as a weighted moving average with triangular weights:

$$k''_j = \frac{\sum_{i=j-\beta n}^{j+\beta n} w_i k'_i}{\sum_{i=j-\beta n}^{j+\beta n} w_i}$$

with  $w_j = \beta n, \dots, w_i = \beta n - |i - j|, \dots, w_{j-\beta n} = w_{j+\beta n} = 0$  and where  $\beta$  is the smoothing parameter to adjust via learning.  $\beta$  describes the proportion of points (relative to the total number of contour points) to include in the smoothing. This implies that the smoothed curvature contour is  $2\beta n$  points shorter ( $\beta n$  on each side) than the smoothed contour.

## 2.7 | Landmark detection

Curvature around the start and end of the contour is noisy; it corresponds to a region of low curvature, at the beginning and end of the contour, outside of the region where we expect to find the five landmarks (Figure 1). In addition, the contour digitization by the user, which tends to start at a precise point and to end in a long stroke, results in a slight left-right asymmetry in the curvature profile. After superimposing all smoothed curvature profiles, we choose to exclude the first and last 20% of the smoothed contour. We find that the axis of symmetry (midline) is at position 0.475 instead of 0.5 for a symmetric profile.

Landmarks are Bookstein's type 2 (local maxima of curvature) (Bookstein, 1992): maxima of the smoothed curvature for landmarks 1, 3, 5 and minima for landmarks 2 and 4. Having detected all minima and maxima, we first define landmark 3 as the maximum closest to the midline position, landmark 2 as the lowest minimum to the left of landmark 3 and landmark 1 as the maximum closest to landmark 2. Following the same logic, we define landmark 4 as the lowest minimum to the right of landmark 3 and landmark 5 as the maximum closest to landmark 4. Having detected all five landmarks, we found that there can seldom be more than one maximum between landmark 2 and landmark 4. In such a situation, we allow resampling of landmark 3 to the highest maximum between landmarks 2 and 4. Finally, we exclude individuals that do not display all five landmarks after detection.

## 2.8 | Spine thrust measure

Having detected all five landmarks, we quantify form using a measure previously introduced (Peluffo et al., 2015), which is highly correlated to the Procrustes analysis principal component measure of interspecific form variation and which we called "spine thrust" (ST). ST is a measure of how much spines are elevated above the central ridge of the ventral branches and is computed as:

$$ST = \frac{1}{2} (Y_{L1} + Y_{L5}) - Y_{L3}$$

where  $Y_{L1}$ ,  $Y_{L3}$ , and  $Y_{L5}$  are the Y coordinate of landmarks 1, 3, and 5, respectively. This measurement depends on the precise definition of X



and Y axes. Here the X-axis is defined as the axis passing by landmarks 2 and 4 and oriented from 2 to 4, and with the Y-axis defined so that (X, Y) is an oriented orthonormal basis.

## 2.9 | Machine learning

Detection of maxima and minima is a simple feature detection that relies on the derivative of the smoothed curvature profile. However, there are two parameters  $\alpha$ ,  $\beta$ , one for each smoothing filter (contour and curvature), which modulate the number and position of these detected maxima and minima. It is possible to explore a set of values for  $\alpha$  and  $\beta$  such that the correlation between manually digitized landmarks and automatically detected landmarks is optimized. Given that humans may introduce bias in the positioning of the landmarks (e.g., if one unconsciously amplifies spine thrust in *D. yakuba* relative to *D. santomea*), the human output may not be optimal over the machine output. This is why we chose not to quantify the learning success rate of our algorithm using the area under the receiver operating characteristic curve but instead to search for combinations of parameter values which yield the highest Pearson correlation value  $r^2$  for ST measured over manually digitized landmarks versus ST measured with automatically digitized landmarks.

## 2.10 | Statistical analyses

All statistical analyses were performed using R version 3.4.3 (R Core Team 2016). We performed two different sets of statistical analyses to investigate how ST changes across species, year of collection and temperature. First, we fitted a standard multiple linear regression with species, year of collection, and temperature as numeric predictors using the standard R function `lm()`. We chose the best model based on the variance explained provided by the  $r^2$  value. Table 2 presents the output of the `lm()` function using the R package `jtools` (v1.0.0; <https://cran.r-project.org/web/packages/jtools/jtools.pdf>) and its function `export_summs()` with "scale" and "transform.response" set to "TRUE" which scales and centers the response variable and reports standardized regression coefficients with their heteroskedasticity-robust standard errors. Second, we performed a regression tree analysis and performed cross-validation using recursive partitioning with the regression trees R package "rpart" version 4.1.13 (Therneau et al., 2018) and the associated function `rpart()` with the "ANOVA" method and obtained the approximate  $r^2$  from a 10-fold cross-validation using the `rsq.rpart()` function. To confirm the importance of each factor on ST change, we also performed random forest regression analysis using the R package "RandomForest" version 4.6.14 and the `randomForest()` function in order. Both sets of statistical analyses investigate the role of predictors in explaining a significant part of the variance, multiple linear regression allows the use of interaction terms while regression trees are easier to interpret (James et al., 2013). In addition to these analyses, we

**TABLE 2** Results of linear model fitting. This best model shows the contribution of each explanatory variable, considered as a numerical value, and their interactions to the overall variance of ST in the full *D. santomea*, *D. yakuba* dataset (shown in Figure 2a,b) at both temperatures (18°C and 25°C) across all years using the standard R function `lm()`. The standardized effect values and their heteroskedasticity-robust standard errors are reported together with the range of their  $p$ -values. For example, species having an overall effect of 1.67 implies that going from *D. santomea* to *D. yakuba* (all other things being equal) increases spine thrust absolute value by a relative (compared to the other effects), dimensionless, value of 1.67. The raw effects together with the full output of the model are provided in Table 3 (see Section 2)

Factor	Effect	SE	$p$ -value	Significance
N = 584, R <sup>2</sup> = 0.76				
Species	1.67	0.06	<0.001	***
Years	0.54	0.06	<0.001	***
Temperature	-0.29	0.05	<0.001	***
Species × Years	-0.56	0.07	<0.001	***
Species × Temperature	0.31	0.09	<0.001	***
Years × Temperature	-0.23	0.09	<0.05	*

Abbreviation: SE, standard error.

\* $p < 0.05$ , \*\*\*:  $p < 0.001$ .

systematically plot distribution of ST across predictors (Figure 2) showing individual values together with mean, standard errors (which directly inform about two-by-two statistical significance between groups), median, quartiles, and estimates of the 95% confidence interval of the medians, calculated as  $\pm 1.58 \times \frac{\text{IQR}}{\sqrt{n}}$  where IQR

is the interquartile range and  $n$  the number of individuals for that IQR (Chambers et al., 1983).

## 3 | RESULTS

### 3.1 | Spine thrust (ST) can be measured semi-automatically

We previously reported that the shape of ventral branches in *D. santomea*, *D. yakuba*, and their hybrids can be characterized with a set of five manually detected landmarks, which allows to calculate via simple arithmetic how much the lateral spines rise above the central ridge, as a quantitative value named "spine thrust" (ST), expressed in micrometers (Peluffo et al., 2015). The manual positioning of landmarks requires each point to be carefully positioned on the exact feature for the ST measure to be exact. It can introduce between-user and between-sample variability. In particular, the positioning of the three central landmarks can be equivocal and may differ between users.

