

Application of mark-recapture models to study fitness landscapes in Darwin's finches

Marc-Olivier Beausoleil

Master's thesis submitted to McGill University

Department of Biology
McGill University, Montréal
April 2017

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree
of Master of Science

© Marc-Olivier Beausoleil, 2017

Table of contents

Abstract	4
Résumé.....	5
Acknowledgements.....	6
Preface and Contribution of Authors	8
List of tables.....	9
List of figures	10
Estimating fitness landscapes within a species complex of Darwin’s finches	12
Introduction.....	12
Fitness Landscapes with multiple species.....	13
Temporal variation in selection regime	14
Material and methods.....	17
Data collection	17
Principal component analysis (PCA) and morphospace.....	19
Statistical Model and inputs.....	19
Modeling the unobserved (state-space) and observed rates of survival	20
Linear and quadratic estimation of survival and recapture probabilities	22
Spline estimation of survival and recapture probabilities.....	23
Definition of model priors	24
Fitness landscape	25
Results.....	26
Ground finch species show distinct beak morphology	26
Species fitness landscapes are variable across species and years	29

(i) Topography: numbers of peaks, positions, gradient elevation.....	29
(ii) Effect of rainfall on the topology of the quadratic model	29
Fitness landscapes are more variable with a species specific effect	30
(i) Topography of a thin-plate spline model with a shared fitness function	31
(ii) Topography of a thin-plate spline model for each species.....	31
Discussion	42
Topologies differ on single species compared to multispecies fitness landscapes	42
(i) Splines with independent fitness surfaces characterize more variable landscapes	43
(ii) Spline analysis of shared fitness surfaces	45
Temporal variation in species-specific patterns of beak evolution.....	46
Niche expansion from <i>G. fortis</i> to <i>G. magnirostris</i>	48
Summary	49
Appendix	50
Principal component Analysis	50
Methods description	50
PCA Results	52
Discussion	54
Validation and Simulations of the Bayesian model	55
Literature Cited	65

Abstract

Ecological speciation, the formation of new species due to divergent selection, is considered to be a major driver of the production of biodiversity (Wright, 1932; Simpson, 1953; Schluter & Nychka, 1994; Benkman, 2003; Gavrilets, 2004; Svensson & Calsbeek, 2012; Martin, 2016). Fitness landscapes, which visually depict the relationship between trait or genotypic variation and fitness, can help to understand the opposing forces of divergent selection and homogenizing gene flow during population differentiation (Wright, 1932; Simpson, 1953; Schluter & Nychka, 1994; Benkman, 2003; Gavrilets, 2004; Svensson & Calsbeek, 2012; Martin, 2016). However, fitness landscapes are typically generated for single species, which impedes their utility for understanding the process of speciation in hybridizing species. In addition, each fitness landscape shows a static representation of a particular selection regime. Thus, they may not capture the longer-term dynamics that occur within complexes of introgressing species, especially when temporal variation in the environment causes shifts in the relationships between trait variation and fitness. To address these issues, I construct fitness landscapes of multiple species across a long-term dataset of Darwin's finches. Specifically, I use mark-recapture models to quantify the relationships between phenotype, climate, and survival probabilities in four species of Darwin's finches (*Geospiza fuliginosa*, *G. fortis*, *G. magnirostris*, and *G. scandens*) coexisting on the island of Santa Cruz in the Galápagos Archipelago. Differences in the topology of these landscapes help to shed light on the role of temporally varying selection in both facilitating and impeding ecological speciation in this species complex.

Résumé

La spéciation écologique, c'est-à-dire la formation de nouvelles espèces en raison d'une sélection divergente, est considérée comme un facteur majeur de production de la biodiversité (Schluter & Nagel, 1995; Schluter, 2001; Nosil, 2012; Shafer & Wolf, 2013). Les paysages adaptatifs représentent visuellement la relation entre un ou plusieurs traits, parfois la variation génotypique, et la valeur sélective (*fitness*). Ils peuvent aider à comprendre l'effet de forces opposées de la sélection divergente et l'homogénéisation du flux de gènes lors de la différenciation de populations (Wright, 1932; Simpson, 1953; Schluter & Nychka, 1994; Benkman, 2003; Gavrilets, 2004; Svensson & Calsbeek, 2012; Martin, 2016). Cependant, les paysages adaptatifs sont souvent générés pour une seule espèce, ce qui empêche leur utilité pour comprendre le processus de spéciation dans les espèces hybrides. En outre, chaque paysage adaptatif montre une représentation statique d'un régime de sélection particulier. Ainsi, ils ne peuvent pas montrer la dynamique à plus long terme qui se produit dans un complexe d'espèces hybride, surtout lorsque la variation temporelle dans l'environnement provoque des changements dans les relations entre les traits et la valeur sélective. Pour résoudre ces problèmes, je construis des paysages adaptatifs de plusieurs espèces de pinsons de Darwin (*Geospiza*) avec un jeu de données d'une étude à long terme. Plus précisément, j'utilise des modèles de capture-marquage-recapture pour quantifier les relations entre le phénotype, le climat et les probabilités de survie de quatre espèces (*Geospiza fuliginosa*, *G. fortis*, *G. magnirostris* et *G. scandens*) qui coexistent sur l'île Santa Cruz aux îles Galápagos. Les différences dans la topologie de ces paysages adaptatifs contribuent à éclairer le rôle temporellement variable de la sélection, parfois en facilitant ou en ralentissant la spéciation écologique dans ce complexe d'espèces.

Acknowledgements

For the past two years, I learned that graduate studies are not only about learning science, but how to develop new responsible citizens with critical thinking. I also realized that science alone is not enough to understand our world. Multidisciplinary and transdisciplinary work are much needed to develop a future based on equality and equity. I wish that we have better grounds to appreciate complexity and variation in our environment through time to integrate and share our knowledge maintaining a cultural and historical light in our mind guiding our research.

My Master's has been a wonderful adventure benefiting from the discussion of many students and knowledgeable faculty members. More specifically, I want to thank my supervisory committee members Rowan Barrett, Andrew Hendry and Jon Sakata for their numerous comments to channel my master's degree on an enlightened path, which I still voyage on, of wisdom.

Rowan Barrett and Andrew Hendry lab members (Ananda Martins, Antoine Paccard, Charles Cong Xu, Krista Oke, José Jonathas Pereira de Lira, Juntao Hu, Léa Blondel, Timothy J. Thurman and Victor Frankel Vilches) for their patience, generous help and insightful discussions. I want to thank Charles Darwin Research Station (CDRS), Galápagos National Parks Service, Caroline Leblond, Kiyoko Gotanda, Sarah Knutie, Angela Hansen, Luis Fernando De León, Joost Raeymaekers, Sofia Carvajal Endara, Jaime A. Chaves, Jeff Podos, Diana Shapre, Nancy Emery, Marc Johnson, Eartwatch volunteers and all the data collectors. They have collected the data and transferred the knowledge about the system.

Coding advices were provided by Luke Owen Frishkoff and Leithen M'Gonigle. They helped me figure out how Bayesian statistics work although I was trained as a "frequentist". Countless

hours of discussion were spent to learn answer the many questions that I have. Also, Andres López-Sepulcre, Olivier Gimenez kindly helped me with the statistics.

I am grateful to the people that support my choices and push me to pursue my objectives. In particular, my parents Lucie Bousquet and Claude Beausoleil put my mind at ease in my graduate degree while keeping my flame lit for biology. Also, I am privileged to have wonderful and brilliant friends inside or outside academia that help me to keep a healthy balance. Especially to Benjamin J. Allard who is always refreshing my mind and Roxanne Richard for her love and care. To all of my friends who helped me build my statistical and programming skills, the ones helping me to realize my ideas respecting common values, and everybody working in our society helping me to concentrate on my research.

This research is supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants Program to Rowan Barrett and an NSERC Canada Graduate Scholarships and Fonds de recherche du Québec, Nature et technologies to Marc-Olivier Beausoleil. This thesis was conducted at McGill University on a traditional and unceded territory of the Kanien'keha:ka (Mohawk). Broken links between our nations need to be restored.

Preface and Contribution of Authors

Data was provided by Andrew Hendry, but was collected by a team of colleagues for twelve consecutive years. I conducted the analyses in collaboration with Luke Frishkoff, Leithen M'Gonigle, and Rowan Barrett. I prepared the thesis, on a manuscript-based style, with inputs from Rowan Barrett, Luke Frishkoff and Leithen M'Gonigle. This thesis will be formatted for future publication in a peer-reviewed journal.

List of tables

Table I Number of birds captured and capture effort per year at El Garrapatero	28
Table II Selection gradients and parameters estimated from the quadratic model.....	33
Table III Parameters estimated from the thin-plate spline model.	36
Table IV Mean and variance of traits used in the simulation analysis of empirical data	57

List of figures

Figure 1 Map of the major Galápagos Islands	18
Figure 2 Correlation double projection of the principal component analysis of all birds including three beak traits; median beak length (MBL), width (MBW) and depth (MBD).	28
Figure 3 Fitness landscape for ground finches for beak size (PC1 scores) and beak shape (PC2 scores) estimated with a quadratic model including linear and quadratic terms.....	35
Figure 4 Fitness landscape per year for ground finches with 95% credible intervals for beak size (the model included only PC1 scores) estimated with a thin-plate spline.	37
Figure 5 Fitness landscape per year for ground finches with 95% credible intervals for beak shape (the model included only PC2 scores) estimated with a thin-plate spline.	38
Figure 6 Fitness landscape per year for ground finches with 95% credible intervals for beak size (the model included only PC1 scores) estimated with a thin-plate spline.	39
Figure 7 Fitness landscape per year for ground finches with 95% credible intervals for beak shape (the model included only PC2 scores) estimated with a thin-plate spline.	40
Figure 8 Climatic data from Puerto Ayora, a town 10 km from El Garrapatero, recorded by Charles Darwin Research Station (CDRS).....	41
Figure S1 Density of the data in relation to the position of the objects in the LDA (x-axis) for the 3 LDA axis.	50
Figure S2 Graphical results from PCA showing PC1 and PC2 of the beak traits (length, width, and depth).....	53
Figure S3 Quadratic model convergence and effective sampling.....	54
Figure S4 Credible interval (95%) of the species per year.	55

Figure S5 Simulation analysis of the mean value of the random intercept of species effect on survival and PCA of simulated data.....	58
Figure S6 Posterior distribution of the generalized linear mixed effect model	59
Figure S7 Traceplot of the generalized linear mixed effect model.....	60
Figure S8 Autocorrelation plot of the generalized linear mixed effect model.....	61
Figure S9 Posterior distribution (density graph) and traceplot of the thin-plate spline model..	62
Figure S10 Autocorrelation plot of the thin-plate spline model.	63
Figure S11 Principal component analysis of the four ground finches.	64

Estimating fitness landscapes within a species complex of Darwin's finches

Introduction

Ecological speciation, the formation of new species due to divergent selection, has three main components: 1. A source of divergent natural selection, 2. A form of reproductive isolation, and 3. A genetic mechanism linking divergent selection to reproductive isolation (Schluter, 2001; Nosil, 2012). It is considered as a major driver of the emergence of novel biodiversity (Schluter & Nagel, 1995; Shafer & Wolf, 2013). Quantifying phenotypic differences among related species within particular ecological contexts can provide crucial information about the role of natural selection during this process. The first attempt to portray the relationship between phenotypes (or genotypes) and fitness as a continuous landscape was described by Wright (Wright, 1932) but many variations of the idea exist (Kimura, 1983; Gavrillets & Gravner, 1997; Arnold *et al.*, 2001; Gavrillets, 2004; Dieckmann *et al.*, 2004; Siepielski *et al.*, 2009). The premise is that individuals with different phenotypes might have different relative fitness, and this relationship can be captured using a two-dimensional topographic-like map or a tridimensional rugged plane representing “mountains” (high fitness phenotype combinations) and “valleys” (low fitness phenotype combinations). Conceptually, the metaphor is simple to understand and also permits a direct heuristic visualization of the form of natural selection (directional, stabilizing, or disruptive) in different populations or species. It has been applied in multiple ecological and evolutionary contexts (McCoy, 1979; Schluter & Nychka, 1994; Benkman, 2003; Gavrillets, 2004; Svensson & Calsbeek, 2012), and is useful for establishing a link between putatively adaptive phenotypes and fitness (but see Kaplan, 2008).

Fitness Landscapes with multiple species

In the context of adaptive radiation, constructing a fitness landscape with individuals from only one species does not permit investigation of the interactions between fitness functions of co-existing species that compete for resources and may also be able to exchange genes (e.g. Hendry *et al.*, 2009). Recently, it has been proposed that the study of multiple species in a shared fitness landscape can help understand niche diversification during adaptive radiation (Martin & Wainwright, 2013). However, using multiple species in fitness landscapes poses a key question about the connectivity of the surface between species: should species sharing similar phenotypes be constrained to have the same fitness (Hendry & Gonzalez, 2008)? For closely related species consuming similar resources, one might assume a fairly consistent relationship between trait variation and fitness (i.e., the topology of the fitness landscape). On the other hand, it has been argued that two individuals from distinct species that have the same trait value (e.g., a certain beak size) might not share the same fitness value (Sober, 2001). Most obviously, other unmeasured traits may contribute to overall fitness, and equating fitness of an organism with only a subset of their traits might not represent all of the components of fitness such as lifetime reproductive success and survival. Additionally, differences in ecology and behaviour between species may cause sharp differences in relationship between trait variation and fitness between species (e.g., niche differentiation). One way to test this assumption would be to compute independent fitness landscapes for each species and investigate differences in their topologies. In this case, the fitness values of individuals from each species are estimated independently of those from other species. An alternative approach would be with a model that computes fitness functions without a species effect. In other words, fitness could be assessed between populations of different species, but without modelling the species in the model. In this case, one assumes that the ecology and behaviour of the different species studied are similar.

Temporal variation in selection regime

Fitness landscapes with multiple peaks can reflect niche diversification if peaks coincide with high frequency trait values (Martin & Wainwright, 2013; Martin, 2016; Roches *et al.*, 2016). For example, the radiation of Darwin's finches is thought to be driven by the availability of distinct food resources, which is in turn influenced by climate. It has been shown that mean beak depth of species of finches living on different islands are positioned where the food is most available (Schluter & Grant, 1984). Moreover, species that share similar trait values in allopatry show differentiation when inhabiting the same islands, suggesting divergent selection and character displacement (Schluter & Grant, 1984; Grant, 1986; Grant & Grant, 2006).

In times of resource limitations, the finches can quickly deplete the resources that they prefer feeding on, necessitating shifts to lower quality food types (Smith *et al.*, 1978; Grant, 1986). In addition, climatic events on the Galápagos islands strongly affect the distribution of available seed sizes (Abbott *et al.*, 1977; De León *et al.*, 2014). In drought years, the plants do not reproduce by seed whereas in wet years, seeds are plentiful. These climatic fluctuations can cause variation in the selection coefficients on beak traits (Grant & Grant, 2002). A consequence of this temporal variation is that computing a fitness landscape at a single point in time may yield misleading inferences about longer-term patterns of evolution. Conversely, using data compiled across multiple years could mask the directionality of selection if there is positive selection for a trait in one year and negative selection in another (no net selection) (Siepielski *et al.*, 2009). This latter scenario could slow the progress towards speciation by reducing the temporal consistency of divergent selection (Hendry *et al.*, 2009). To enable more robust understanding of the connections between selection regime and evolutionary trajectories of this species complex, I use

a long-term dataset to find the association between beak traits and the effect of climate (rainfall) on the estimation of survival probabilities.

Here, I investigate fitness landscapes using mark-recapture models applied to four finch species (*Geospiza fuliginosa*, *G. fortis*, *G. magnirostris*, and *G. scandens*) on Santa Cruz Island. I ask two questions: 1) What are the topologies (location of adaptive peaks and gaps) of fitness landscapes when considering multiple hybridizing species on a single landscape versus independent fitness landscapes for each species? 2) What is the impact of temporal variation in the environment (rainfall in our model) on phenotype-fitness relationships?