To use a less biased approach and automate the process, we decided to develop a new measurement method that relies on manually digitized contours of the ventral branches, which are easier to define

**TABLE 3** Linear statistical model output using R lm (formula = ST ~ Species \* Year \* Temperature)

Residuals				
Min	1Q	Median	3Q	Max
-5.9035	-1.4710	-0.2689	1.3520	7.7223
Coefficients				
	Estimate	Standard error	t Value	Pr(> t )
(Intercept)	-750.5	189.2	-3.97	8.1e-05 <sup>***</sup>
spYak	680.5	213.8	3.18	1.5e-03 <sup>**</sup>
year	0.4	0.1	4.00	7.3e-05 <sup>***</sup>
temp	21.9	8.7	2.51	0.012 <sup>*</sup>
spYak:year	-0.3	0.1	-3.17	0.0016 <sup>**</sup>
spYak:temp	-16.5	9.8	-1.68	0.093
year:temp	-0.01	4.3e-03	-2.54	0.011 <sup>*</sup>
spYak:year:temp	0.01	4.8e-03	1.70	0.090

Note: Residual standard error = 2.125 on 576 degrees of freedom; Multiple R-squared: .7556; Adjusted R-squared: .7526; F-statistic: 254.3 on 7 and 576 df; p-value: <2.2e-16.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ .

than landmarks. The position contour of the ventral branches was digitized by hand at an approximately four times faster rate than landmark detection, because it can be done in a single stroke with a digital pen and the resulting ST measure is barely sensitive to the exact pen position. We designed a pipeline that automatically identifies the five landmarks based on the curvature of the manually digitized contours of the ventral branches and then calculates ST (Figure 1). The typical rounded form of *D. santomea* is then characterized by a null or negative value of ST (Figure 1, central panel) whereas the spiny form of *D. yakuba* is characterized by a positive value of ST (Figure 1, right panel). Note that our method does not separate size and shape (Klingenberg, 2016), but considers morphological form as a single quantifiable entity.

To assess repeatability, we digitized twice, at one-month interval (at the beginning and at roughly the midpoint of the digitizing effort), 30 individuals of the most characteristic *D. yakuba* strain (Oku, sharp spines) and 31 individuals of the *D. santomea* strain which is the most divergent from this *D. yakuba* strain (1563, extremely rounded shape and small spines). Despite a few outliers, we found a good correspondence, in the statistical sense, between the two sets of automatic measures (Figure 3), indicating that our pipeline produces robust statistical quantification of ventral branch form trends to distinguish both species and the continuum of forms between them.

### 3.2 | Learned $\alpha$ and $\beta$

We find that the same set of 30 *D. yakuba* and 31 *D. santomea* individuals is enough to identify optimal parameter values for  $\alpha$  (contour smoothing) and  $\beta$  (curvature smoothing).

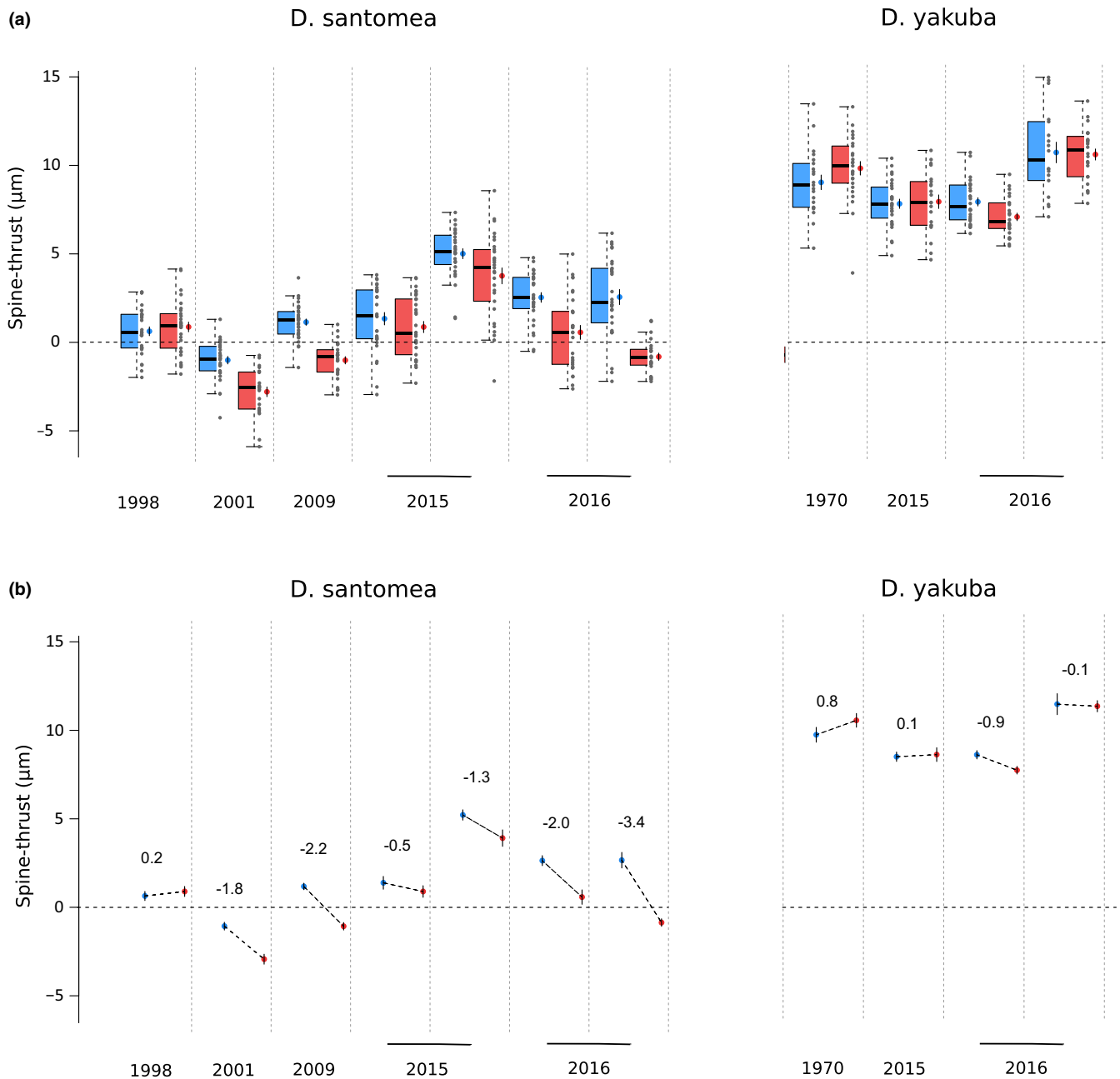
We find that with  $\alpha = 0.025$  and  $\beta = 0.055$  we obtain  $r^2 = 0.91$  (Figure 4). Although a few other combinations of  $\alpha$  and  $\beta$  yield the

same  $r^2$  (Figure 4), we choose this set because it is the one which applies the lowest degree of smoothing.

### 3.3 | Strong interspecific difference in ST

In total, with our semi-automated method (and after removing  $n = 71$  individuals incorrectly dissected or mounted, 12% of total samples, with no apparent distribution bias), we phenotyped 684 individuals raised at 18°C or 25°C throughout their development, corresponding to four *D. yakuba* lines and seven *D. santomea* lines collected between 1998 and 2016 (Table 1). We checked all the automatically detected landmarks by eyes and found that 30 individuals were incorrectly digitized, with a few landmarks either missing or aberrantly positioned (see Figure 5 for a sample of such individuals), and we excluded these individuals (4% of 684) from subsequent analysis. These aberrant landmark profiles were found in almost all the lines and at both temperatures, with no apparent distribution bias.