1) Topologies of fitness landscapes in a species complex

On a multi-species fitness landscape, I hypothesize that a model that does not include a species-specific term in the fitness function will simplify the fitness landscape by reducing the peak number, because the fitness function will account for less variation and will be more influenced by the global mean of overall fitness. In contrast, when using independent fitness functions for each species, I predict that the topologies will generally show multiple high fitness peaks around the means of the respective beak morphologies of each species, because these should correspond to distinct feeding habits (Schluter & Grant, 1984).

To investigate the topology of fitness landscapes, I first build a morphospace (Raup, 1967) using a principal component analysis on three beak traits (length, depth, and width). I then use a generalized mixed model and thin-plate spline (connected polynomials, which offer a smoother fit of the function (Schluter & Nychka, 1994)) to construct a "Simpsonian" fitness landscape based on mark-recapture data (Simpson, 1944; 1953; Kéry & Schaub, 2012). I use the generalized linear mixed model to compute deterministic selective gradients as well as the

topology of the fitness landscapes, and the thin-plate spline to infer the topology of the fitness landscapes. An important caveat is that selection gradients are estimated through viability selection only (i.e. survival estimated from mark-recapture data). Thus, I am constructing landscapes using survival probability, a component of fitness (e.g. Gimenez *et al.*, 2009; Martin, 2016) and I am omitting differences in fitness due to variation in reproductive success that may be associated with trait variation (here, beak morphology) (Ratcliffe & Grant, 1983). By estimating survival based on a hierarchical model, I construct a fitness landscape for multiple-species under the assumption of an open population model.

2) Do changes in rainfall alter the topologies of fitness landscapes?

Darwin's finch species show specialization on different food types (De León *et al.*, 2014), and the availability of food types may be altered under wet versus dry conditions. Therefore, I expect that not all species will respond in the same way to changes in rainfall (e.g., a drought). For example, the cactus finch that eats *Opuntia* flowers has a reliable source of food every year, whereas other species rely on seed production by plants are strongly impacted by climatic conditions. To test the effect of rainfall on the topology of the fitness landscape, I will use the generalized mixed model (later referred to as the "quadratic model" for simplicity) mark-recapture models, described in the previous section, and add an interaction term between phenotypes (beak traits summarized in a principal component analysis) and rainfall. I use rainfall (the amount of rain in millimetres in one year) as a climate variable because it is the main factor that affects the abundance of food resources for the ground finches (Grant, 1986; Grant & Grant, 2008).

In summary, fitness landscapes can be used to test how environmental variation alters relationships between traits and fitness (or a surrogate such as survival) in closely related, interacting species. Here, I pursue this goal using a mark-recapture model to estimate survival probabilities in sympatric finch species. Investigating changes in the topology of multispecies fitness landscapes using long-term datasets will help to understand the role of temporally varying selection in both facilitating and impeding ecological speciation in Darwin's finches.

Material and methods

Data collection

Our team captured individuals from four ground finch species (*Geospiza magnirostris*, *G. fortis*, *G. fuliginosa*, and *G. scandens*) between 2003 and 2014 at El Garrapatero, an arid coastal zone (coordinates: 0° 4' 15.7" S, 90° 13' 18.3" W) on Santa Cruz Island in the Galápagos Archipelago, Ecuador (Huber *et al.*, 2007; De León *et al.*, 2010; 2012; Raeymaekers *et al.*, n.d.). Our team collected data during the breeding season of the finches between January and April. This period of the year corresponds to the wet season, marked by an increase in precipitation. All species are able to consume similar food types during the wet season, but during the dry season, *G. magnirostris*, *G. fortis*, *G. fuliginosa* forage more heavily on seeds while *G. scandens* consumes mostly flower-based food types composed primarily of nectar, fruits and buds (Smith *et al.*, 1978; Grant, 1986; De León *et al.*, 2014).

Our team captured and banded birds using mist-nets placed at haphazard locations (Figure 1). We determined the sex of each individual by examining beak and feather colour as well as the brood patch if it was caught during the mating season. Following (Grant & Grant, 1995), we measured beak length (anterior edge of nares to anterior tip of the upper mandible), beak depth

(at the nares), and beak width (at the base of the lower mandible) with calipers. To control for variation in the duration of sampling across years that may impact recapture estimates, our team recorded sampling effort as the number of hours that the nets were deployed in one year.

A meteorological database is available for Puerto Ayora on the Charles Darwin foundation website (<http://www.darwinfoundation.org/datazone/climate/>). Instead of using the raw climatic data between calendar years, I calculated the amount of rain (as a continuous variable) per year between the median date of sampling seasons (which usually has a duration of 4 months) by summing rainfalls between the median dates. As such, the capture history reflects what birds have experienced in the year prior to our sampling.

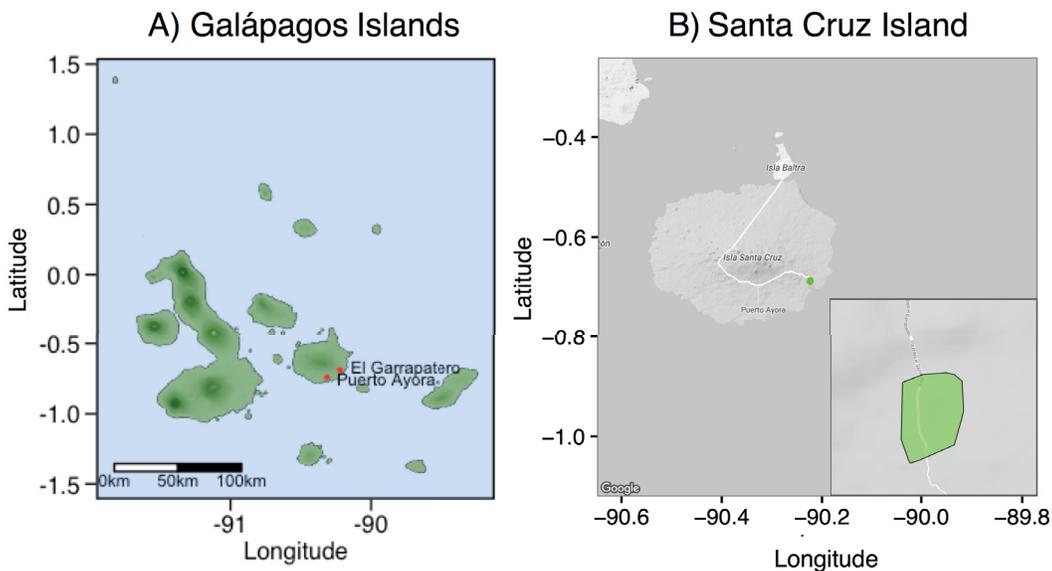


Figure 1 Map of the major Galápagos Islands

(A) The two red dots represent Puerto Ayora, one of the largest cities on Santa Cruz Island, and our sampling site, El Garrapatero. The map and the elevation data comes from CARTO (with the OpenStreetMap (OSM) source) and utilize the Web Mercator projection, Datum WGS84. (B) Santa Cruz Island and El Garrapatero (inset) is near a

road, but 10 km away Puerto Ayora. In the inset, the green area represents the sampling site of the birds. Maps from Google Maps (2017), projection Web Mercator, Datum WGS84.

Principal component analysis (PCA) and morphospace

I pooled all of the species sampled from the beginning of our study in order to characterize the morphospace of the beak traits. This enables me to use the principal component scores as phenotypes that are indicative of species identity. I analyze only adults because they tend to be loyal to their breeding localities. Also, banded birds are rarely seen beyond our sampling sites (Hendry *et al.*, 2009). I used a principal component analysis based on the covariance matrix because all the beak traits were on the same scale (millimetre). This ordination technique has been used previously with Galápagos finch populations (Grant & Schluter, 1984; Grant, 1986). It is useful because it reduces the number of traits (originally three: median beak length, depth and width) by taking only the first two component axes (PC1: beak size, PC2: beak shape) while accounting for the correlation between them (Schluter & Grant, 1984; Mitchell-Olds & Shaw, 1987; De León *et al.*, 2012; Legendre & Legendre, 2012; Martin, 2016). Concretely, I computed the principal component analysis on the three beak traits (length, width and depth) for the four species with a scaling 2 (correlation biplot), preserving the Mahalanobis distances among the objects in the matrix. This scaling takes into account the collinearity of the traits (Mahalanobis, 1936; Legendre & Legendre, 2012).

Statistical Model and inputs

To compute rates of recapture and individual survival probability, I used a Cormack-Jolly-Seber (CJS) capture recapture model (Cormack, 1964; Jolly, 1965; Seber, 1965; Plummer, 2003a; Kéry

& Schaub, 2012). The CJS model, which forms the base of our analyses, uses capture histories to simultaneously estimate recapture probabilities and survival probabilities. I assumed that the population size of the four species could vary through processes such as birth, and death, which is in line with open-population model assumptions (Amstrup *et al.*, 2005; MacKenzie *et al.*, 2005; Kéry & Schaub, 2012). Also, migration is negligible since our data collection was conducted between January and April, it should be less affected by the movement of birds that usually disperse after the breeding season in May (Schluter, 1984), but see Hendry *et al.* (2009).

These models also require that species identification be without error, if I am to include species-level effects. This presents a challenge when studying species that hybridize as is the case for these finches (Salvin, 1876; Lack, 1947; Grant, 1986). I use a linear discriminant analysis to assess our accuracy with bird identification (Figure S1; Legendre & Legendre, 2012). In the next sections, I will explain how the recapture model is mathematically defined and discuss differences between using a quadratic model and a thin-plate spline model to calculate selective gradients as well as selection surfaces for the former and to construct topological selection surfaces, for the latter.

Modeling the unobserved (state-space) and observed rates of survival

Estimating rates of survival in unobserved birds — In the conceptual framework of the CJS, capture histories can be arranged in an $N \times T$ array, where N is the total number of unique individuals and T is the number of years over which sampling occurred. Because some birds may evade capture, this capture history array does not necessarily reflect the actual state (dead or alive) of all birds in the community. To deal with these false-absences, inferences about an individual's survival, denoted as $\Phi[i,t]$ for individual i in year t , is statistically separated from

that individual's recapture history ($p[sp_{[i]},t]$ denotes the recapture probability of individual i of species $sp_{[i]}$ in year t). A matrix $z[i,t]$ denotes the latent dead/alive state for birds in year t . I consider each bird from the time it was first captured and all subsequent sampling events (because the CJS considers only captures after the individual was first marked). In the first year that a bird is captured, the latent state $z[i,t_{first}]=1$. An individual's dead/alive state in future years ($t > t_{first}$) is then given by

$$z[i,t] \sim \text{Bern}(z[i,t-1]\Phi[i,t-1])$$

where, *Bern* is a bernoulli distribution, $\Phi[i,t-1]$ is the probability that individual i survived from year $t-1$ to year t . The matrix z , representing the dead or alive state of a bird in a year, will have a length of N rows and T columns.

Estimating rates of survival in observation process. — The CJS model uses the raw capture history to estimate rates of recapture. I let $p[sp_{[i]},t]$ denote the recapture probability of an individual of species ($sp_{[i]}$) in year t . Thus, the observation process is conditional on the individual being alive ($z[i,t]=1$) and, consequently, the individual having survived from year $t-1$. I let $X[i,t]$ denote the observation matrix which I assume is Bernoulli distributed based on the state process and the species recapture probability in year t :

$$X[i,t] \sim \text{Bern}(z[i,t]p[sp_{[i]},t])$$

The matrix X is also called the observed capture history. It represents the presence and absence of the birds in the community across years which can be determined by a Bernoulli distribution because it describes a binary outcome (0 or 1) from a probability (in this case $z[i,t]p[sp_{[i]},t]$).

Thus, the X matrix is used to infer information about the state process (z), the recapture probabilities of each species (p), and the survival of the individuals (Φ).

Linear and quadratic estimation of survival and recapture probabilities

The quadratic model is a deterministic model particularly well suited to compute selective gradients (Lande & Arnold, 1983). Unfortunately, it is less suited to capture the topology of the fitness surface compared to a spline since it is less flexible because of the lack of a smoothing function. But it can still give an idea of the shape of fitness function. Below, I describe the mathematical formulation of the quadratic model. In order to facilitate comparison of effect sizes across variables, I standardized all continuous variables, by subtracting the mean and dividing by the standard deviation for each data point $Z[i,\sigma]=(Z[i]-Z_{mean})/\sigma$, where Z is the variable to be standardized.

Survival probabilities — Survival (Φ) and recapture (p) probabilities are estimated in the model. The phenotypic measurements and climatic data were used to inform individual survival probabilities. Specifically, I define the probability that individual i survives from year t to year $t+1$ as:

$$\begin{aligned} \text{logit}(\Phi[i,t]) = & \Phi.sp[sp_{[i]}] \\ & + \Phi.year[t] \\ & + \Phi.rain \times rain[t] \\ & + \Phi.PC1[sp_{[i]}] \times pc1[i] \\ & + \Phi.PC2[sp_{[i]}] \times pc2[i] \\ & + \Phi.PC1.PC2[sp_{[i]}] \times pc1[i] \times pc2[i] \\ & + \Phi.PC1.2[sp_{[i]}] \times pc1[i]^2 \\ & + \Phi.PC2.2[sp_{[i]}] \times pc2[i]^2 \\ & + \Phi.PC1.rain[sp_{[i]}] \times pc1[i] \times rain[t] \\ & + \Phi.PC2.rain[sp_{[i]}] \times pc2[i] \times rain[t] \end{aligned}$$

I use a logit link function ($\text{logit}(\Phi[i,t])$) to ensure that rates of survival remain between 0 and 1. A species-specific intercept $\Phi.sp[sp_{[i]}]$ accounts for the differences between the species. The principal component scores of individual i of species sp is included as both a linear, quadratic, and interaction effect ($\Phi.PC1[sp_{[i]}]$ quantifies the linear effect of PC1 on survival and $\Phi.PC1.2[sp_{[i]}]$ the quadratic effect; $\Phi.PC2[sp_{[i]}]$ quantifies the linear effect of PC2 on survival and $\Phi.PC2.2[sp_{[i]}]$ the quadratic effect; $\Phi.PC1.PC2[sp_{[i]}]$ quantifies the interaction effect between PC1 and PC2). The linear effect is added to account for directional selection. The quadratic term models a curvilinear regression, which can account for divergent selection. An interaction term accounts for correlational selection and allows the fitness landscape to be oriented in different directions in the phenotypic space (Lande & Arnold, 1983; Arnold *et al.*, 2001). I included a fixed effect of precipitation $\Phi.rain$ to account for the seasonality (wet vs dry). The sum of the rainfall is used in the model with the $rain[t]$ parameter. I also include a random effect of year ($\Phi.year[t]$) in order to capture other year to year variation in survival.

Capture probabilities — To compute the year-specific recapture probabilities I used a logit link function to ensure that capture probabilities remain between 0 and 1. Specifically,

$$\text{logit}(p[i,t]) = p_0[sp_{[i]}] + p_1 \text{effort}[t]$$

Here, $p_0[sp_{[i]}]$ is a species-specific intercept and p_1 is a fixed effect of sample effort ($\text{effort}[t]$) calculated in hours of net use in year t .

Spline estimation of survival and recapture probabilities

The thin-plate spline, a nonparametric procedure, is used here to infer the topology of the fitness landscapes. It is a collection of polynomials connected together with a smoothing factor controlling the fit of the function (Schluter & Nychka, 1994).

Survival probabilities — I also estimated survival (Φ) probabilities using splines (Gimenez *et al.*, 2009).

$$\begin{aligned} \text{logit}(\Phi[i,t]) = & \Phi.\text{year}[t] \\ & + \Phi.\text{rain} \times \text{rain}[t] \\ & + \Phi.\text{PC1}[i] \\ & + \Phi.\text{PC2}[i] \\ & + S[i,k]u[k] \end{aligned}$$

where u is a vector of random effects. In our model, the individual variation is captured in the S matrix (also called a design matrix), where the rows correspond to the individuals and the columns correspond to connection points between the polynomials. The number of joints between the polynomials was controlled with $k = \max(20, \min(N/4, 150))$, where k is the number of joints and N is the sample size (Ruppert *et al.*, 2003). The S matrix is computed using a Swapping Algorithm with the space-filling coverage designs from the fields package in R (Nychka *et al.*, 2015; R Core Team, 2016; Fields Development Team, 2006). Four models are explored: one including only PC1 effect with all species, and a second including only the effect of PC2 with all species. The last two models are run independently for each species and trait combination and then combined on the same fitness landscape to remove any effect of species on the estimation of the different populations of finches.