At 18°C and 25°C, for *D. santomea* and *D. yakuba*, in all 11 wild isofemale strains, we observed within-strain variability in ST values (Figure 2a, for all groups,  $n$  per group is between 26 and 31). At both temperatures, the mean ST of each of the seven *D. santomea* strains is inferior to the mean ST of any of the four *D. yakuba* strains (Figure 2a). All *D. yakuba* individuals have a positive ST, while most *D. santomea* strains have a mean ST close to 0 (Figure 2a). Accordingly, multiple linear regression analysis where the best fit model is  $ST \sim species \times years \times temperature$  shows that the species independent variable explains a significant part of the variance in ST ( $p < .001$ , Table 2). Overall, and despite within-strain variability and sensitivity to temperature variation, we confirm a morphological difference of ventral branches between wild strains of *D. santomea* and *D. yakuba* using our semi-automatic method of form quantification based on ST (Figure 2a; Table 2).



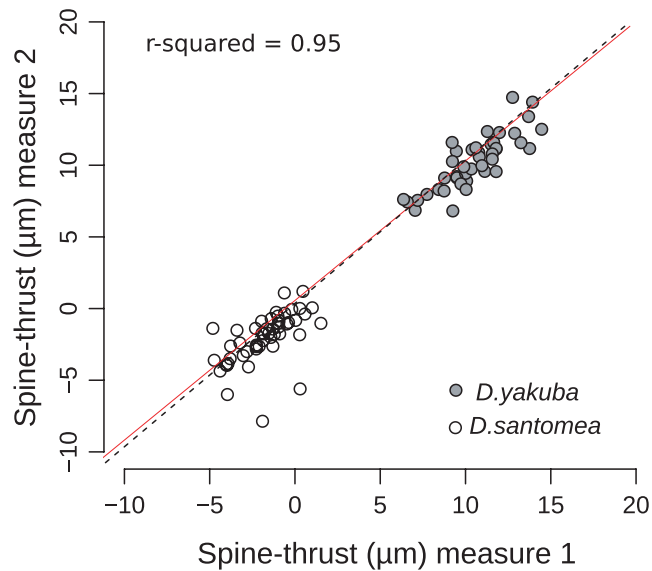
**FIGURE 2** Ventral branches form is sensitive to temperature in *Drosophila santomea*. Isofemale lines arranged by year of collection (see Table 1). (a) For each line, individual values (gray points), median (thick black line), quartiles (colored box plots), mean (colored points), standard errors (black vertical segment over the colored points) and 95% confidence interval estimates of the median (top and bottom notches) of automatically measured spine thrust are shown. Each line was reared at 18°C (blue) or 25°C (red). (b) For each line, the same mean and standard errors as in panel (a) are shown, together with the effect slope and corresponding value of that effect (in µm)

### 3.4 | Ventral branches of *D. santomea* are plastic to temperature whereas *D. yakuba* ventral branches are not plastic

For *D. santomea*, in all strains but the oldest one collected in 1998, the mean ST is systematically smaller at 25°C compared to 18°C and standard errors do not overlap (Figure 2). In contrast, no significant difference in mean ST between 25°C and 18°C is observed for *D. yakuba* strains, except for one strain collected in 2016 (*D. yakuba* Raphia)

(Figure 2b). Multiple linear regression analysis supports a negative effect of temperature, as seen with *D. santomea* ( $p < 0.001$ , Table 2) and that effect is dependent on species ( $p < 0.001$ , Table 2). For the most recently collected wild strain of *D. santomea* (BM16.2), we compared the contours of the two most representative individuals raised at 18°C and 25°C, that is, the two individuals with ST values closest to the median value of their group. We observed that the individual raised at 18°C has a more *D. yakuba*-like shape of ventral branches compared to the individual raised at 25°C (Figure 6).





**FIGURE 3** Correlation between two sets of automatic measures from the same dataset. For the training dataset (31 *Drosophila santomea* 1563 and 30 *Drosophila yakuba* Oku), the same user digitized the same contours twice at one-month interval and spine thrust was automatically measured. Each point represents one individual. The  $y = x$  (black dashed line) and linear regression (full red line) are shown

We find that the statistically significant effect of *temperature* on *D. santomea* is also statistically dependent on the *year* at which the strain was collected ( $p < 0.05$ , Table 2). In order to interpret our statistical analysis with multiple regression, we performed a 10-fold cross-validated regression tree analysis on the full dataset (2 species, 11 strains, 584 individuals). The 10-fold cross-validated error rate is 0.3% and using an additive model of the shape  $ST \sim species + years + temperature$ . We found that the variance in the dataset is first best partitioned by *species* and that *temperature* partitions the dataset best for strains collected in 2015 and 2016 (Figure 7, total variance explained as assessed by cross-validation  $r^2$  is 0.77). To confirm those results, we also performed a random forest regression analysis with the same model as for the regression tree and found that the overall variance explained is  $r^2 = 0.74$  and that the rank of importance of each independent variable is *species* > *years* > *temperature*. Altogether, our results show that in *D. santomea*, but not in *D. yakuba*, ventral branches are sensitive to temperature during development and that this effect is stronger in recently collected strains.

### 3.5 | The effect of temperature on spine thrust is as high as the effect of the major QTL between *D. yakuba* and *D. santomea*

To compare the effects of temperature and of interspecific genetic variation on ventral branch form, we used our previous QTL mapping dataset of ventral branch form between *D. santomea* and *D. yakuba*, which comprises 365 *D. santomea* backcross individuals (Peluffo et al., 2015). In this previous study, all flies were reared at 25°C as we found that

this temperature was optimal to rear both species. The five landmarks were placed manually on images of the ventral branches. A generalized Procrustes analysis was performed on a set of 365 backcross progeny individuals and a larger dataset including the backcross progeny, F1 hybrids, and parents. We found that, in both cases, the principal component PC1 explains an important part of the variance (58% in the full dataset and 41% in the backcross), that they are highly correlated ( $r^2 = 0.996$ ) and that PC1 in the backcross is highly correlated to ST (spine thrust;  $r^2 = 0.841$ ) and not to centroid size ( $r^2 = 0.038$ ).

This QTL mapping study revealed that a 2.7Mb locus on chromosome 3L explains 30% of the mean species difference in ST, meaning that replacing one *D. santomea* allele at this locus with a *D. yakuba* allele leads to an increase in ST of about 3 µm (30% of 9 µm, Peluffo et al., 2015). Pooling all the *D. santomea* lines examined in the present study, we find that a change in the raising temperature from 18°C to 25°C leads to an increase in ST of about 3.4 µm (Figure 8). We conclude that the effect of temperature is as high as the effect of genetic variation at the major interspecific genetic locus.

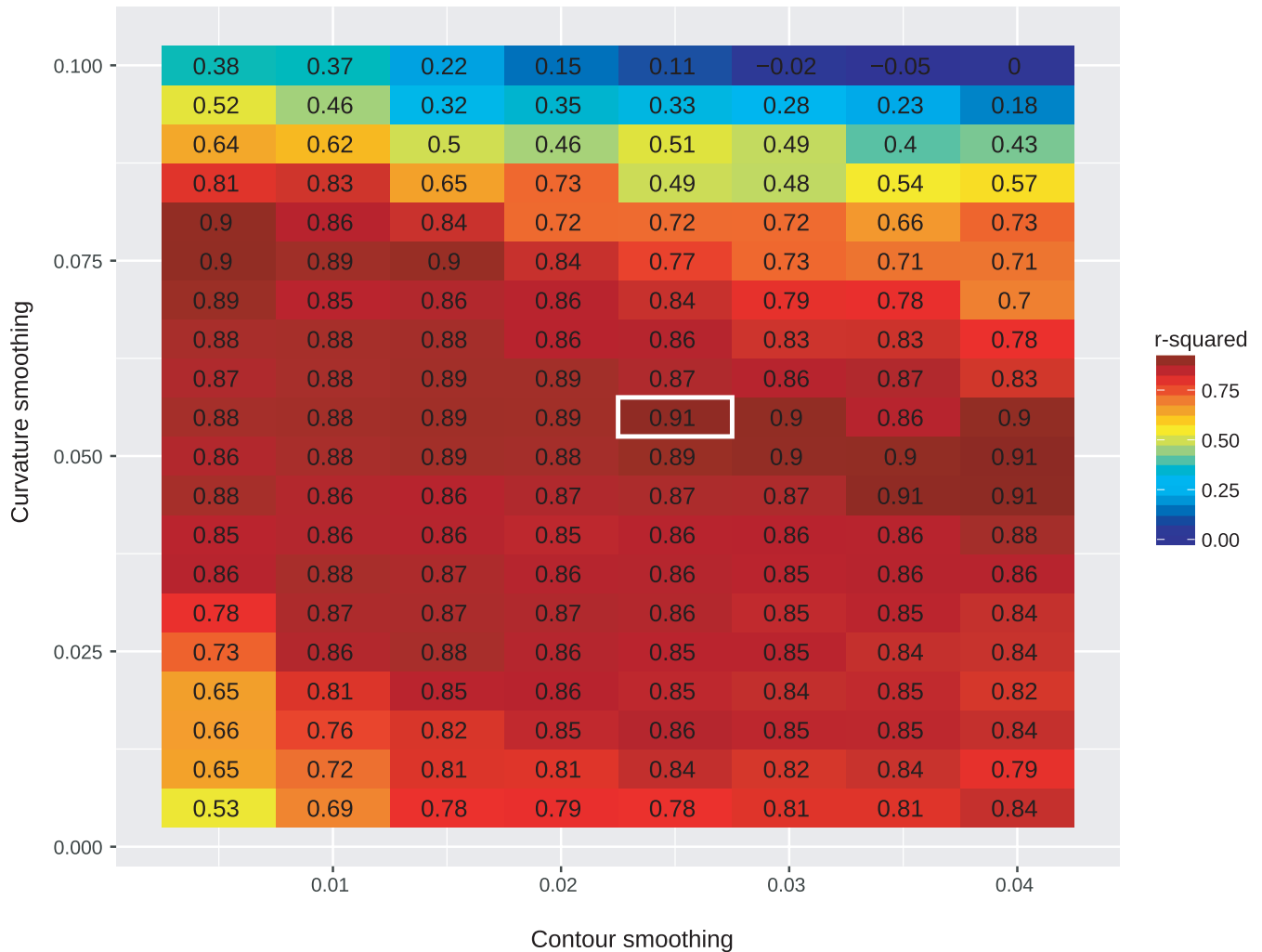
While the ST density distribution of all *D. yakuba* individuals shows little overlap with the ST density distribution of all *D. santomea* individuals reared at 25°C, it overlaps more with *D. santomea* individuals reared at 18°C (Figure 8). Overall, our results show that decreasing the temperature from 25°C to 18°C yield *D. santomea* males with spinier, *D. yakuba*-like, ventral branches in the same way as introgressing a *D. yakuba* alleles in place of a *D. santomea* allele at the major interspecific locus.

## 4 | DISCUSSION

### 4.1 | A dataset-independent simultaneous quantification of shape and size

Our semi-automatic method, which relies on two simple layers of contour transformation adjusted by regression based learning, is fast and allows the measure of form variation through the simple outlining of ventral branches on 2D pictures. We note that in the future, progress in edge detection algorithms (which for now introduce too much error to measure with precision variations of the order of a few micrometers) might allow full automation from pictures to form quantification.

Having drawn contours, we could also have relied on Fourier based analyses. However, such methods require closed contours which in our case are difficult to draw since the base of the ventral branches is a complex structure which cannot be easily delimited from the cuticle of the ventral branches (Figure 1). In addition, our method is more suitable for contours in which very large and very small curvatures coexist. Furthermore, an important limitation of morphometrics analyses on landmark data (e.g., Procrustes principal component analysis) is that the PC values are dimensionless (Klingenberg, 2010) and may be difficult to relate to physical features. With our simple measure of ST obtained from the automatically detected landmarks, we are able to quantify and compare forms



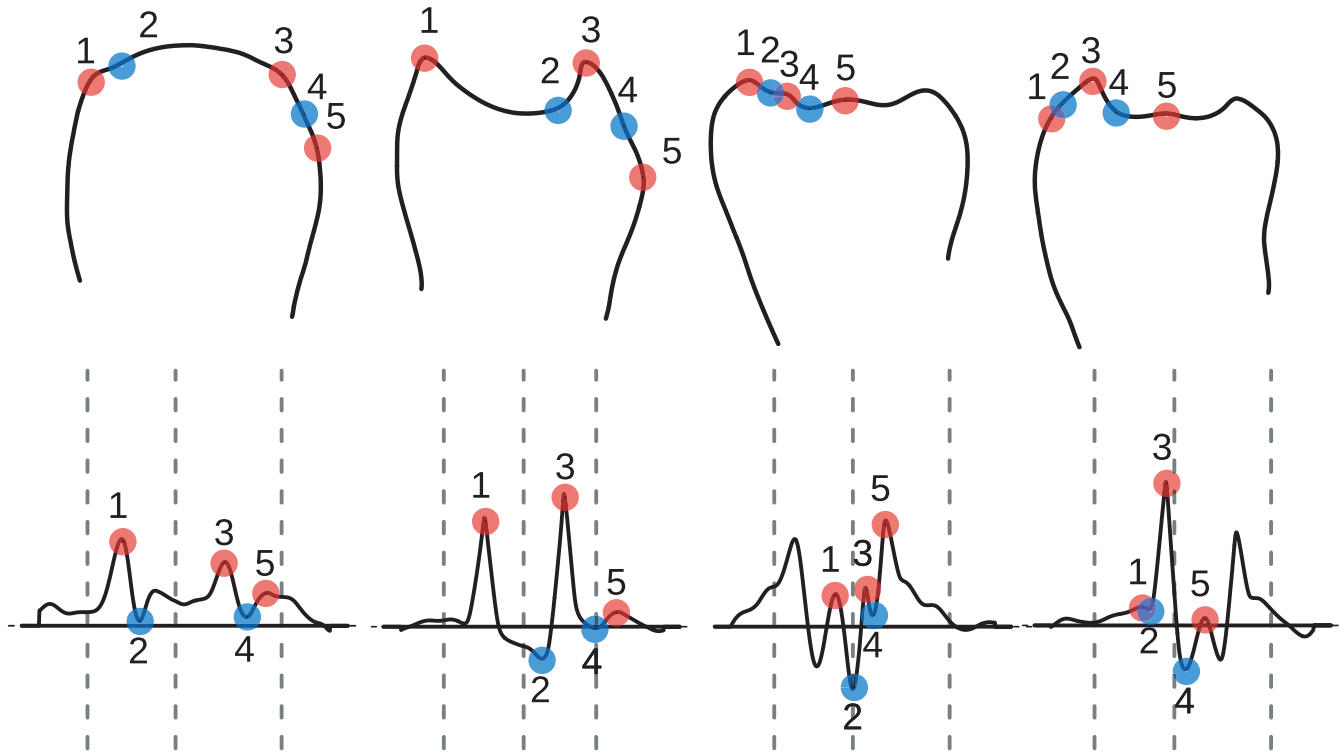
**FIGURE 4** Parameter adjustment for the machine detection algorithm. Training the algorithm relies on two layers of transformation that are each dependent on one parameter: contour coordinate smoothing (horizontal axis) and curvature profile smoothing (vertical axis). Training was performed using a set of 61 individuals, 31 *Drosophila santomea* 1563 and 30 *Drosophila yakuba* Oku (see Table 1) for which we manually digitized both landmarks and contours. For each value of the two smoothing parameters, we performed linear regression of spine thrust from manually digitized landmarks against spine thrust derived from automatically digitized landmarks. The colors and values represent the  $r^2$  from that regression. The value used for all detections is contoured in white

across studies. Importantly, because we deal with absolute geometric measurements, our method simultaneously analyzes shape and size, unlike most morphometric approaches (Claude, 2008; Klingenberg, 2010). We believe this to be a strength in our case since both shape and size of ventral branches probably contribute to the lock-and-key mechanism; for example, spiny but short *D. yakuba* ventral branches may not harm *D. santomea* females (Kamimura, 2012; Kamimura & Mitsumoto, 2012b).