Definition of model priors

Quadratic model — All of the climatic and phenotypic parameters are fixed effects in the model with a normal prior distribution $\mathcal{N}(\mu = 0, \tau = 0.01)$, where τ is the precision $\tau = 1/\sigma^2$. The random parameters were distributed with $\mathcal{N}(\mu, \tau[i])$ where $\mu = \mathcal{N}(0, 0.01)$ and the variance (the σ^2 in the equation $\tau = 1/\sigma^2$) a wide uniform distribution $\mathcal{U}(a = 0, b = 10)$. Only the random effect of year on survival was distributed with a mean of 0, $\Phi.\text{year}[t] \sim \mathcal{N}(0, \tau[t])$.

Thin-plate spline model — I used the same priors as for the quadratic model, with the exception that the random effects of the spline parameters were distributed according to a Half-Cauchy distribution (Gelman *et al.*, 2013).

$$u[k] = \mathcal{N}(0, \tau[k]), \quad \tau[k] = G(0.5, 0.5)$$

G denotes the Half-Cauchy distribution and k is the number of joints between the polynomials of the spline. Our models were executed in JAGS using a Gibbs sampler (Geman & Geman, 1984; Plummer, 2003a). Markov Chain Monte Carlo (MCMC) simulations were set with a number of 105 000, 15 000, 3 000 iterations with 5000 of initial burn-in iterations respectively for the quadratic model, the thin-plate spline and the independent thin-plate splines. Using three independent chains, the sample rate (thinning) was 15, 5 and 5 (respectively for the quadratic model, the thin-plate spline and the independent thin-plate splines). All the analysis was performed in R, version 3.3.1 (Plummer, 2003b; Su & Yajima, 2015; R Core Team, 2016). Parameters whose estimated 95% credible intervals do not include zero were considered "significant". Model diagnostics are provided in the supplementary material from Figure S3 to Figure S10.

Fitness landscape

To characterize the topology (number of adaptive peaks and gaps) of multiple species fitness landscapes, I use three models: 1. A quadratic model, 2. Independent thin-plate spline for each species, and 3. A thin-plate spline without the species effect but including all the individuals of this study. I use survival probabilities Φ as a surrogate of fitness. Survival is only one component of fitness and might not represent individual's actual fitness (Orr, 2009). But it has been used extensively (Hendry *et al.*, 2009; Kéry & Schaub, 2012; Martin & Wainwright, 2013; Martin,

2016; Roches *et al.*, 2016). The quadratic model serves two purposes. First, the selective gradients (linear and quadratic) are estimated, and I implement an interaction term between climate and traits to analyze temporal variation in the fitness landscape. Second, the quadratic model can be used to interpret the change of topology in the different years based on the interaction term between climate and traits.

The two spline models are used to characterize the topology of multiple species fitness landscapes on the two phenotypic axis (PC1: beak size and PC2: beak shape). First, a spline will be fitted for each species independently and then combined in a common fitness landscape. This way, the fitness functions will be independently fitted through each species without trying to connect the fitness functions between the species. By “connecting”, I mean that the end points of the fitness functions, that is the extreme phenotypes of the population of finches, are joined between species, thus removing the gaps between the surface of the species. Second, another spline model will be used but all the species at the same time to estimate one general fitness function. This way, there will be a connection of the fitness surface between the species.

Results

Ground finch species show distinct beak morphology

Through our long-term study, our team collected a total of 2598 individuals (1224 *G. fortis*, 675 *G. fuliginosa*, 61 *G. magnirostris* and 164 *G. scandens* across all years). In two years, some populations were not captured in our mist nets despite the fact those finches were present on the island (in 2005, no *G. scandens* were captured and in 2008, only *G. fortis* was captured Table I).

To characterize the fitness landscapes, I first establish a phenotypic space. I used three beak traits (length, depth, and width) to describe the morphospace. The first principal component explains

89.65% of the variation (Figure 2). This component summarizes the variation for beak size with *G. fuliginosa* being the smallest and *G. magnirostris* the largest. There is a high correlation between median beak width and depth (using a linear model slope = 0.74; $r^2 = 0.95$; p-value = 0 on the raw phenotypic values). The second principal component explains 9.64% of the variation. This component summarizes the variation in beak shape from blunt to sharp. *G. scandens*, possesses a sharp beak relative to other species, which it uses to probe into cactus flowers.

Table I Number of birds captured and capture effort per year at El Garrapatero

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
<i>G. fortis</i>	35	109	181	127	57	94	97	188	162	170	203	143
<i>G. fuliginosa</i>	6	8	13	6	2	0	36	142	155	128	180	88
<i>G. magnirostris</i>	1	1	9	10	6	0	2	9	9	9	11	4
<i>G. scandens</i>	5	5	0	10	2	0	10	21	34	30	47	33
Capture effort (hours)	36	140	212	120	52	56	132	300	128	120	128	104

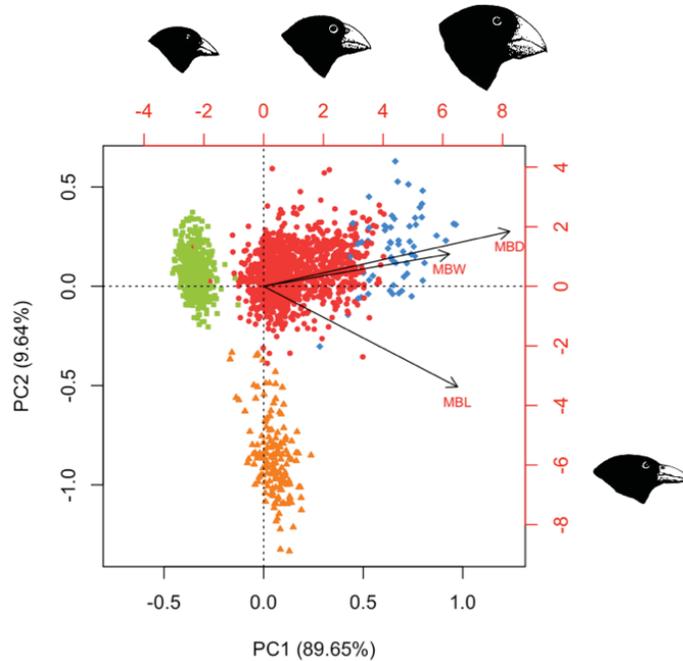


Figure 2 Correlation double projection of the principal component analysis of all birds including three beak traits; median beak length (MBL), width (MBW) and depth (MBD).

Green squares represent *G. fuliginosa*, red dots are *G. fortis*, blue diamonds are *G. magnirostris*, and orange triangles are *G. scandens*. PC1 explains 89.65% of the variation and can be interpreted as beak size, whereas PC2 explains 9.64% of the variance and can be interpreted as beak shape. The red axes (top and right) are scaled for the trait vectors while the black axes (bottom and left) are scaled for the points. Most individuals are clustered by species.

Species fitness landscapes are variable across species and years

We found no main effects of species phenotypes on their survival probabilities using the quadratic model. However, we did find evidence of a significant interaction between beak shape (PC2) and rainfall for *G. fortis* (-1.56, CI: [-3.1 ; -0.08]; Table II), but the 95% credible intervals for all other variables contain zero.

(i) Topography: numbers of peaks, positions, gradient elevation

There is one fitness function per species because there is a random intercept of species ($\Phi_{.sp[sp_{ij}]}$), and the species specific effects on the two traits (PC1 and PC2). First, all the species have different fitness functions in terms of shape and directions (Figure 3). In general, the four species have distinct adaptive peaks. Overall, *G. fuliginosa* and *G. scandens* have a higher survival probability than the two other species (mean of 0.69 and 0.70 respectively). *G. fuliginosa* shows evidence of positive selection for smaller beaks, but no effect of selection on beak shape. For all the other species, selection acts on both beak size and beak shape. In *G. fortis*, the survival values show two adaptive peaks for extreme values of beak size and beak shape. The fitness landscape for *G. magirostris* is represented by a bell curve in both PC1 and PC2 directions, with the curve steeper on the beak shape axis. *G. scandens* have an adaptive peak on the PC1 axis, but this peak changes shape depending on the year switching between directional selection and stabilizing selection.

(ii) Effect of rainfall on the topology of the quadratic model

Climate does not affect survival probability strongly overall with the quadratic model (slope -0.10, 95% credible intervals overlapping 0, [-0.99; 0.66]). Nevertheless, there is a relevant interaction term for climate on the beak shape axis for *G. fortis* (-1.56, CI: [-3.1; -0.08]).

Four years are categorized as wet years (2008, and 2010-2012, (mean \pm standard deviation) 602.20 ± 69.08 mm; Figure 8) and all the other years are considered as dry years (208.57 ± 67.98 mm). The year 2013 is exceptionally dry with only 76.3 mm of rain. There is a change in the fitness landscape based on climatic condition (wet vs dry) specifically for *G. fortis*. For some species, the direction of selection is switched from one year to another, but not statistically different. Stabilizing selection seems to be more pronounced in dry years for *G. scandens* (Figure 3). But in wet years, there is directional selection towards larger beaks. In *G. magnirostris*, stabilizing is the norm in dry years, but there is directional selection for blunter beaks in a dry environment. *G. fortis* always shows 2 peaks of different heights; in wet years smaller and blunter beaks are advantageous whereas in dry years, larger pointier beaks are favoured. *G. fuliginosa* is consistently selected for smaller beaks but pointier vs blunter beaks are favoured following wet and dry years respectively.

Fitness landscapes are more variable with a species specific effect

As it was the case for the quadratic model, only the effect of year on survival probabilities for 2003 and 2011 have credible intervals not overlapping zero in the thin-plate spline model (Table III). The recapture probabilities for all the species are 27.49%. The fitness landscapes from the thin-plate splines are shown in Figure 4 to Figure 7. The results are shown with different models: 1. fitness functions that were independently drawn for each species and plotted in the same graph for each trait (PC1 Figure 4; PC2 Figure 5), and 2. one fitness function for all the species per trait (PC1 Figure 6; PC2 Figure 7). Since I am using a mark-recapture model, the fitness landscapes were drawn using the individuals that were inferred to be in the population (state process). In the first year (2003), there are fewer individuals whereas in the last year (2014), all the individuals

are drawn in the fitness landscape. Also, the model is not built to show different shapes for the fitness functions, but only different intercepts (or height in the survival values).

(i) Topography of a thin-plate spline model with a shared fitness function

A thin-plate spline was used to draw the fitness landscapes, but considering only one trait at the time (PC1 Figure 6; PC2 Figure 7). For beak size (PC1), three peaks are observed: one for *G. fuliginosa* which tends to have its maximum towards smaller beaks, one that shows stabilizing selection for a small morph of *G. fortis* and *G. scandens*, and one that corresponds to stabilizing selection for a large morph of *G. fortis* and *G. magnirostris*. The fitness function of the beak shape axis (PC2) shows a higher survival probability towards pointier beaks across all species. Overall, *G. scandens* shows the highest survival rates across all species.

(ii) Topography of a thin-plate spline model for each species

When the fitness landscapes are computed independently for each species, more flexible landscapes are possible for both beak size (Figure 4) and beak shape (Figure 5). Because the sample size varies between the species, the number of knots varied for each species. There were 50, 50, 41, and 20 knots for *G. fortis*, *G. fuliginosa*, *G. scandens*, and *G. magnirostris* respectively.

First, for beak size (PC1), the survival of *G. fortis* shows a bimodal landscape with high survival probability for the small morphs and large morphs and low survival for intermediate sizes.

Overall, there is directional selection for smaller beaks in *G. fuliginosa*. *G. scandens* shows a higher survival probability towards bigger beaks. Interestingly, the survival of the cactus finch depicts only the left side of a chopped bell curve. *G. magnirostris* has a slightly curved fitness function where smaller beaks for this species are advantageous. Generally, the overall survival

probability decreases through the years for all species. However, for *G. magnirostris*, there is a sharp decrease in survival probability from 2010 until the end of the study. It should be noted that the credible intervals around survival probabilities are very large and as such the single species fitness landscapes do not show statistically significant variation across years.

Second, for beak shape axis (PC2), *G. fortis* appears to have a small bimodal fitness landscape which favours blunter and shaper beaks. *G. fuliginosa* has two peaks; one showing a linear gradient favouring pointier beaks and one showing a bell-curve function towards blunter beaks. *G. scandens* has a general adaptive peak where mean morphotype has a higher probability to survive. But this peak is subdivided in two showing a smaller peak favouring blunter beaks and one favouring pointier beaks. *G. magnirostris* has a bimodality for beak shape where blunter beaks and pointier beaks have approximately the same survival probabilities. As for PC1, there is a general tendency of reduced survival probabilities throughout the years for all the species. But *G. magnirostris* has a reduced survival probability starting from 2010 onward.

Table II Selection gradients and parameters estimated from the quadratic model.

\hat{R} is the scale reduction factor, a measure of the convergence of the model; a smaller value means a better convergence. All values are on a logit scale. The effective sampling size is the number of independent samples from the posterior distribution. The model was set with 105 000 iterations with a burn-in of 5000 and a thinning of 15 on three independent chains. Numbers in bold mean a significant result.

Parameters	Species / year	Posterior Mean [†]	95 % Credible interval	\hat{R}	Effective sampling size
Recapture probabilities	<i>G. fortis</i>	-0.82	[-1.05 ; -0.59]	1.00	20000
	<i>G. fuliginosa</i>	-1.87	[-2.38 ; -1.35]	1.00	4300
	<i>G. magnirostris</i>	-2.15	[-3.57 ; -0.99]	1.02	150
	<i>G. scandens</i>	-1.40	[-2.13 ; -0.7]	1.00	2700
Sampling effort	-	0.21	[0.01 ; 0.4]	1.00	3800
Linear selection gradient PC1	<i>G. fortis</i>	-0.75	[-3.04 ; 1.55]	1.00	20000
	<i>G. fuliginosa</i>	-3.63	[-15.24 ; 6.56]	1.02	160
	<i>G. magnirostris</i>	3.20	[-8.68 ; 16.63]	1.01	690
	<i>G. scandens</i>	2.67	[-11.35 ; 16.93]	1.00	1700
Linear selection gradient PC2	<i>G. fortis</i>	0.53	[-1.08 ; 2.14]	1.00	20000
	<i>G. fuliginosa</i>	-3.77	[-11.08 ; 3.44]	1.00	1900
	<i>G. magnirostris</i>	0.27	[-15.52 ; 15.74]	1.01	470
	<i>G. scandens</i>	-3.15	[-13.15 ; 7]	1.04	1400
Quadratic selection gradient PC1 ²	<i>G. fortis</i>	1.12	[-3.89 ; 6.07]	1.00	20000
	<i>G. fuliginosa</i>	5.46	[-11.16 ; 23.16]	1.01	180
	<i>G. magnirostris</i>	0.67	[-12.18 ; 16.42]	1.01	400
	<i>G. scandens</i>	-1.55	[-20.8 ; 17.87]	1.00	20000
Quadratic selection gradient PC2 ²	<i>G. fortis</i>	1.53	[-3.99 ; 7.22]	1.00	4400
	<i>G. fuliginosa</i>	8.91	[-6.56 ; 24.74]	1.00	16000
	<i>G. magnirostris</i>	-2.50	[-20.49 ; 15.94]	1.00	1800
	<i>G. scandens</i>	-1.56	[-8.69 ; 7.09]	1.01	370
Interaction PC1xPC2 (correlational selection)	<i>G. fortis</i>	-2.54	[-8.99 ; 3.79]	1.00	11000
	<i>G. fuliginosa</i>	-2.27	[-20.58 ; 16.06]	1.00	4200
	<i>G. magnirostris</i>	0.77	[-16.66 ; 18]	1.00	1900
	<i>G. scandens</i>	-1.12	[-17 ; 14.36]	1.00	1100
Interaction PC1xclimate	<i>G. fortis</i>	0.19	[-1.01 ; 1.41]	1.00	20000
	<i>G. fuliginosa</i>	0.65	[-1.3 ; 3.82]	1.02	1100
	<i>G. magnirostris</i>	-3.30	[-10.27 ; 9.22]	1.03	130
	<i>G. scandens</i>	5.12	[-8.58 ; 19.39]	1.00	2900
Interaction PC2xclimate	<i>G. fortis</i>	-1.56	[-3.1 ; -0.08]	1.00	5700
	<i>G. fuliginosa</i>	-3.06	[-7.86 ; 1.81]	1.00	4800
	<i>G. magnirostris</i>	7.76	[-10.44 ; 23.9]	1.01	320
	<i>G. scandens</i>	0.33	[-1.15 ; 3.27]	1.03	730
Climate (precipitation)	-	-0.10	[-0.99 ; 0.66]	1.00	5000
Effect of species on survival	<i>G. fortis</i>	0.33	[-1.15 ; 3.27]	1.03	730
	<i>G. fuliginosa</i>	0.28	[-0.49 ; 1.15]	1.00	940
	<i>G. magnirostris</i>	-0.70	[-3.53 ; 1.52]	1.02	140

Effect of year on survival	<i>G. scandens</i>	0.07	[-4.49 ; 5.4]	1.01	1000
	2003	1.52	[0.09 ; 3.74]	1.00	1800
	2004	0.29	[-1.04 ; 1.58]	1.00	1400
	2005	-0.18	[-1.35 ; 0.84]	1.00	1300
	2006	-0.44	[-1.62 ; 0.54]	1.00	1700
	2007	0.91	[-0.33 ; 2.65]	1.00	2200
	2008	0.69	[-0.8 ; 2.45]	1.00	2300
	2009	0.16	[-0.93 ; 1.39]	1.00	850
	2010	-0.06	[-1.4 ; 1.4]	1.00	1100
	2011	-0.53	[-1.84 ; 0.91]	1.00	1600
	2012	-0.69	[-1.77 ; 0.46]	1.01	870
	2013	-1.53	[-3.1 ; -0.29]	1.00	1600
	2014	0.01	[-2.47 ; 2.56]	1.00	20000

† The posterior mean is on the logit scale [$f(x) = \log(x/(1-x))$]. It can be transformed on the inverse logit with: $f^l(x) = 1/(1+e^{(-x)})$.