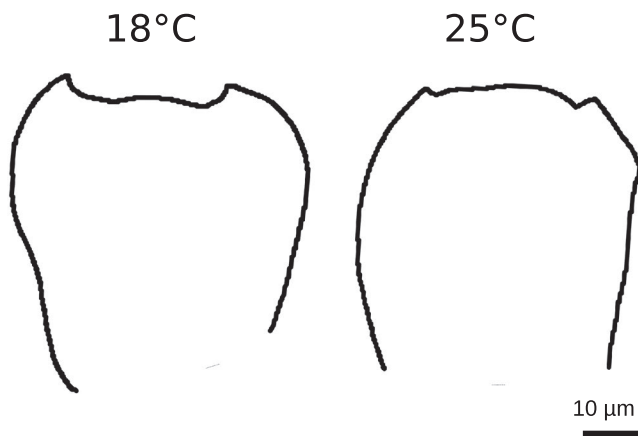
#### 4.2 | Effect of temperature on size and shape

In most insects and other ectotherms, adult body size typically increases with lower temperatures (Angilletta et al., 2004). Bergmann's rule, which posits an increasing body size with higher altitude, has been observed within the São Tomé island for the

terrestrial caecilian *Schistometopum thomense*, over a temperature range of 9°C (Measey & Van Dongen, 2006). In contrast to other body parts, the genitalia of insects, and of *D. melanogaster* in particular, have been reported as not, or little, plastic in response to temperature or other types of environmental variation (Eberhard, 2009; Shingleton et al., 2009; Wheeler et al., 1993). We find here that this is true also for *D. yakuba* but not for *D. santomea*: changing the rearing temperature from 25°C to 18°C leads to an increase in spine thrust in *D. santomea* male genitalia that is similar to what is observed between *D. santomea* and *D. yakuba*. In the present study, we only analyzed the effect of temperature on spine thrust, a scalar quantifier which captures one and only one characteristic of the whole shape. We did not examine whether the plasticity-induced change is affecting the entire shape of the ventral branches in the same way as the interspecific change. It would be interesting to include additional landmarks to capture



**FIGURE 5** Representative samples of landmarks incorrectly identified with the machine detection algorithm. For each example, we show the smoothed contour, the corresponding curvature profile and identified landmarks



**FIGURE 6** Difference in contour shape at 18°C versus 25°C within the same *Drosophila santomea* isofemale strain collected in 2016. Here are shown the contours of the two individuals that have the closest spine thrust value to the median value for *D. santomea* BM16.2 (right most strain on Figure 2a) at 18°C and 25°C

the entire shape of the genital structure and compare the changes in shape resulting from temperature variation and from interspecific difference (Noble et al., 2019).

The automatically detected landmarks could in principle be used to calculate the centroid size of the anatomical structure, and then, test whether the changes in ventral branches form triggered by temperature reflect heterogeneity in organ size variation within a strain among temperatures, among strains for a given temperature, or even a combination of both.

For each species, we find that strains raised in the same conditions display different averages in ST, showing that the ventral branches form is influenced by genetic factors and is able to evolve.

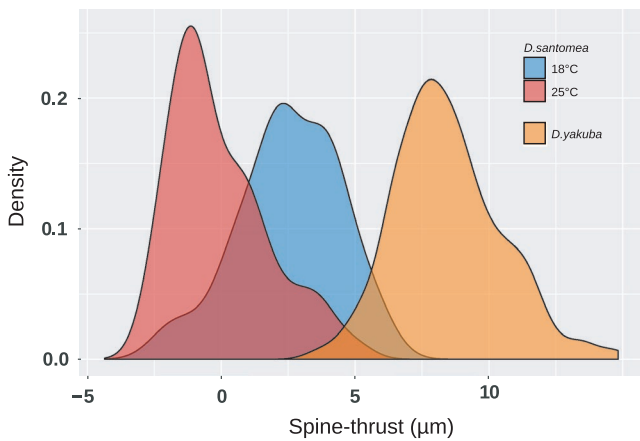
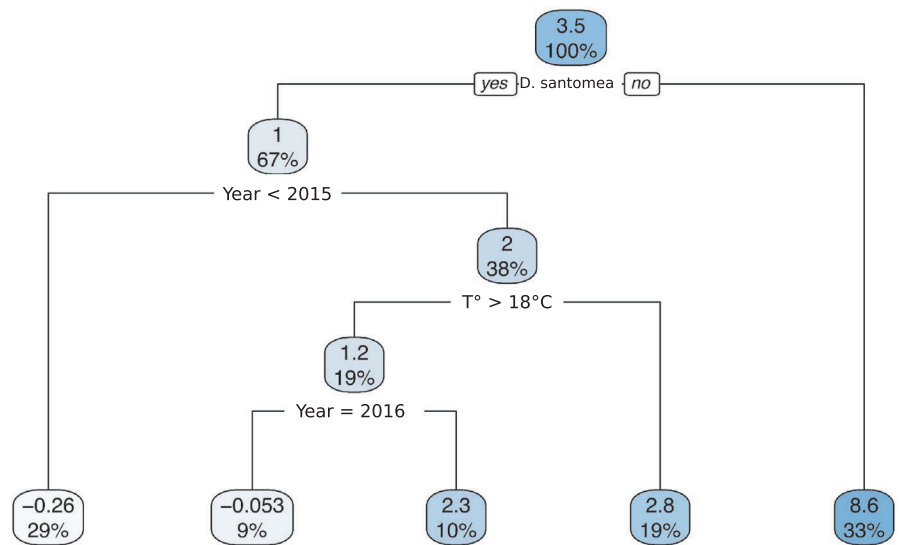
Plasticity of ventral branches form was detected for all the tested *D. santomea* strains except the one that was maintained for the longest time in the laboratory. Furthermore, the strains collected recently (in 2009, 2015, and 2016) display more pointed ventral branches at 18°C than the ones collected earlier. This suggests that as flies adapt to the laboratory environment, the plasticity of ventral branches form toward temperature tends to be lost and ventral branches tend to be more rounded. Recent studies show that *Drosophila* flies can adapt to a laboratory environment in 20 generations only, which corresponds to about 8 months (Langmüller & Schlötterer, 2020).

Based on our experiments, we cannot fully rule out plasticity in *D. yakuba*. It is possible that their genital morphology would be altered in external conditions outside of the specific ones that we assayed here. In any case, we find that in our experimental conditions the plasticity of genital form with respect to temperature is higher in *D. santomea* than *D. yakuba*.

### 4.3 | Laboratory observations should be complemented by analysis of wild-caught flies

Tests in the laboratory show that *D. santomea* flies appear to be poorly adapted to high temperatures (Matute et al., 2009). The optimal temperature for larval survival is 21°C for *D. santomea* and

**FIGURE 7** Regression tree for spine thrust measures of all *Drosophila santomea* and *Drosophila yakuba* isofemale strains at both 18°C and 25°C. Each node gives the spine thrust mean of all samples included in that node and the proportion. Below each node are two alternatives: to the left the condition is true and to the right the condition is false. Note that the split between *D. santomea* and *D. yakuba* happens at the top, thereby suggesting that neither temperature nor years have an effect on spine thrust within *D. yakuba*



**FIGURE 8** Density distribution of spine thrust values for *Drosophila santomea* lines at 18 and 25°C and *Drosophila yakuba* lines. Density distribution inferred from *D. santomea* males raised at 18°C (blue) and 25°C (red), *D. yakuba* males at both temperatures (orange). Distributions are inferred from the total data shown in Figure 2a

24°C for *D. yakuba*. Furthermore, when adult flies initially raised at 24°C are allowed to distribute themselves along a thermal gradient, they show a preference for 23°C for *D. santomea* and between 26°C and 27°C for *D. yakuba* (Matute et al., 2009). These observations are in agreement with *D. santomea* being collected at higher altitudes than *D. yakuba* in Sao Tomé. However, the reasons why the exact preferred temperature values observed in the laboratory are different from the temperature values measured in the geographic areas of the two species are unknown. Fly collections in Sao Tomé have mostly been done on the north slopes of the island and in these areas *D. santomea* flies are found at an altitude of 1,150 m or above (Lachaise et al., 2000), which corresponds to temperatures around 18°C or below (<https://en.climate-data.org/>, [www.worldclim.org/bioclim](http://www.worldclim.org/bioclim)). However, we note that on the southern slopes of the island a few *D. santomea* flies have also been collected at lower altitudes (650 m) in the dense mist forest near Rio

Queijo (Matute & Coyne, 2010; Nagy et al., 2018). This suggests that *D. santomea* flies can also inhabit warmer regions of the island and that they might be found across the native forest of Sao Tomé, which goes down to sea level on the western slope of the island (Bell & Irian, 2019). Interestingly, this type of coexistence is not unique on the island: two sister species of frogs closely match the distribution of *D. santomea* and *D. yakuba*, respectively, with the endemic species *Hyperolius thomensis* tied to wet forest habitats while its sister species *H. molleri* is in dry, human-disturbed areas, and *H. thomensis* frogs have also been found in the southern forest at 150 m (Bell & Irian, 2019).

It would be interesting to examine the genitalia of wild-caught individual males of *D. santomea* to check the form of their ventral branches at various altitudes. One possibility is that at low altitudes in the southern part of the island *D. santomea* flies display rounded ventral branches while in the hybrid zone with *D. yakuba* at 1,150 m, and at higher altitude, where temperatures are 18°C or below, they have spiner ventral branches. Of note, *D. santomea* flies have always been collected from traps and have never been observed directly in their native environment. It is possible that they live in microenvironments whose temperature is distinct from the one measured by climate stations (Feder et al., 2000; Negoua et al., 2019).

#### 4.4 | Evolution of the plasticity of ventral branches form

To understand the relevance of this temperature sensitivity of genital form for the past and present evolution of *D. santomea* and *D. yakuba*, more needs to be learnt about their ecology and the plasticity of the ventral branches form of their closely related species, *D. teissieri*. Ventral branches are only found in the three species of the *D. yakuba* complex, *D. santomea*, *D. yakuba*, and *D. teissieri* (Yassin & Orgogozo, 2013). Since ventral branch form plasticity has not been studied in *D. teissieri*, it is unclear whether this plasticity to

temperature is an ancestral trait which has been lost in *D. yakuba* or if it is a novel trait which evolved in *D. santomea* only. The species *D. teissieri* is not found in São Tomé but on the mainland and a few islands of the African continent; it can hybridize with *D. yakuba* (Cooper et al., 2018; Turissini & Matute, 2017). In *D. teissieri* males, the spines are very long and no layer of cuticle is present between them (Kamimura & Mitsumoto, 2012a; Yassin & Orgogozo, 2013). In any case, even if the extent of ventral branch form plasticity in *D. teissieri* was known, it would still be difficult to reconstruct ancestral trait states based on only three species.

The female protective pouches, into which the spiny ventral branches of *D. yakuba* males fit during copulation, were observed in *D. yakuba* but not in *D. santomea* females raised at 21°C and 25°C (Kamimura & Mitsumoto, 2012b; Yassin & Orgogozo, 2013). It would be interesting to check whether such pouches form in *D. santomea* females raised at 18°C, coinciding with the emergence of spiny ventral branches in males. Furthermore, whether more pointed ventral branches in *D. santomea* males due to lower temperatures affects copulation, reproduction, and female physiology after mating is unknown.

If we assume that the São Tomé island species *D. santomea* arose from a *D. yakuba*-like ancestor living on the African continent, one can hypothesize that regression in ventral branch size and their plasticity evolved recently in the lineage leading to *D. santomea*. Such a scenario is opposite to the most common view that posits that morphological diversification tends to proceed through losses of plasticity, rather than gains of plasticity (“flexible stem hypothesis”; Schneider & Meyer, 2017; West-Eberhard, 2003; “plasticity-first” model; Levis & Pfennig, 2016). It is possible that the decrease in spine thrust that occurred during evolution in the lineage leading to *D. santomea* was accompanied by a gain of ventral branches form plasticity toward temperature. It is unclear whether the plasticity of ventral branches form to temperature is adaptive. More knowledge about the ecology of *D. santomea* and its sister species will be required to elaborate a convincing scenario to interpret the role of the ventral branch form plasticity that we discovered.

## 5 | CONCLUSION

Our data show that genitalia can be plastic to temperature and that this plasticity can evolve coincidentally with speciation. Whereas the sensitivity of insect genitalia shape to temperature or nutrition has been used previously as a proof against the lock-and-key hypothesis (Andrade et al., 2005; Arnqvist & Thornhill, 1998), our work suggests that genitalia can be plastic without rejecting the lock-and-key hypothesis if the environmentally induced changes do not hamper reproduction within each sister species lineage.

## ACKNOWLEDGMENTS

We are grateful to São Tomé authorities for allowing us to collect flies. We thank David Stern and Daniel Matute for fly strains. We thank the Courtier laboratory for helpful discussions. We also thank

Philippe Rinaudo for feedback on the statistical analyses. The research leading to this paper has received funding from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 337579) to VCO and from the labex “Who am I?” (ANR-11-LABX-0071) and the Université de Paris IdEx (ANR-18-IDEX-0001) funded by the French government through grant no. ANR-11-IDEX-0005-02 to AEP.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTION

**Virginie Courtier-Orgogozo:** Conceptualization (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Project administration (equal); Validation (equal); Visualization (equal); Writing-review & editing (equal). **Alexandre E Peluffo:** Conceptualization (lead); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing-original draft (lead); Writing-review & editing (equal). **Mehdi Hamdani:** Data curation (equal); Resources (equal). **Alejandra Vargas-Valderrama:** Data curation (equal); Resources (equal). **Jean R. David:** Data curation (equal); Resources (equal); Writing-review & editing (equal). **François Mallard:** Methodology (equal); Software (equal); Writing-review & editing (equal). **François Graner:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing-review & editing (equal).

## DATA AVAILABILITY STATEMENT

The images, contours, scripts, and measurement values are available on DRYAD. <https://doi.org/10.5061/dryad.kpr4xh1f>.