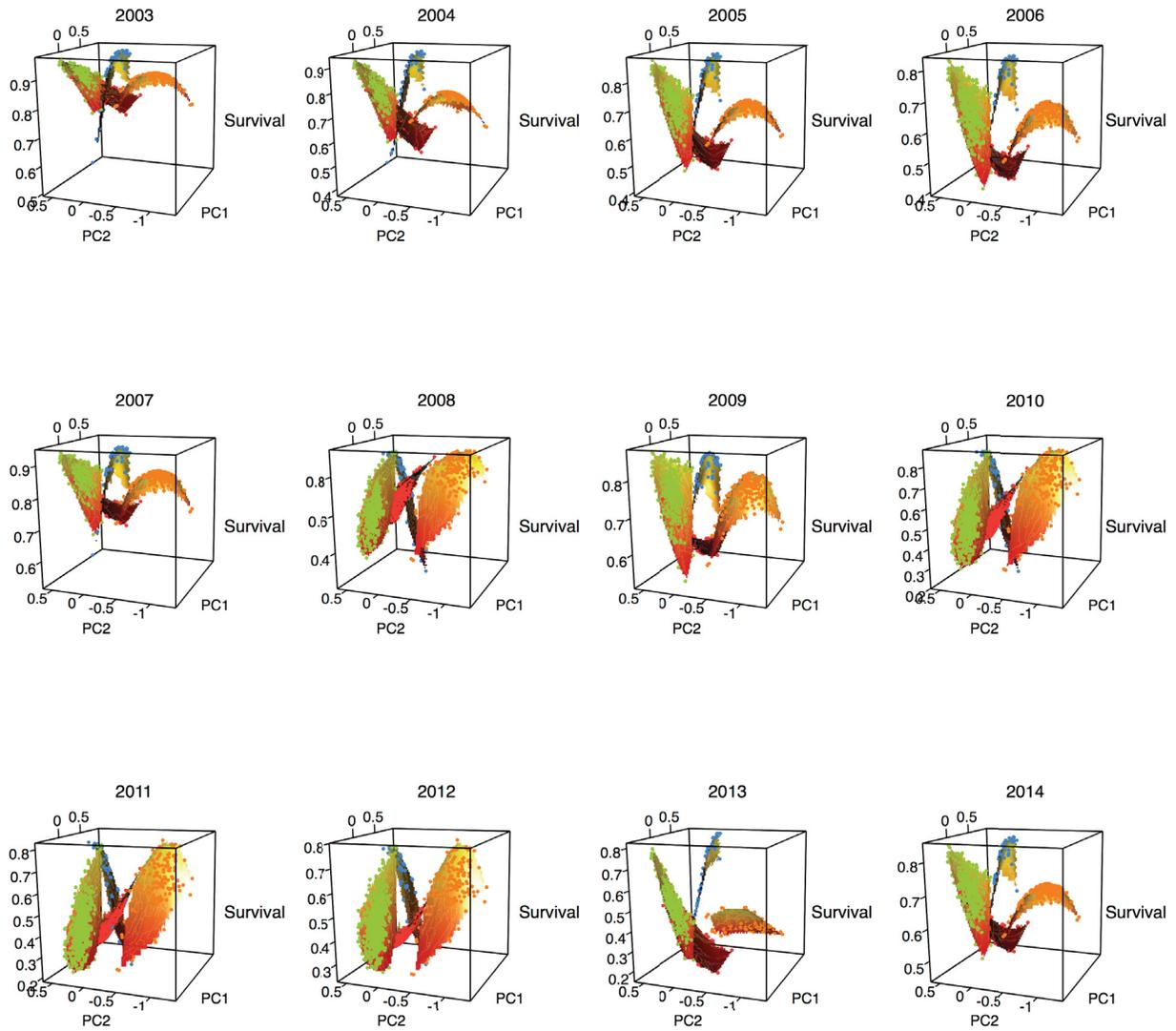


Figure 3 Fitness landscape for ground finches for beak size (PC1 scores) and beak shape (PC2

scores) estimated with a quadratic model including linear and quadratic terms.

Survival is the surrogate of fitness. The model is estimating survival per year for each species independently. Green dots: *G. fuliginosa*, red dots: *G. fortis*, blue dots: *G. magnirostris*, orange dots: *G. scandens*.

Table III Parameters estimated from the thin-plate spline model.

\hat{R} is the scale reduction factor, a measure of the convergence of the model; a smaller value means a better convergence. All values are on a logit scale. The effective sampling size is the number of independent samples from the posterior distribution. The model was set with 15 000 iterations with a burn-in of 5000 and a thinning of 10 on three independent chains. Numbers in bold mean a significant result.

Parameters	Species / year	Posterior Mean [†]	95 % Credible interval	\hat{R}	Effective sampling size
Recapture probabilities	<i>G. fortis</i>	-0.98	[-1.17 ; -0.78]	1.02	140
	<i>G. fuliginosa</i>	-0.98	[-1.17 ; -0.78]	1.02	140
	<i>G. magnirostris</i>	-0.98	[-1.17 ; -0.78]	1.02	140
	<i>G. scandens</i>	-0.98	[-1.17 ; -0.78]	1.02	140
Sampling effort	-	0.10	[-0.09 ; 0.29]	1.00	6000
Intercept	-	0.38	[-0.44 ; 1.28]	1.03	99
Individual effect of PC1	-	-0.05	[-0.67 ; 0.5]	1.00	3800
Individual effect of PC2	-	-0.52	[-1.15 ; 0.14]	1.01	3900
Climate (precipitation)	-	-0.23	[-1.23 ; 0.51]	1.07	37
Effect of year on survival	2003	1.61	[0.09 ; 3.97]	1.00	6000
	2004	0.61	[-0.9 ; 2.2]	1.01	720
	2005	-0.36	[-1.56 ; 0.74]	1.01	870
	2006	-0.59	[-1.76 ; 0.41]	1.01	820
	2007	0.92	[-0.43 ; 2.67]	1.01	260
	2008	0.73	[-0.84 ; 2.72]	1.07	37
	2009	0.56	[-0.65 ; 2.12]	1.00	1500
	2010	-0.33	[-1.65 ; 1.09]	1.08	31
	2011	-0.62	[-2 ; 0.83]	1.08	32
	2012	-0.70	[-1.78 ; 0.39]	1.07	37
	2013	-1.73	[-3.34 ; -0.43]	1.02	130
2014	0.01	[-2.7 ; 2.61]	1.00	6000	

[†] The posterior mean is on the logit scale [$f(x) = \log(x/(1-x))$]. It can be transformed on the inverse logit with: $f^{-1}(x) = 1/(1+e^{-x})$.

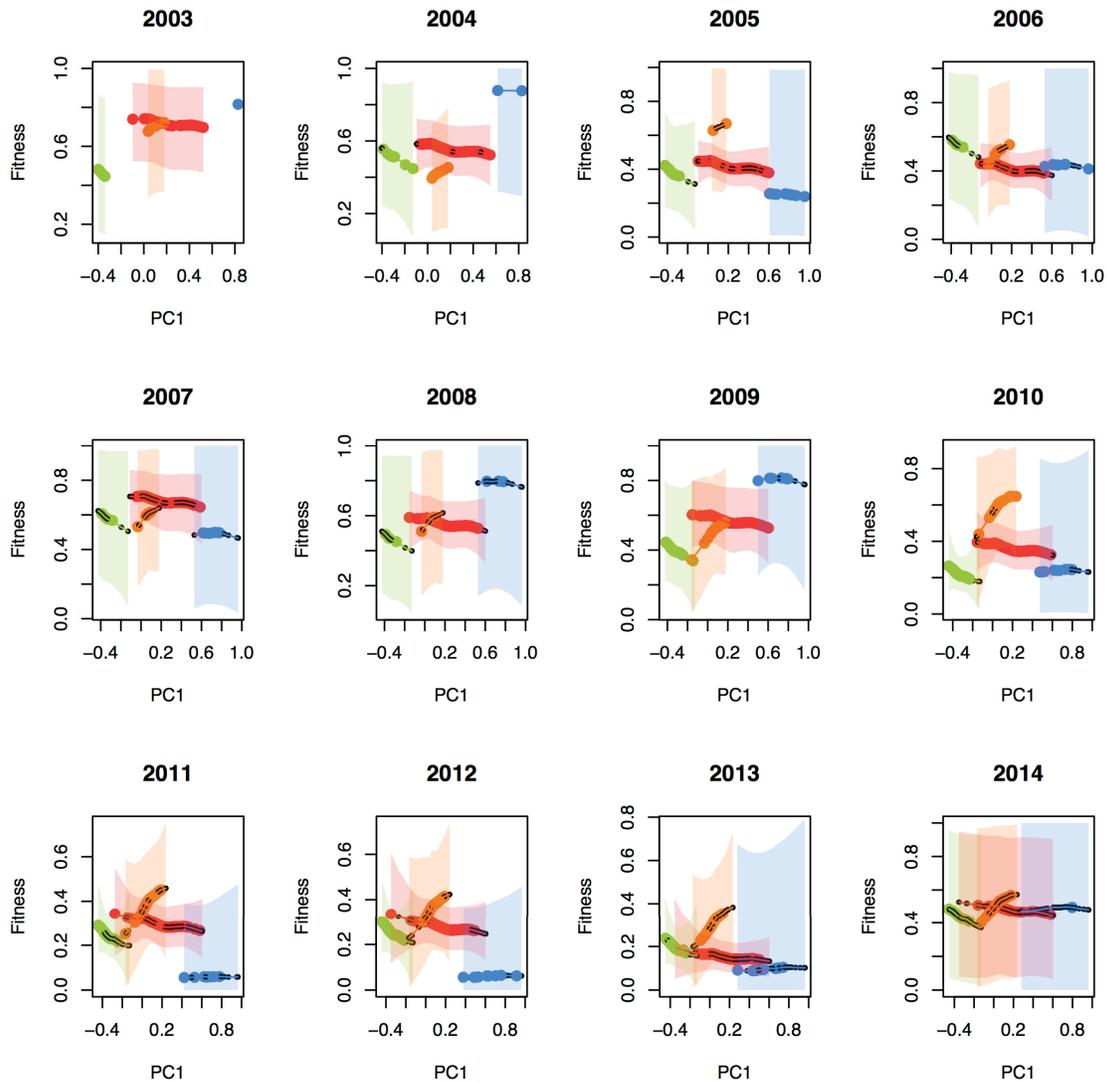


Figure 4 Fitness landscape per year for ground finches with 95% credible intervals for beak size (the model included only PC1 scores) estimated with a thin-plate spline.

Survival is the surrogate of fitness. The model is estimating survival per year for each species independently. A large dot suggests a high probability of an organism to be present in a year whereas smaller black dots represent a low probability of survival for an individual. At the beginning, fewer individuals can have an estimate of their survival, but at the end, all individuals have a survival estimate. Green dots: *G. fuliginosa*, red dots: *G. fortis*, blue dots: *G. magnirostris*, orange dots: *G. scandens*.

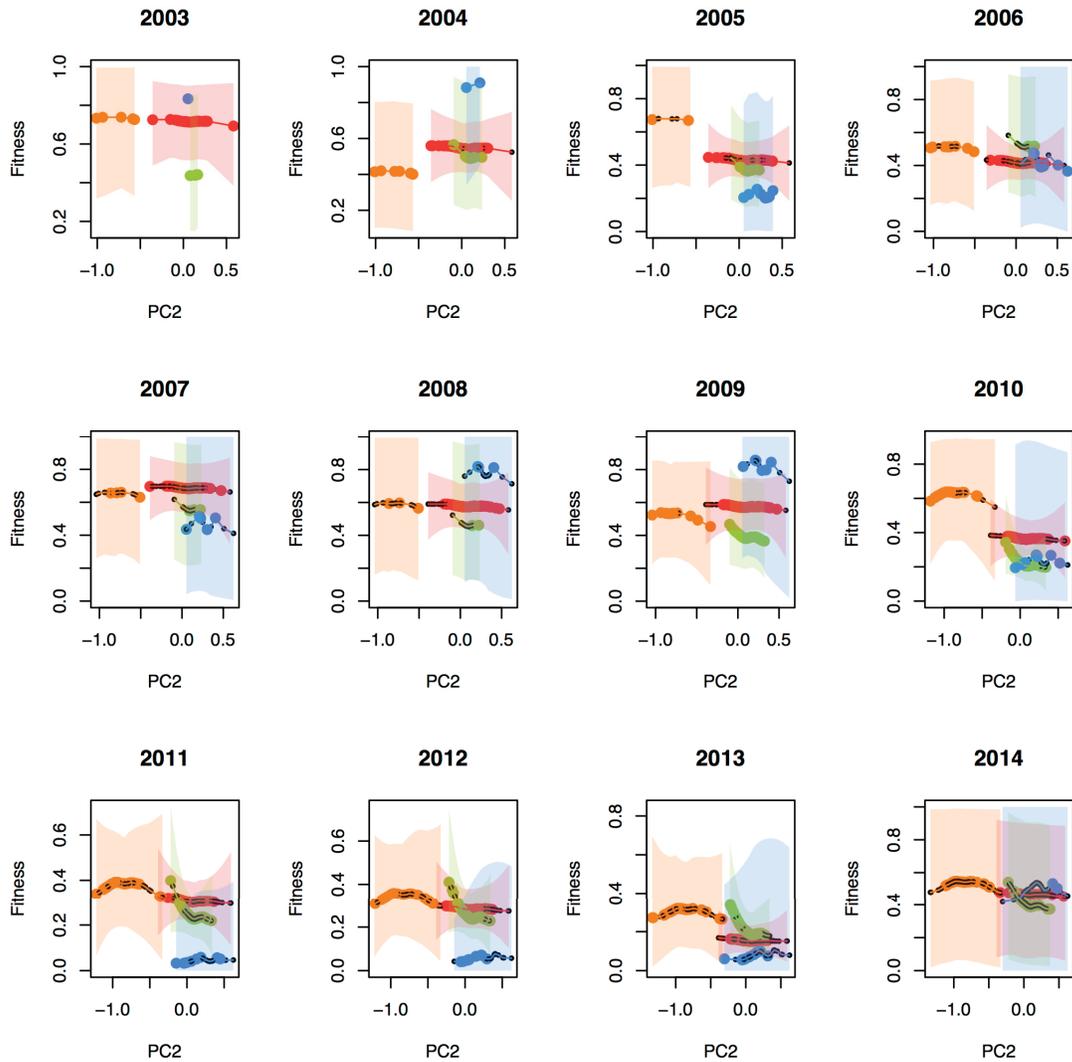


Figure 5 Fitness landscape per year for ground finches with 95% credible intervals for beak shape (the model included only PC2 scores) estimated with a thin-plate spline.