## ORCID

Alejandra Vargas-Valderrama  <https://orcid.org/0000-0002-6003-3459>

François Mallard  <https://orcid.org/0000-0003-2087-1914>

François Graner  <https://orcid.org/0000-0002-4766-3579>

Virginie Courtier-Orgogozo  <https://orcid.org/0000-0002-9297-9230>

## REFERENCES

- Andrade, C. A., Hatadani, L. M., & Klaczko, L. B. (2005). Phenotypic plasticity of the aedeagus of *Drosophila mediopunctata*: Effect of the temperature. *Journal of Thermal Biology*, 30, 518-523. <https://doi.org/10.1016/j.jtherbio.2005.05.011>
- Angilletta, Jr., M. J., Steury, T. D., & Sears, M. W. (2004). Temperature, growth rate, and body size in ectotherms: Fitting pieces of a life-history puzzle. *Integrative and Comparative Biology*, 44, 498-509.
- Arnqvist, G., & Thornhill, R. (1998). Evolution of animal genitalia: Patterns of phenotypic and genotypic variation and condition dependence of genital and non-genital morphology in water strider (Heteroptera: Gerridae: Insecta). *Genetical Research*, 71, 193-212. <https://doi.org/10.1017/S0016672398003279>



- Bell, R. C., & Irian, C. G. (2019). Phenotypic and genetic divergence in reed frogs across a mosaic hybrid zone on São Tomé Island. *Biological Journal of the Linnean Society*, 128, 672-680.
- Bookstein, F. L. (1992). *Morphometric tools for landmark data: Geometry and Biology*. Cambridge University Press.
- Cande, J., Andolfatto, P., Prud'homme, B., Stern, D. L., & Gompel, N. (2012). Evolution of multiple additive loci caused divergence between *Drosophila yakuba* and *D. santomea* in wing rowing during male courtship. *PLoS One*, 7, e43888.
- Cariou, M. L., Silvain, J. F., Daubin, V., Da Lage, J. L., & Lachaise, D. (2001). Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line. *Molecular Ecology*, 10, 649-660.
- Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis*. Wadsworth & Brooks.
- Claude, J. (2008). *Morphometrics with R*. Springer Science & Business Media.
- Comeault, A. A., Venkat, A., & Matute, D. R. (2016). Correlated evolution of male and female reproductive traits drive a cascading effect of reinforcement in *Drosophila yakuba*. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20160730.
- Cooper, B. S., Sedghifar, A., Nash, W. T., Comeault, A. A., & Matute, D. R. (2018). A maladaptive combination of traits contributes to the maintenance of a *Drosophila* hybrid zone. *Current Biology*, 28, 2940-2947.
- Cooper, B. S., Vanderpool, D., Conner, W. R., Matute, D. R., & Turelli, M. (2019). Wolbachia acquisition by *Drosophila yakuba*-clade hosts and transfer of incompatibility loci between distantly related Wolbachia. *Genetics*, 212, 1399-1419.
- Coyne, J. A., Elwyn, S., Kim, S. Y., & Llopart, A. (2004). Genetic studies of two sister species in the *Drosophila melanogaster* subgroup, *D. yakuba* and *D. santomea*. *Genetical Research*, 84, 11-26.
- Coyne, J. A., Kim, S. Y., Chang, A. S., Lachaise, D., & Elwyn, S. (2002). Sexual isolation between two sibling species with overlapping ranges: *Drosophila santomea* and *Drosophila yakuba*. *Evolution*, 56, 2424-2434.
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates.
- Debat, V., & David, P. (2001). Mapping phenotypes: Canalization, plasticity and developmental stability. *Trends in Ecology & Evolution*, 16, 555-561.
- Dreyer, A. P., & Shingleton, A. W. (2011). The effect of genetic and environmental variation on genital size in male *Drosophila*: Canalized but developmentally unstable. *PLoS One*, 6, e28278. <https://doi.org/10.1371/journal.pone.0028278>
- Dufour, L. (1844). Anatomie générale des diptères. *Annales Des Sciences Naturelles*, 1, 244-264.
- Eberhard, W. G. (1988). *Sexual selection and animal genitalia*. Harvard University Press.
- Eberhard, W. G. (2009). Static allometry and animal genitalia. *Evolution*, 63, 48-66. <https://doi.org/10.1111/j.1558-5646.2008.00528.x>
- Eberhard, W. G., Huber, B. A., Briceño, R. D., Salas, I., & Rodriguez, V. (1998). One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. *Evolution*, 52, 415-431.
- Fairbairn, D. J. (2005). Allometry for sexual size dimorphism: testing two hypotheses for Rensch's rule in the water strider *Aquarius remigis*. *The American Naturalist*, 166, S69-S84.
- Feder, M. E., Roberts, S. P., & Bordelon, A. C. (2000). Molecular thermal telemetry of free-ranging adult *Drosophila melanogaster*. *Oecologia*, 123, 460-465.
- Gavin-Smyth, J., & Matute, D. R. (2013). Embryonic lethality leads to hybrid male inviability in hybrids between *Drosophila melanogaster* and *D. santomea*. *Ecology and Evolution*, 3, 1580-1589.
- Gibert, J.-M. (2017). The flexible stem hypothesis: Evidence from genetic data. *Development Genes and Evolution*, 227, 297-307.
- Gibert, J.-M., Mouchel-Vielh, E., De Castro, S., & Peronnet, F. (2016). Phenotypic plasticity through transcriptional regulation of the evolutionary hotspot gene tan in *Drosophila melanogaster*. *PLoS Genetics*, 12, e1006218.
- Hribar, L. J. (1996). Larval rearing temperature affects morphology of *Anopheles albimanus* (Diptera: Culicidae) male genitalia. *Journal of the American Mosquito Control Association*, 12, 295-297.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (Eds.) (2013). Tree-based methods. In *An Introduction to Statistical Learning with Applications in R* (pp. 303-335). Springer.
- Jeong, S., Rebeiz, M., Andolfatto, P., Werner, T., True, J., & Carroll, S. B. (2008). The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell*, 132, 783-793. <https://doi.org/10.1016/j.cell.2008.01.014>
- Kamimura, Y. (2012). Correlated evolutionary changes in *Drosophila* female genitalia reduce the possible infection risk caused by male copulatory wounding. *Behavioral Ecology and Sociobiology*, 66, 1107-1114. <https://doi.org/10.1007/s00265-012-1361-0>
- Kamimura, Y., & Mitsumoto, H. (2012a). Genital coupling and copulatory wounding in *Drosophila teissieri* (Diptera: Drosophilidae). *Canadian Journal of Zoology*, 90, 1437-1440.
- Kamimura, Y., & Mitsumoto, H. (2012b). Lock-and-key structural isolation between sibling *Drosophila* species. *Entomol. Sci.*, 15, 197-201. <https://doi.org/10.1111/j.1479-8298.2011.00490.x>
- Klingenberg, C. P. (2010). Evolution and development of shape: Integrating quantitative approaches. *Nature Reviews Genetics*, 11, 623-635. <https://doi.org/10.1038/nrg2829>
- Klingenberg, C. P. (2016). Size, shape, and form: concepts of allometry in geometric morphometrics. *Development Genes and Evolution*, 226, 113-137. <https://doi.org/10.1007/s00427-016-0539-2>
- Klingenberg, C. P. (2019). Phenotypic plasticity, developmental instability and robustness: The concepts and how they are connected. *Frontiers in Ecology and Evolution*, 7, 56.
- Lachaise, D., Cariou, M.-L., David, J. R., Lemeunier, F., Tsacas, L., & Ashburner, M. (1988). Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evolutionary Biology*, 159-225.
- Lachaise, D., Harry, M., Solignac, M., Lemeunier, F., Bénassi, V., & Cariou, M. L. (2000). Evolutionary novelties in islands: *Drosophila santomea*, a new melanogaster sister species from São Tomé. *Proceedings. Biological Sciences*, 267, 1487-1495.
- Lafuente, E., & Beldade, P. (2019). The genomics of developmental plasticity: Recent progress in animal models. *Frontiers in Genetics*, 10, 720.
- Langmüller, A. M., & Schlötterer, C. (2020). Low concordance of short-term and long-term selection responses in experimental *Drosophila* populations. *Molecular Ecology*, 29, 3466-3475. <https://doi.org/10.1111/mec.15579>
- LeVasseur-Viens, H., Polak, M., & Moehring, A. J. (2015). No evidence for external genital morphology affecting cryptic female choice and reproductive isolation in *Drosophila*. *Evolution*, 69, 1797-1807.
- Levis, N. A., & Pfennig, D. W. (2016). Evaluating 'plasticity-first' evolution in nature: Key criteria and empirical approaches. *Trends in Ecology & Evolution*, 31, 563-574. <https://doi.org/10.1016/j.tree.2016.03.012>
- Liu, Y., Ramos-Womack, M., Han, C., Reilly, P., Brackett, K. L. R., Rogers, W., Williams, T. M., Andolfatto, P., Stern, D. L., & Rebeiz, M. (2019). Changes throughout a genetic network mask the contribution of hox gene evolution. *Current Biology*, 29, 2157-2166. <https://doi.org/10.1016/j.cub.2019.05.074>
- Llopart, A., Elwyn, S., Lachaise, D., & Coyne, J. A. (2002). Genetics of a difference in pigmentation between *Drosophila yakuba* and *Drosophila santomea*. *Evolution*, 56, 2262-2277. <https://doi.org/10.1111/j.0014-3820.2002.tb00150.x>
- Llopart, A., Lachaise, D., & Coyne, J. A. (2005a). An anomalous hybrid zone in *Drosophila*. *Evolution*, 59, 2602-2607. <https://doi.org/10.1111/j.0014-3820.2005.tb00972.x>



- Llopart, A., Lachaise, D., & Coyne, J. A. (2005b). Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. *Genetics*, *171*, 197-210.
- Masly, J. P. (2011). 170 years of "lock-and-key": Genital morphology and reproductive isolation. *International Journal of Evolutionary Biology*, *2012*, 1-10. <https://doi.org/10.1155/2012/247352>
- Matute, D. R. (2010). Reinforcement of gametic isolation in *Drosophila*. *PLoS Biology*, *8*, e1000341.
- Matute, D. R., & Coyne, J. A. (2010). Intrinsic reproductive isolation between two sister species of *Drosophila*. *Evolution*, *64*, 903-920.
- Matute, D. R., Novak, C. J., & Coyne, J. A. (2009). Temperature-based extrinsic reproductive isolation in two species of *Drosophila*. *Evolution*, *63*, 595-612.
- Measey, G. J., & Van Dongen, S. (2006). Bergmann's rule and the terrestrial caecilian *Schistometopum thomense* (Amphibia: Gymnophiona: Caeciliidae). *Evolutionary Ecology Research*, *8*, 1049-1059.
- Moehring, A. J., Llopart, A., Elwyn, S., Coyne, J. A., & Mackay, T. F. (2006). The genetic basis of postzygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to hybrid male sterility. *Genetics*, *173*, 225-233.
- Nagy, O., Nuez, I., Savaisar, R., Peluffo, A. E., Yassin, A., Lang, M., Stern, D. L., Matute, D. R., David, J. R., & Courtier-Orgogozo, V. (2018). Correlated evolution of two copulatory organs via a single cis-regulatory nucleotide change. *Current Biology*, *28*, 3450-3457.
- Negoua, H., Chakir, M., David, J. R., & Cappy, P. (2019). Climatic adaptation in *Drosophila*: Phenotypic plasticity of morphological traits along a seasonal cycle. *Annales De La Société Entomologique De France*, *55*, 48-60.
- Noble, D. W., Radersma, R., & Uller, T. (2019). Plastic responses to novel environments are biased towards phenotype dimensions with high additive genetic variation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 13452-13461.
- Peluffo, A. E., Nuez, I., Debat, V., Savaisar, R., Stern, D. L., & Orgogozo, V. (2015). A major locus controls a genital shape difference involved in reproductive isolation between *Drosophila yakuba* and *Drosophila santomea*. *G3 Genes Genomes Genetics*, *5*, 2893-2901. <https://doi.org/10.1534/g3.115.023481>
- Price, T. D., Qvarnström, A., & Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*, 1433-1440.
- R Core Team (2016). *R: A language and environment for statistical computing (R version 3.4.3)*. R Foundation for Statistical Computing. <https://www.R-project.org>
- Rice, G., David, J. R., Kamimura, Y., Masly, J. P., Mcgregor, A. P., Nagy, O., Noselli, S., Nunes, M. D. S., O'Grady, P., Sánchez-Herrero, E., Siegal, M. L., Toda, M. J., Rebeiz, M., Courtier-Orgogozo, V., & Yassin, A. (2019). A standardized nomenclature and atlas of the male terminalia of *Drosophila melanogaster*. *Fly*, *13*, 51-64.
- Schneider, R. F., & Meyer, A. (2017). How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Molecular Ecology*, *26*, 330-350.
- Shapiro, A. M., & Porter, A. H. (1989). The lock-and-key hypothesis: evolutionary and biosystematic interpretation of insect genitalia. *Annual Review of Entomology*, *34*, 231-245.
- Shingleton, A. W., Estep, C. M., Driscoll, M. V., & Dworkin, I. (2009). Many ways to be small: Different environmental regulators of size generate distinct scaling relationships in *Drosophila melanogaster*. *Proceedings of the Royal Society B-Biological Sciences*, *276*, 2625-2633.
- Simmons, L. W. (2014). Sexual selection and genital evolution. *Austral Entomology*, *53*, 1-17.
- Stern, D. L., Crocker, J., Ding, Y., Frankel, N., Kappes, G., Kim, E., Kuzmickas, R., Lemire, A., Mast, J. D., & Picard, S. (2017). Genetic and transgenic reagents for *Drosophila simulans*, *D. mauritiana*, *D. yakuba*, *D. santomea*, and *D. virilis*. *G3 Genes Genomes Genetics*, *7*, 1339-1347.
- Therneau, T., Atkinson, B., & Ripley, B. (2018). rpart: Recursive partitioning and regression trees. R Package. version, 4.1-13.
- Turissini, D. A., & Matute, D. R. (2017). Fine scale mapping of genomic introgressions within the *Drosophila yakuba* clade. *PLoS Genetics*, *13*, e1006971.
- Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, *150*, 563-565. <https://doi.org/10.1038/150563a0>
- West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.
- Wheeler, D., Wong, A., & Ribeiro, J. M. (1993). Scaling of feeding and reproductive structures in the mosquito *Aedes aegypti* L. (Diptera: Culicidae). *Journal of the Kansas Entomological Society*, *66*, 121-124.
- Yassin, A., & David, J. R. (2016). Within-species reproductive costs affect the asymmetry of satyrization in *Drosophila*. *Journal of Evolutionary Biology*, *29*, 455-460.
- Yassin, A., & Orgogozo, V. (2013). Coevolution between male and female genitalia in the *Drosophila melanogaster* species subgroup. *PLoS One*, *8*, e57158. <https://doi.org/10.1371/journal.pone.0057158>

**How to cite this article:** Peluffo AE, Hamdani M, Vargas-Valderrama A, et al. A morphological trait involved in reproductive isolation between *Drosophila* sister species is sensitive to temperature. *Ecol Evol*. 2021;11:7492-7506. <https://doi.org/10.1002/ece3.7580>