Survival is the surrogate of fitness. The model is estimating survival per year for each species independently. A large dot suggests a high probability of an organism to be present in a year whereas smaller black dots represent a low probability of survival for an individual. At the beginning, fewer individuals can have an estimate of their survival, but at the end, all individuals have a survival estimate. Green dots: *G. fuliginosa*, red dots: *G. fortis*, blue dots: *G. magnirostris*, orange dots: *G. scandens*.

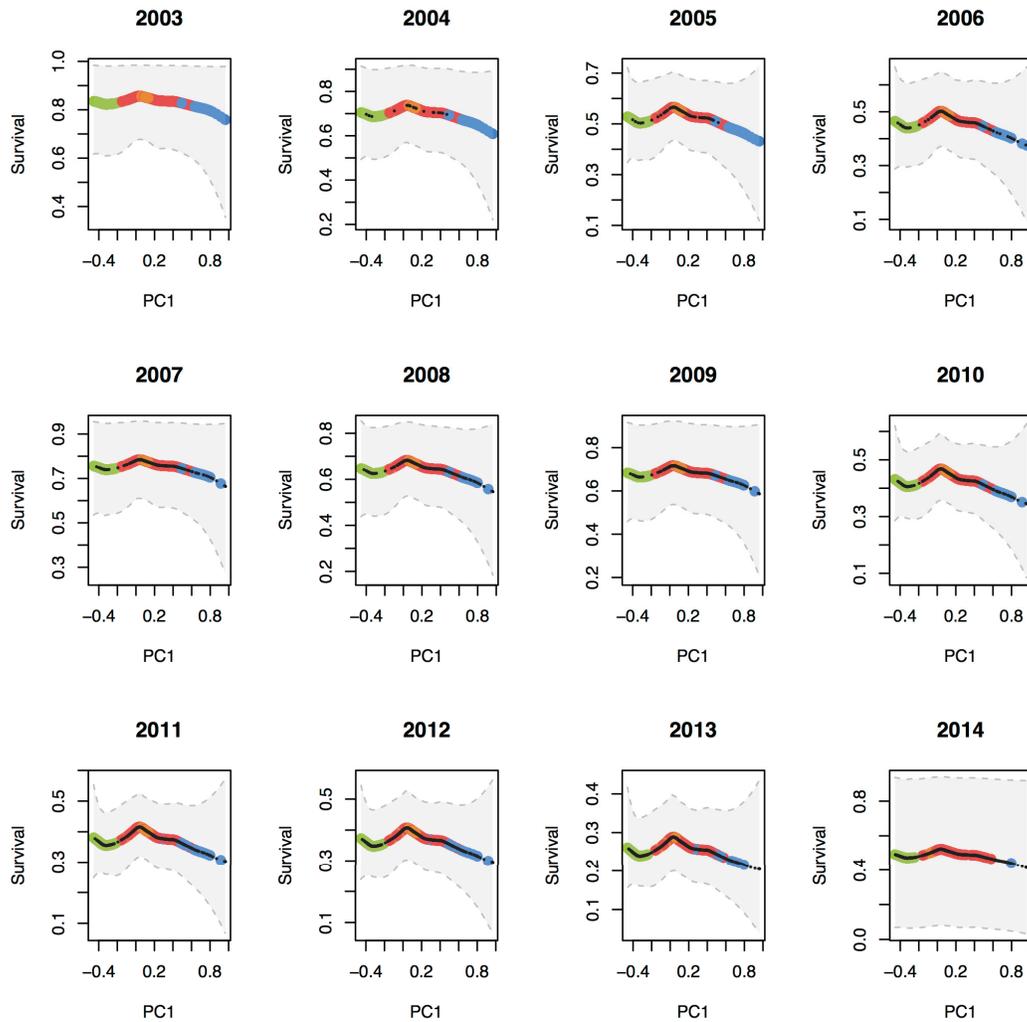


Figure 6 Fitness landscape per year for ground finches with 95% credible intervals for beak size (the model included only PC1 scores) estimated with a thin-plate spline.

Survival is the surrogate of fitness. The model is estimating survival per year considering only one species with PC1 as the only trait. A large dot suggests a high probability of an organism to be present in a year whereas smaller black dots represent a low probability of survival for an individual. At the beginning, fewer individuals can have an estimate of their survival, but at the end, all individuals have a survival estimate. Green dots: *G. fuliginosa*, red dots: *G. fortis*, blue dots: *G. magnirostris*, orange dots: *G. scandens*

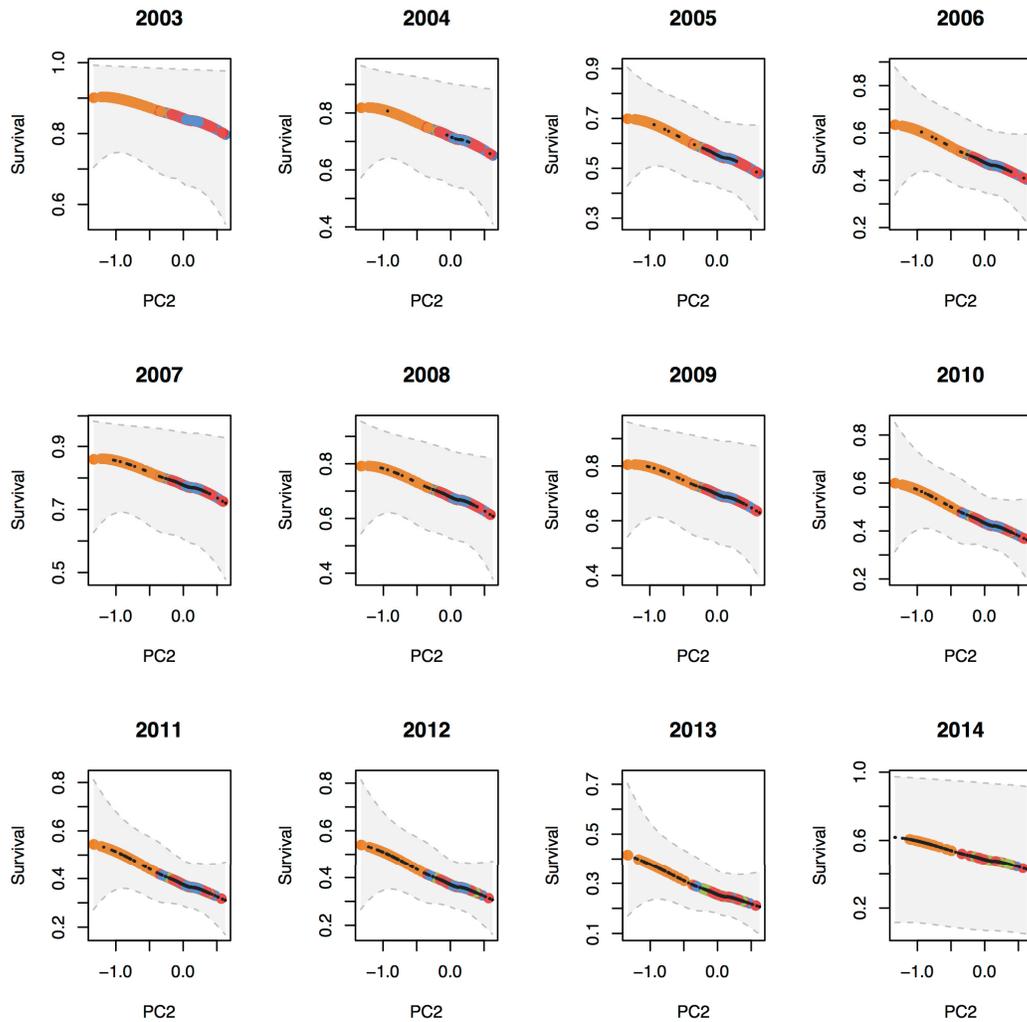


Figure 7 Fitness landscape per year for ground finches with 95% credible intervals for beak shape (the model included only PC2 scores) estimated with a thin-plate spline.

Survival is the surrogate of fitness. The model is estimating survival per year considering only one species with PC2 as the only trait. A large dot suggests a high probability of an organism to be present in a year whereas smaller black dots represent a low probability of survival for an individual. At the beginning, fewer individuals can have an estimate of their survival, but at the end, all individuals have a survival estimate. Green dots: *G. fuliginosa*, red dots: *G. fortis*, blue dots: *G. magnirostris*, orange dots: *G. scandens*.

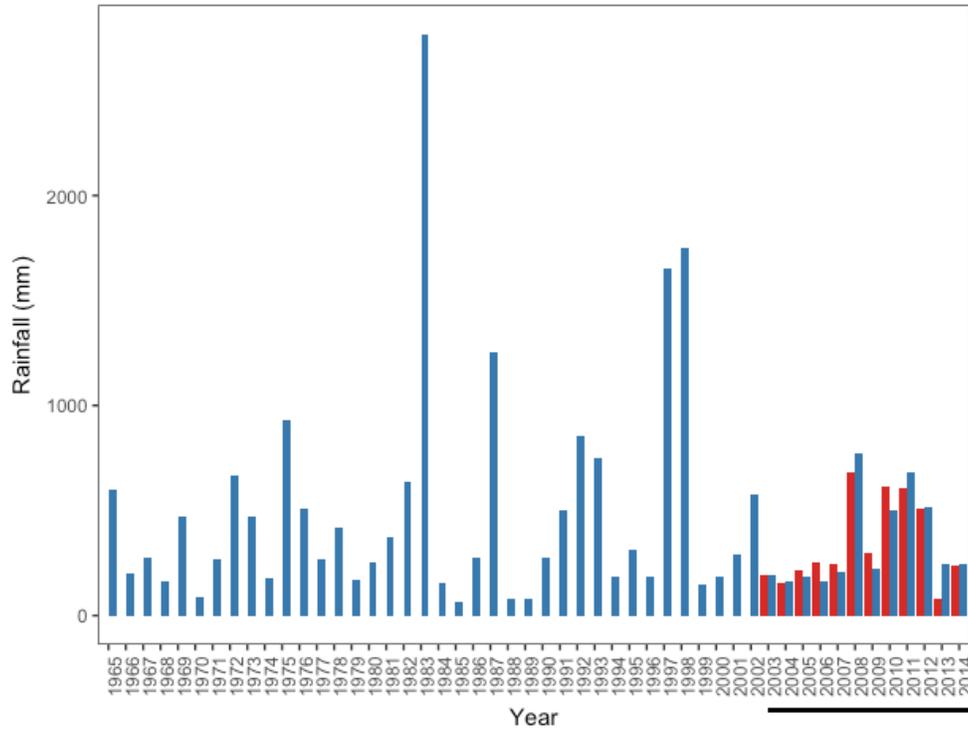


Figure 8 Climatic data from Puerto Ayora, a town 10 km from El Garrapatero, recorded by Charles Darwin Research Station (CDRS).

The y axis corresponds to the cumulative sum of rainfall per year between 1965 and 2014 (blue bars). The mark recapture data is from 2003 to 2014 (shown by the black line and red bars).

Discussion

In order to understand ecological speciation, it is necessary to link phenotypes and the environment with fitness (Nosil, 2012). Here, I investigate these links by using a mark-recapture model to quantify survival probabilities and determine the topology of fitness surfaces in multiple sympatric species of Galápagos finches. I first discuss the differences observed in the shapes of the fitness surfaces from the different models used. I then talk about temporal variation in the fitness landscapes. Finally, I discuss the expansion of morphospace observed in *G. fortis*.

Topologies differ on single species compared to multispecies fitness landscapes

My results show that the topology of fitness landscapes is strongly impacted by whether fitness functions include an independent species-specific effect. As hypothesized, the multi-species landscapes constructed from a thin-plate spline model without an effect of species yield less complex topologies than the independent fitness landscapes constructed for each species. This is because removing any species effect assumes that identical phenotypes between species have the same survival values. Consequently, there is only one fitness surface for all the species. The independent fitness landscapes were variable in their survival probabilities and thereby captured subtle patterns in the topologies of their fitness surfaces. This variation could reflect biologically relevant differences such as niche specialization, and thus help to understand how divergent selection may lead to ecological speciation. However, by using distinct fitness functions for each species, these landscapes do not account for the influence of distinct species on each other's fitness. Moreover, hybrid and backcrossed individuals must be assigned to a single species and its fitness landscape, which may result in misleading inferences about the topology of the landscape if these individuals show distinct fitness functions from "pure" species individuals. Recent analyses have reaffirmed that hybridization might be high between these species (Grant

and Grant 2008, 2010, Chaves et al. 2016), and hybrids have been shown to possess high fitness when dry events lead to food limitations (Grant and Grant 1996).

(i) Splines with independent fitness surfaces characterize more variable landscapes

I did not find evidence to support my hypothesis that fitness peaks will occur in regions of phenotypic space occupied by a higher density of individuals (i.e., higher fitness among birds with common phenotypes). The spline analysis of independent fitness surfaces showed more variable survival probabilities, with ridges or planes that do not necessarily correspond to phenotypic space occupied by high numbers of individuals, as opposed to simple peaks around mean phenotypes. This suggests that rare morphotypes do not necessarily have lower fitness, and shows that these species are not experiencing stabilizing selection. This is particularly noticeable for two species; in the small (*G. fuliginosa*) and cactus (*G. scandens*) ground finches' survival probabilities indicate directional selection towards smaller beaks and larger beaks respectively. This could be due to competition between the species because there is an overlap in their diet (Smith *et al.*, 1978). However, beak shape shows a distinct pattern with directional selection for pointier beaks in *G. fuliginosa* and stabilizing selection for pointy beaks in *G. scandens*. This may suggest that these species are able to partition resources that require pointy beaks based on differences in beak size alone.

Consistent with Hendry *et al.* (2009), the survival probabilities of *G. fortis* indicates a bimodality in beak size that favours both large and small morphotypes of this species, but with small morphs consistently having the highest survival. Interestingly, the large ground finch (*G. magnirostris*) shows the opposite pattern, with small beaked individuals having a greater survival probability than those with larger beak size. Unfortunately, the representation of independent fitness landscapes for each species does not provide a way to connect the fitness functions between

G. fortis and *G. magnirostris*, which limits inferences about potential hybridization of the two species. Connecting the fitness functions for these species could reveal that birds with intermediate phenotypes between two species (which are likely to be hybrids) show a lower survival probability (a valley), thus indicating hybrid inviability. On the other hand, hybrids could have a higher fitness at the junction (a peak) of the fitness function as described between *G. fuliginosa* and *G. scandens* (Grant & Grant, 1996).

The large ground finch also shows two clear survival peaks for pointy and blunt beaks. This pattern could potentially be explained by sex differences because it is known that large ground finches have a slight sexual dimorphism in size, with females tending to have pointier beaks (Price, 1984). While female *G. magnirostris* possessed pointier beaks in this dataset, due to the small sample size of this species, and a small effect size, differences between males and females were not significant (Figure S11).

The independent fitness functions are well suited to interpret the fitness (or survival) consequences of trait variation in each species. These functions depict more variable survival probabilities because the parameter estimating it are not averaged out across species. Thus, a clear interpretation of the type of selection (directional, stabilizing or disruptive) acting on each species can be made. The major problem with computing independent fitness surfaces for each species is the impossibility of discussing the connection between the distinct fitness functions. For example, testing whether extreme phenotypes of different species should have similar survival probabilities, or if hybrids between species show higher or lower survival rates. Thus, at present, this approach seems better suited to understanding the independent selection regimes that can drive initial population differentiation within each species.

(ii) Spline analysis of shared fitness surfaces

In order to understand ecological speciation, it is necessary to determine if populations are experiencing divergent natural selection. Fitness landscapes that show a fitness valley at intermediate phenotypes can help to reveal this process. For this purpose, sharing the fitness functions between closely related species can help to understand how selection against hybrids with intermediate phenotypes may drive or maintain reproductive isolation. But this advantage comes with a drawback: because the shared fitness function assumes that identical phenotypic values result in the same fitness (or survival probability), the topology of the fitness landscape is not as variable as those obtained with independent fitness functions. As a consequence, the shared landscape can potentially obscure fine-scale differences in the form of selection acting on each species, in particular when there is overlap in the phenotypic values of different species. For example, *G. scandens* shows directional selection in the independent fitness landscape. But in the shared surface, the relationship between phenotype and survival probabilities in *G. scandens* is mostly driven by the more common species *G. fortis*. Likewise, high overlap in the beak shape values of the small, medium and large ground finches lead to the shared fitness surface obscuring the distinct topologies observed when using independent fitness functions. In contrast, beak size shows the same form of selection in the shared landscape as it does on the independent landscapes. This may reflect the clearer partitioning of the phenotypic distribution of beak size across these three species, as well as the consistency of selection on this trait.

Thus, the clearest understanding of the process of ecological speciation may arise from construction of fitness landscapes with independent functions for each species. The main reason for this is that assuming equal fitness for the same trait values across species masks subtle variation in the shape of the landscapes (with correspond to types of selection) that may be

crucial for driving initial population differentiation, or permitting introgression and the survival of hybrids. This issue is particularly relevant when trait values overlap. For instance, I found considerable overlap in beak morphology between *G. fortis* and *G. fuliginosa*, *G. magnirostris* (Figure 3; Figure 4; Figure 5). The two peaks on the fitness surface of *G. fortis* could be formed because *G. fuliginosa* and *G. magnirostris* have a higher survival than *G. fortis* in some years (e.g. the dry year in 2006; Figure 4). If *G. fortis* is hybridizing with the two other species at those times, I would expect to see new intermediate phenotypes such as the small and large morphs of the medium ground finch. A profitable direction for future work would be to identify hybrids using genetic data and explicitly consider how these individuals alter the topography of species-specific fitness landscapes.

Temporal variation in species-specific patterns of beak evolution

Temporal variation in environmental conditions (in this case, precipitation) resulted in shifting topologies of the fitness landscapes across years. Inspection of fitness landscapes estimated from the quadratic model show that 1) the fitness values (height of the whole fitness function) of the finch populations are not the same from year to year, and 2) there is a change in the shape of the fitness surface. Several species-specific patterns emerge from these changes. *G. fuliginosa* shows evidence of disruptive selection, with evolution towards pointier and blunter ends of the beak shape spectrum. Similarly, *G. fortis* has a general tendency to evolve towards either pointier or blunter beaks, but it is consistently evolving smaller and larger beaks. In contrast, the *G. scandens* population shows either evolution towards bigger beaks or stabilizing selection for all traits. Finally, *G. magnirostris* showed evidence for directional selection towards either a blunter beak or a smaller beak. Importantly, these evolutionary trajectories reflect movement of fitness peaks across the landscape as opposed to changes in the number of peaks. Peak number is

consistent for all fitness surfaces estimated from the quadratic model: *G. fuliginosa* always shows one peak (but since it is switching between pointier and blunter, it increases the variance in the trait distribution), *G. fortis* two peaks, *G. magnirostris* one peak, and *G. scandens* one. Thus, over the time frame of our study we observe no changes in the form of selection in specific species (e.g., directional selection changing to disruptive, and resulting in the creation of two fitness peaks from one).

Perhaps surprisingly given past work in this system, the form of selection experienced by these populations does not appear to be driven by rainfall since the climate variable in the quadratic model showed credible intervals that overlap 0. One exception to this pattern is that the medium ground finch (*G. fortis*) evolves towards a pointier beak shape in years when there is high precipitation (reflected by a negative interaction effect between beak shape and climate in the quadratic model). This suggests that an extended period of rain over many years could cause the medium ground finch to evolve to resemble a cactus finch-like beak shape. This scenario coupled with mis-imprinting of the song of *G. fortis* from *G. scandens* might explain why the two species hybridize, especially in years with high precipitation (Grant & Grant, 1997). Besides this result, there were no significant effects of climate on the direction of selection on any other traits in any of the species. It is possible that the weak effect of climate observed in this study relative to the strong impact of this variable in earlier work is that previous studies encompassed some very extreme fluctuations in precipitation (e.g., precipitation in 1983 was 2769 mm but dropped to just 157 mm in 1984; Figure 8). Indeed, I do find a significant effect of year on survival in 2003 and in 2013, which are two of the driest years in this study (190 and 76 mm of rain). Similarly, in the year with the highest precipitation in the dataset (2008), the survival probabilities of the large ground finch, *G. magnirostris*, were higher than both the small and

large morphotype of the medium ground finch, *G. fortis*. This might present a situation in which it could be advantageous for *G. fortis* to introgress with *G. magnirostris*. Thus, while extreme drought or rainfall may occur rarely, these events could have out-sized effects on finch abundance and ensuring evolutionary dynamics. As a potential caveat, the rainfall might not represent exactly the climate at El Garrapatero since rain is rarely falling at the same time on El Garrapatero and Puerto Ayora (personal observation; Figure 8). I would conclude that these estimates are overestimating the actual quantity of rain for El Garrapatero.

Niche expansion from G. fortis to G. magnirostris

Darwin's finches are continuously evolving in their resource use, which can either "expand" or "pack" their morphospace (Pigot et al. 2016). Morphospace expansion is the extension of trait distributions to new phenotypic values whereas packing is the development of already existing phenotypes, but increasing their density. Expansion of the morphospace is observed in *G. fortis*, which is currently extending towards *G. magnirostris* morphospace. This process can be identified through two patterns in the independent fitness landscapes: 1) *G. fortis* is being selected for larger beaks while *G. magnirostris* is being selected for smaller beaks, which suggests that there might be an optimum beak size between the two species. 2) Low abundance of *G. magnirostris* suggests that it may not be completely filling its niche at our sampling sites, which may result in selection favouring large beaked *G. fortis* that can expand into this morphospace. Once the expansion of a morphotype reaches a fitness peak, as currently suggested in our quadratic model and slightly in the thin-plate spline between *G. fortis* and *G. magnirostris*, the new phenotypic distribution will increase in density, i.e. "packing" the phenotypes.

Summary

Comparing fitness landscapes of multiple species that take into account environmental variation can improve understanding about the process of ecological speciation in sympatric populations. Here, I showed that the most informative way to construct fitness landscapes that can inform about the environmental conditions that promote or inhibit divergence is to use independent fitness functions for each species. The alternative approach, to build shared fitness landscapes for multiple species, makes the restrictive assumption of equal fitness between individuals sharing the same phenotypic trait values. The consequence of this assumption is that subtle variation in the form of selection acting on different species is obscured. This is particularly the case when traits values overlap heavily between species. However, further work is still required to understand how hybrids should be included in independent fitness landscapes. My work also shows that environmental variation can result in temporal changes to the fitness surface for some species. This finding reiterates the importance of understanding how the environment might shape the phenotypic distribution of sympatric species. Collectively, my thesis demonstrates that constructing fitness landscapes for interacting species complexes can help to understand how temporally varying selection can both facilitate and impede ecological speciation and drive the emergence of novel biodiversity.

Appendix

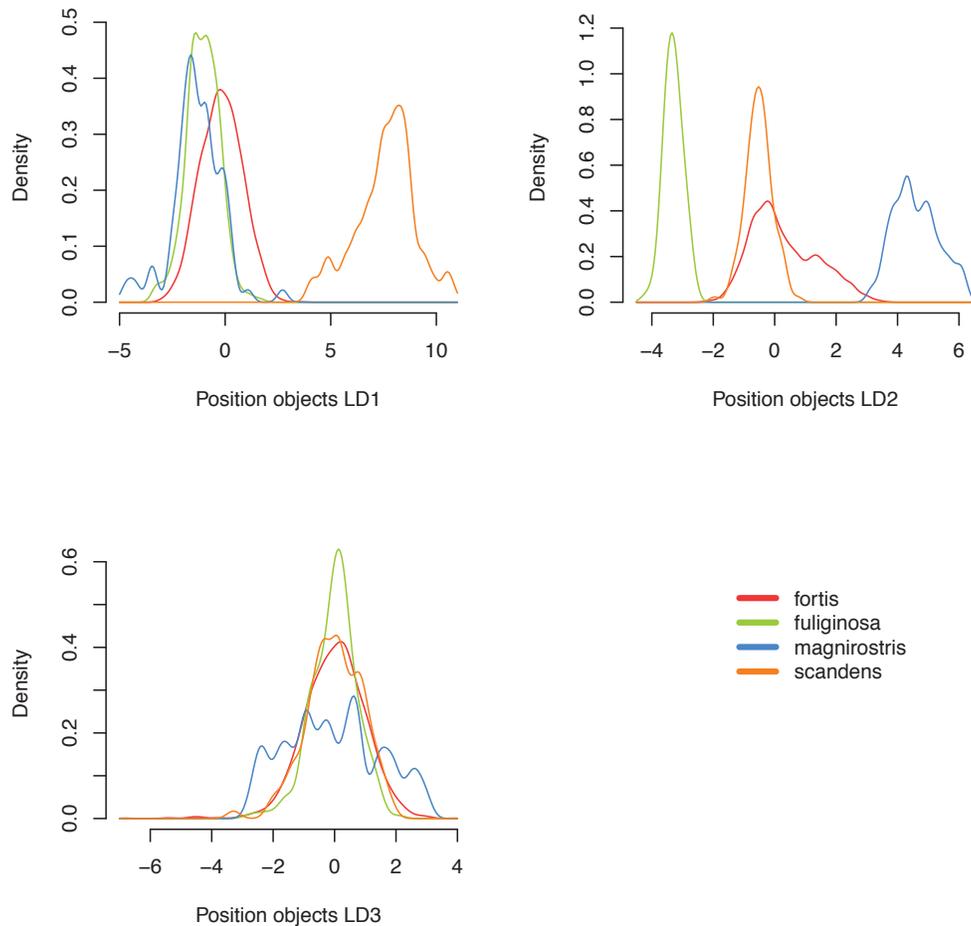


Figure S1 Density of the data in relation to the position of the objects in the LDA (x-axis) for the 3 LDA axis.

Principal component Analysis

Is this possible to ecologically interpret in less dimensions the variance explained by beak traits?

Methods description

The principal component analysis (PCA) is a multivariable ordination method without constrain on the quantitative dependent variables (Legendre & Legendre, 2012). This method is applied in many contexts such as a complement for cluster analysis, when someone wants to understand the

structure in variation of many quantitative variables while reducing the dimensionality of the dataset. Ordination methods are thus useful to interpret complex ecological data. PCAs, a method that uses eigenvalues and eigenvectors based on a covariance (or correlation) matrix \mathbf{S} , have three main properties. The \mathbf{S} matrix is symmetrical and positive definite meaning that the 1. principal axes are orthogonal and that 2. the eigenvalues will be positive non-null representing the amount of variance of the data on the principal axes. 3. The last property stipulates that this method explain a maximum of the variation of the explanatory variables with less dimensions.

The suppositions of the PCA are to have quantitative descriptors with homogeneous units, reasonably multinormal and not having too many zeros. The principal axes are found using the analysis of eigenvalues and eigenvectors: $(\mathbf{S} - \lambda_k \mathbf{I})\mathbf{u}_k = 0$, where \mathbf{I} , is an identity matrix, $\mathbf{S} = (n-1)^{-1} \mathbf{Y}_c' \mathbf{Y}_c$, where \mathbf{Y}_c is the column centroid matrix ($\mathbf{Y}_c = [y - \bar{y}]$), n is the number of objects, and the index k represents the principal axes. The eigenvalues and eigenvectors are respectively λ_k and \mathbf{u}_k . One of the properties of the equation is that the determinant has to equal 0, making it possible to compute this equation: $(|\mathbf{S} - \lambda_k \mathbf{I}| = 0)$. The \mathbf{U} matrix is a column matrix of the scaled eigenvectors to 1. To represent the objects graphically, in relation to the axes, two types of scaling exist. (1) For scaling 1, or biplot preserving the Euclidean distances, the matrix of principal components is computed using $\mathbf{F} : \mathbf{F} = \mathbf{Y}_c \mathbf{U}$. In addition, the orthogonal projection of an object on a descriptor gives an approximation of the position of the object in the multidimensional space. The scaling 1 is used to interpret the relation between the points. In this case, the interpretation is similar to a cluster analysis where the closer are objects, the more they are similar. The angles between the axes of the descriptors is meaningless with this scaling. But the length of a descriptor's projection on an axis in the reduced space is an index of its contribution. The comparison of the descriptors can be done only on one axis at the time because

the principal axes have different proportionality factors. The equilibrium circle of contribution of radius $\sqrt{(d/p)}$, where d is the number of dimensions of space and p is the number of orthogonal axes is drawn to represent the contribution of a descriptor. The length of the eigenvectors is compared to the circle. If an eigenvector is smaller than the radius, its contribution is reduced whereas a vector longer than the radius means that it contributes more to the space.

(2) The second scaling or biplot preserving correlations retains the Mahalanobis distance between the objects. The angle between descriptors represents their correlation (0° and 180° means a positive and negative covariance whereas 90° means a null covariance). In scaling 2, the eigenvectors are scaled by the square root of their eigenvalues ($\sqrt{\lambda_k}$) which transform the length of the projection of the descriptors to equate their standard error. The \mathbf{G} matrix represent the position of the objects in the biplot: $\mathbf{G} = \mathbf{F}\mathbf{\Lambda}^{-1/2}$, where $\mathbf{\Lambda}$ is the diagonal matrix of the eigenvalues. The circle of equilibrium contribution of descriptors is measured by: $s_j\sqrt{(d/p)}$ where s_j is the length of the j descriptor. If the descriptors are scaled, the radius of the equilibrium contribution circle will be $\sqrt{(d/p)}$. Finally, the scaling 2 is used if the relation between the descriptors is the question of interest.

PCA Results

Darwin's finches are an example of adaptive radiation where each "species" adapted to a particular niche. Variation in the adaptive traits, such as the beak, is thus expected and a cluster representing closely related individuals. There is indeed variation in beak traits which correspond to the traits that are forming different morphotypes (it is notably the case for beak length, width and depth, Figure S2). The first two axes explain a total of 98.88% of the variance (90.9 % and 7.98 % respectively). The rest of the variance is in the last axis. Each species of finches is grouped by their beak morphology. The circle of equilibrium contribution is shown for scaling 1

(Figure S2A), but not for scaling 2 (since all the beak traits are in millimetres, there was no need to scale them). The variable that contributes the most to the projection is the median beak length since the eigenvector is longer than the radius of the circle. For scaling 2 (Figure S2B), the median beak width is highly correlated to the median beak depth since the vectors are pointing in the same direction. The median beak length is slightly correlated to the other two traits. Thus, PC1 axis may be biologically interpreted as the size of the beak and PC2 beak shape (Grant, 1994).

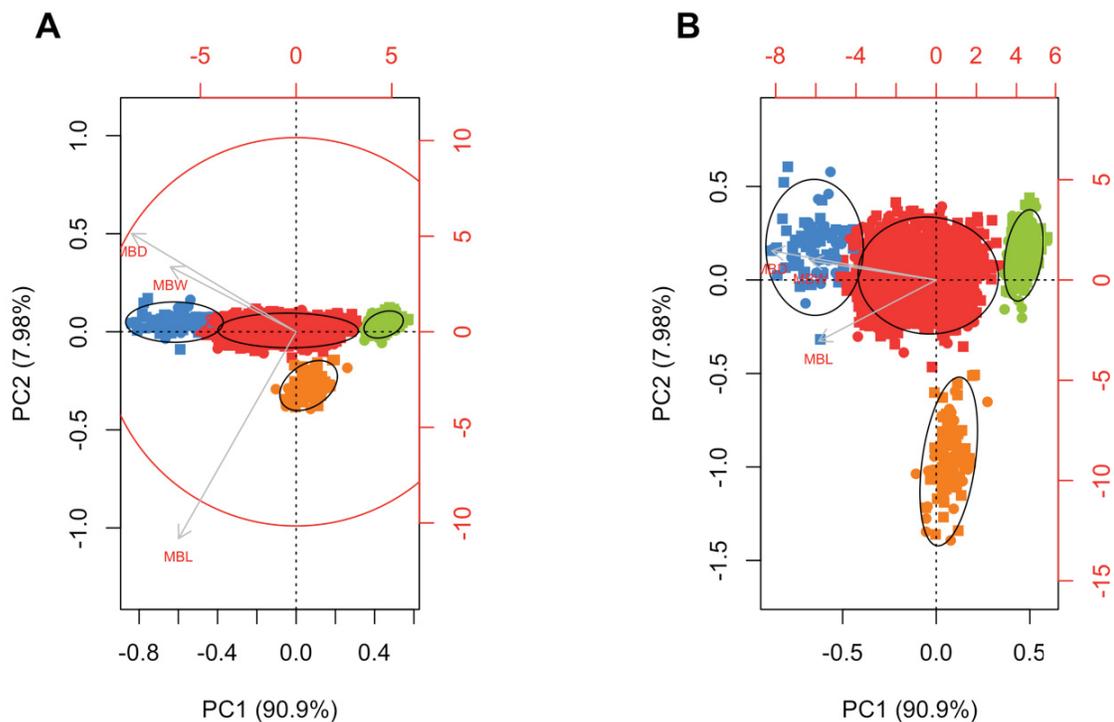


Figure S2 Graphical results from PCA showing PC1 and PC2 of the beak traits (length, width, and depth).

A) double projection preserving distances (scaling 1) B) double projection preserving correlations (scaling 2). Squares (■) represent the site Academy Bay and circles (●) represent the site El Garrapatero. MBL : Median beak length, MBW : Median beak

width, MBD : Median beak depth. The proportion of the variance explained by the first axis is 90.9% whereas PC2 is 7.98%. Both sites (Academy Bay and El Garrapatero) are clustering together. Red: *G. fortis*; Green: *G. fuliginosa* ; Blue *G. magnirostris*; Orange *G. scandens*. The ellipsis comes from the library *vegan* with a 95 % confidence limit.

Discussion

The different beak traits of Darwin's finches are forming two major components explaining most (98.88%) of the variation. It is possible to ecologically interpret the two major components: the first principal component (PC1) can be interpreted as beak size whereas the second component (PC2) can be attributed to beak shape (Grant, 1986). This has implications in the construction of a multispecies fitness landscape using the community of Darwin's finches in a bidimensional morphospace using the principal component scores. Another observation is that the majority of the variation is explained by the first component.

The amount of variation in the first principal component can be explored on an evolutionary basis. For example, perhaps the finches have a genetic constrain to diversify their beak length. Or perhaps, it may be more advantageous to develop more diverse beak sizes than beak shapes.

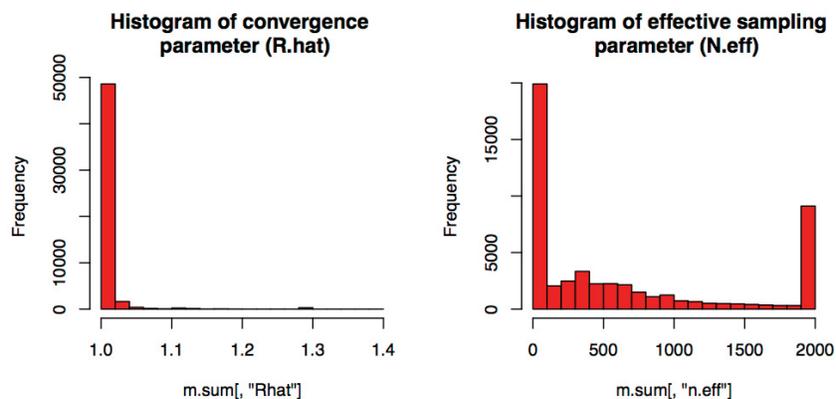


Figure S3 Quadratic model convergence and effective sampling.

These histograms will vary depending on the parameter tracked. There is a lot of small value in the effective sampling parameter histogram because the state process matrix was tracked. The model was run for 105 000 iterations.

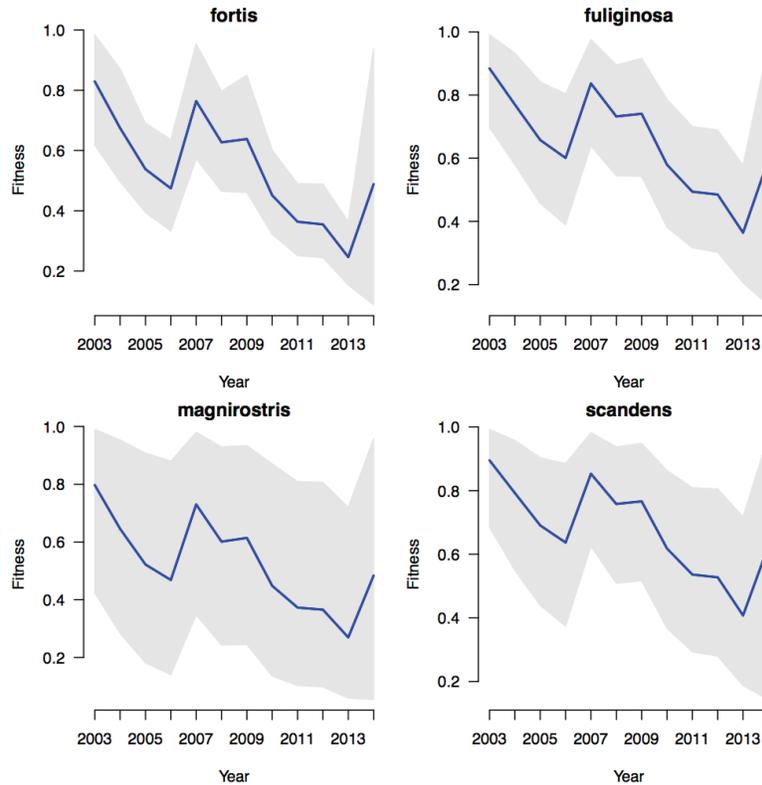


Figure S4 Credible interval (95%) of the species per year.

Validation and Simulations of the Bayesian model

To validate our model, I tested it using simulated data, examining accuracy and precision of estimated parameters. I first created a script to numerically simulate populations containing multiple species. I created a new matrix representing the capture history (the observational state) and a state process matrix of the simulated species. The dimensions of the matrix are given by a targeted number of individuals per simulated species and the number of years simulated.

I decided to add a burn-in period to discard the first years in the simulation for the capture history to avoid any border effects at the beginning of the simulation. In other words, in order to generate a capture history that will reflect the recapture rates in the simulation, I need to remove the beginning of the simulated capture history to remove any artefact in the generated populations. For our simulated analysis, I kept only the last generated values in the simulation (in this case the last twelve years of the simulation).

The recapture and survival probabilities are set with a mean and a standard deviation to generate probabilities of survival sampled from a normal distribution.

$$\begin{aligned}\Phi.simulated[sp, year] &= \mathcal{N}[sp](\mu_1, \sigma^2) + \mathcal{N}[year](\mu_3, \sigma^2) \\ p.simulated[sp, year] &= \mathcal{N}[sp](\mu_2, \sigma^2) + \mathcal{N}[year](\mu_3, \sigma^2)\end{aligned}$$

The two matrix $\Phi.simulated$ and $p.simulated$ contain in rows the species and in columns the years of the combined effect of species and year on survival and recapture probabilities respectively. I will use my Bayesian mark-recapture model to recover the mean estimate of the effect of species on survival.

I generated simulated the bird's phenotypes using a normal distribution such that individual trait values are from species specific mean and variance of empirical distribution of the three beak traits (beak length, depth, and width; Table 1). A principal component analysis is computed on the traits to extract the axes scores.

Table IV Mean and variance of traits used in the simulation analysis of empirical data

	<i>G. fortis</i>	<i>G. fuliginosa</i>	<i>G. magnirostris</i>	<i>G. scandens</i>
Beak Width	9.83 (0.99)	6.84 (0.1)	13.17 (0.98)	8.26 (0.19)
Beak Length	11.64 (0.94)	8.51 (0.22)	14.48 (0.76)	14.21 (0.89)
Beak Depth	11.09 (1.75)	7.09 (0.16)	15.27 (1.99)	8.59 (0.29)

All traits are in millimetre.

These scores will be the input for the Bayesian simulation analysis. I assumed that survival probability is a linear combination of the species survival intercept and the survival probability of the different traits:

$$\begin{aligned}\text{logit}(\Phi[i]) &= \Phi.sp[sp_{[i]}] \\ &+ \Phi.year[t] \\ &+ \Phi.trait[k] \times trait[k,i]\end{aligned}$$

where $\Phi.trait[k]$ is the vector probability of survival having trait k and $trait[k,i]$ is the matrix of traits for all individuals (i).

Allowing for imperfect detection enabled me to insert missing recaptures in the capture history. I accomplished this by selecting all of the individuals that survived ($n[i]$, represented as 1s) and switching them to 0 with the probability of recapture ($p[i]$) of our simulated species with a binomial distribution $CH_{survived} = B(n[i], p[i])$. Since the number of simulated species is generated from a probability of being detected, the individuals that are never recorded in the simulated capture history (have a line showing no capture or filled with “0”s) are dropped modifying the targeted total number of individual simulated. Thus, I allow the matrix to be off (i.e. containing more or less simulated individuals) of the initial target by 5%. If the number of dropped individuals is above the threshold of the initial target, new simulated individuals will be added to the capture history with the survival and recapture rates set in the simulation.

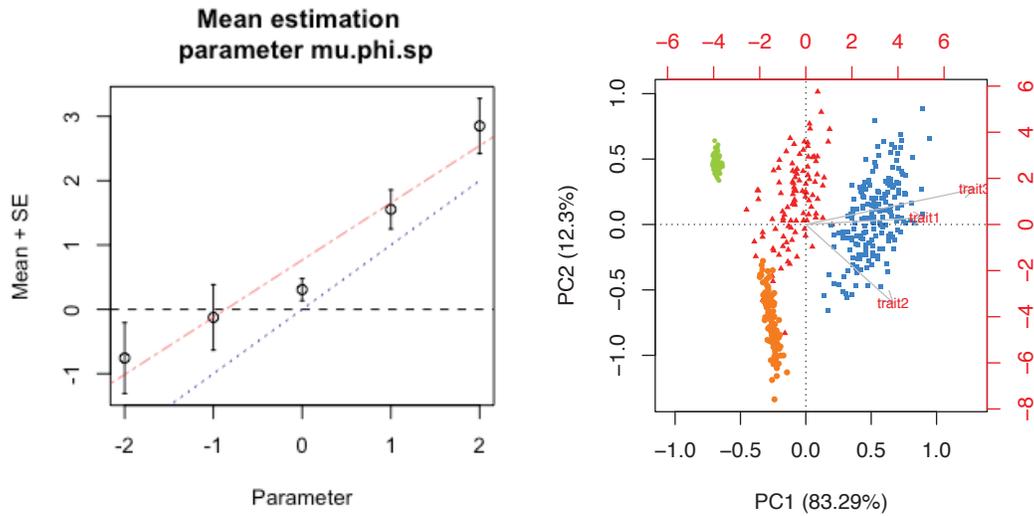


Figure S5 Simulation analysis of the mean value of the random intercept of species effect on survival and PCA of simulated data.

Left panel: The red line is a linear model from the parameter computer by the Bayesian model. Each point is the mean of 12 model run of the mean effect of survival on each species (on a logit scale, -2, -1, 0, 1, and 2, correspond to probabilities of survival of 0.12, 0.73, 0.50, 0.73, and 0.88). Each model ran with 15000 iterations with a burn-in of 5000 and a thinning of 5. The blue line is the value expected to be found by the Bayesian model. Right panel: Principal component analysis of the simulated data based on Table IV.

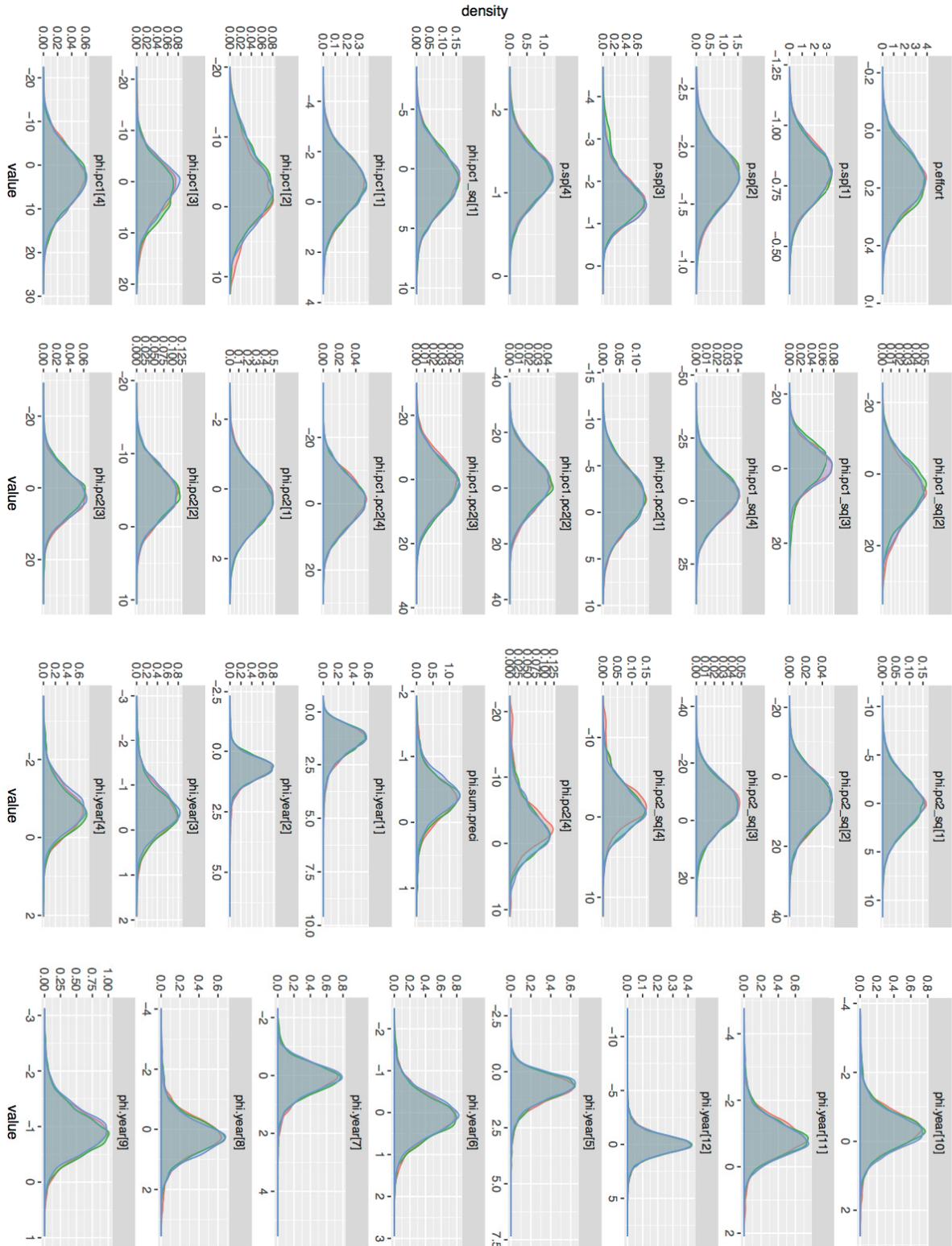


Figure S6 Posterior distribution of the generalized linear mixed effect model

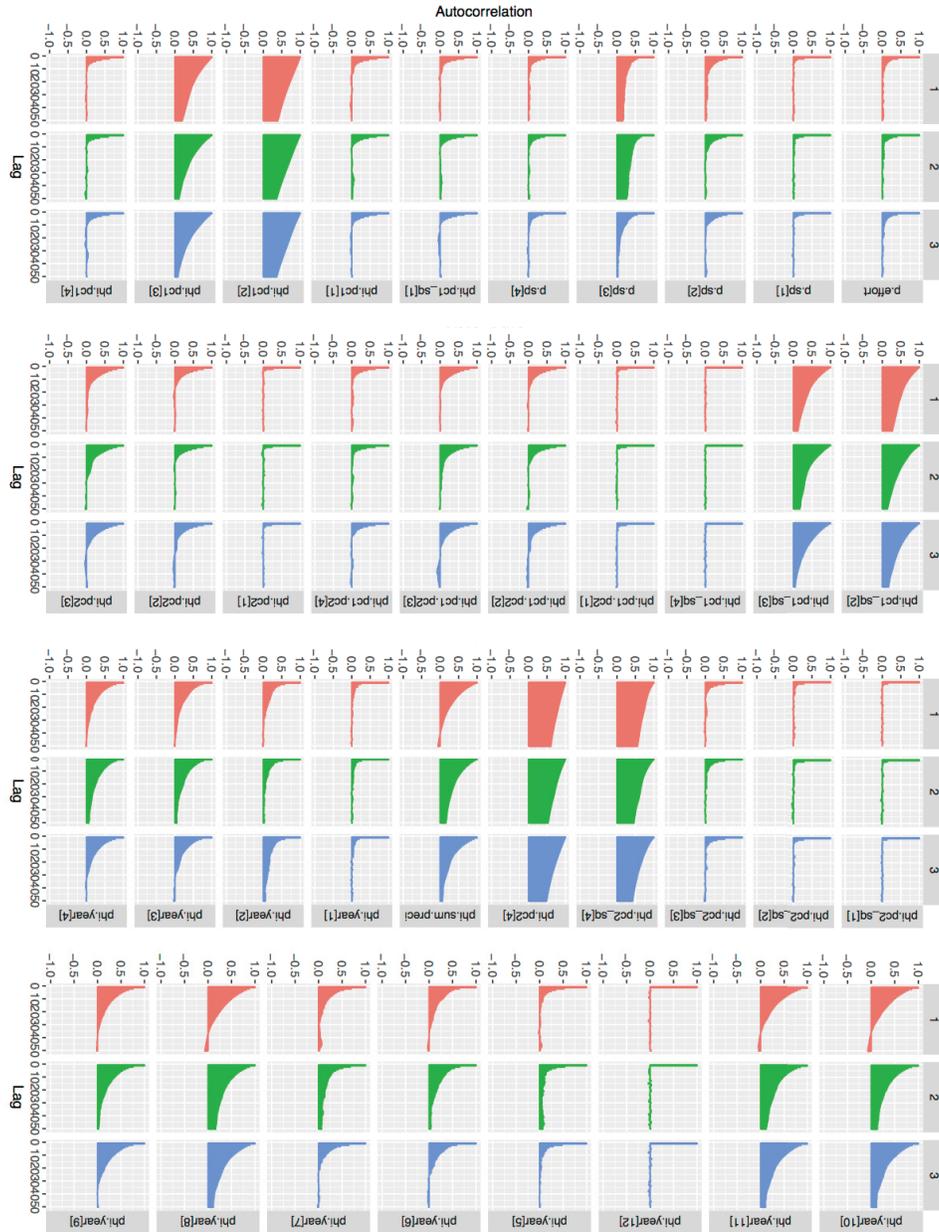


Figure S8 Autocorrelation plot of the generalized linear mixed effect model.

It is expected to observe an autocorrelation near 0 at the end of the execution of the model. A positive autocorrelation means that the values tested in the Gibbs algorithm are too similar, preventing a proper exploration of the parameter space. On the other hand, a negative autocorrelation means that the values are too dissimilar.

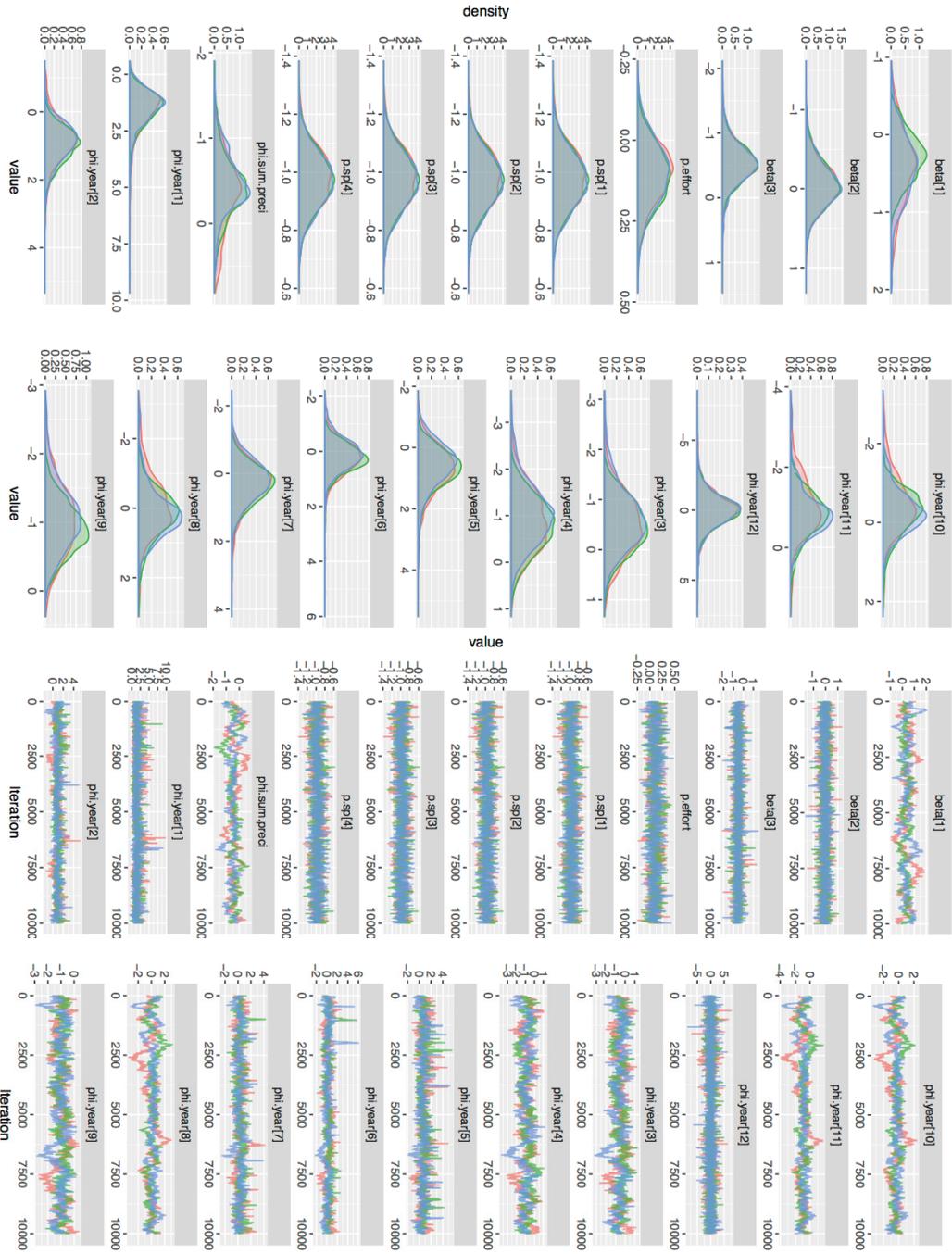


Figure S9 Posterior distribution (density graph) and traceplot of the thin-plate spline model.

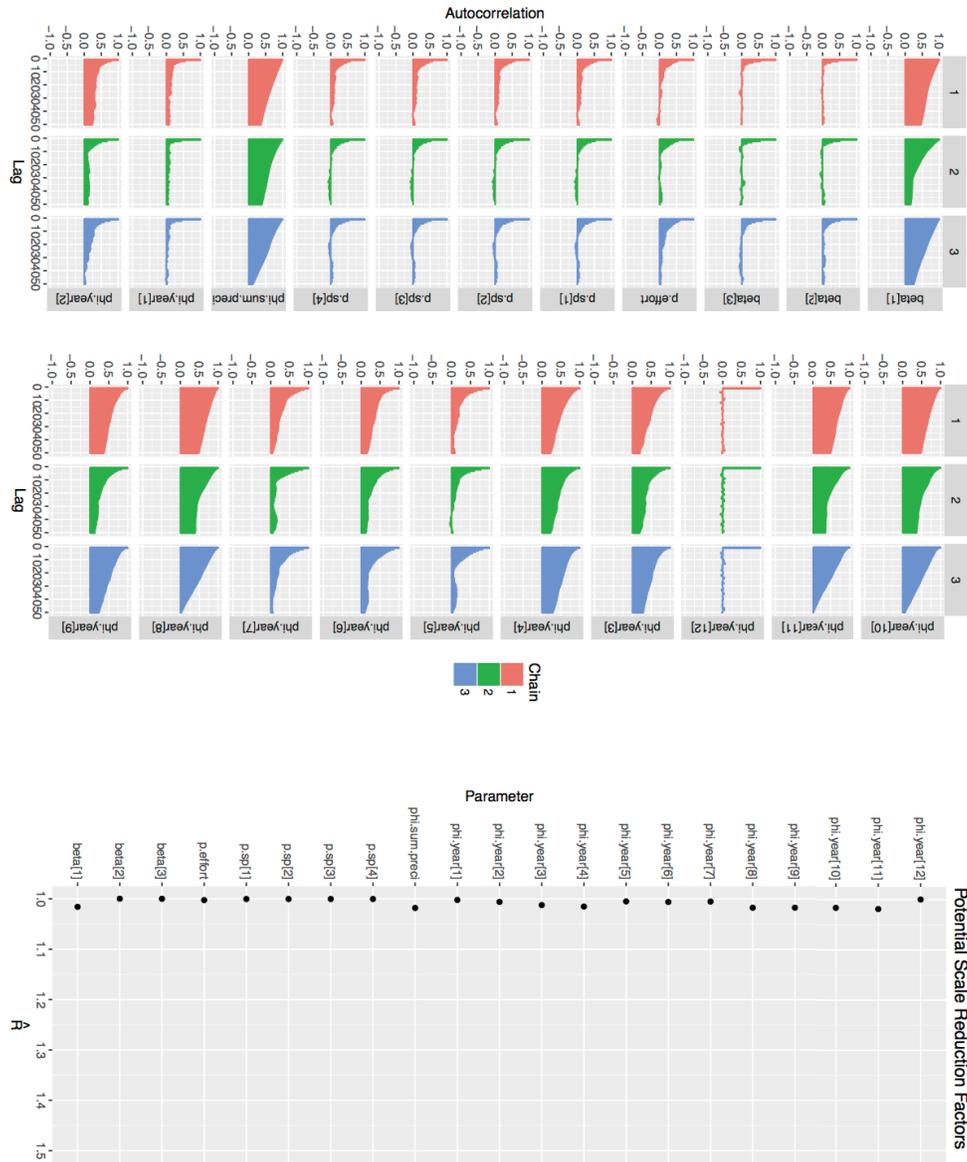


Figure S10 Autocorrelation plot of the thin-plate spline model.

It is expected to observe an autocorrelation near 0 at the end of the execution of the model. A positive autocorrelation means that the values tested in the Gibbs algorithm are too similar, preventing a proper exploration of the parameter space. On the other hand, a negative autocorrelation means that the values are too dissimilar. \hat{R} (Potential scale reduction factor) are shown in the bottom panel.

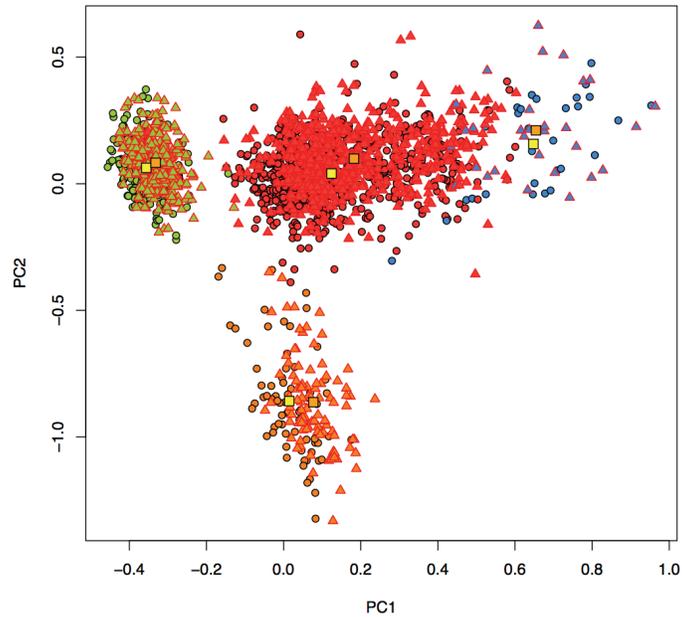


Figure S11 Principal component analysis of the four ground finches.

The squared represents the median phenotype of the principal component scores per sex (yellow for females and orange for males). There is a statistical difference, using an ANOVA, between the sexes for both axes in *G. fortis* (PC1 F-value = 45.33, p -value = 0; PC2 F-value = 71.76, p -value = 0) and *G. fuliginosa* (PC1 F-value = 57.58, p -value = 0; PC2 F-value = 8.24, p -value = 0.004), but only a significant difference for the first axis for *G. scandens* (F-value = 41.43, p -value = 0). There is no difference for *G. magnirostris* (p -value > 0.25 for both axes). Red: *G. fortis*; Green: *G. fuliginosa*; Blue *G. magnirostris*; Orange *G. scandens*. Triangles: females; Circles: males. (Grant, 1993)

Literature Cited

- Abbott, I., Abbott, L.K. & Grant, P.R. 1977. Comparative ecology of Galápagos ground finches (*Geospiza* Gould): evaluation of the importance of floristic diversity and interspecific competition. *Ecol. Monogr.* **47**: 151–184.
- Amstrup, S.C., McDonald, T.L. & Manly, B.F.J. 2005. *Handbook of Capture-Recapture Analysis*. Princeton University Press.
- Arnold, S.J., Pfrender, M.E. & Jones, A.G. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112-113**: 9–32.
- Benkman, C.W. 2003. Divergent selection drives the adaptive radiation of crossbills. *Evolution* **57**: 1176–1181.
- Cormack, R.M. 1964. Estimates of survival from the sighting of marked animals. *Biometrika* **51**: 429–438.
- De León, L.F., Bermingham, E., Podos, J. & Hendry, A.P. 2010. Divergence with gene flow as facilitated by ecological differences: within-island variation in Darwin's finches. *Phil. Trans. R. Soc. B* **365**: 1041–1052.
- De León, L.F., Podos, J., Gardezi, T., Herrel, A. & Hendry, A.P. 2014. Darwin's finches and their diet niches: the sympatric coexistence of imperfect generalists. *J. Evol. Biol.* **27**: 1093–1104.
- De León, L.F., Rolshausen, G., Bermingham, E., Podos, J. & Hendry, A.P. 2012. Individual specialization and the seeds of adaptive radiation in Darwin's finches. *Evol. Ecol. Res.* **14**: 365–380.
- Dieckmann, U., Doebeli, M., Metz, J.A.J. & Tautz, D. 2004. *Adaptive Speciation*. Cambridge University Press.
- Fields Development Team. 2006. fields: Tools for Spatial Data. National Center for Atmospheric Research. Boulder, CO.
- Gavrilets, S. 2004. *Fitness Landscapes and the Origin of Species*. Princeton University Press.
- Gavrilets, S. & Gravner, J. 1997. Percolation on the fitness hypercube and the evolution of reproductive isolation. *J. Theor. Biol.* **184**: 51–64.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. 2013. *Bayesian Data Analysis*. CRC Press.
- Geman, S. & Geman, D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**: 721–741.

- Gimenez, O., Grégoire, A. & Lenormand, T. 2009. Estimating and visualizing fitness surfaces using mark-recapture data. *Evolution* **63**: 3097–3105.
- Grant, B.R. & Grant, P.R. 2008. Fission and fusion of Darwin's finches populations. *Philos. Trans. R. Soc. B* **363**: 2821–2829.
- Grant, B.R. & Grant, P.R. 1996. High survival of Darwin's finch hybrids: effects of beak morphology and diets. *Ecology* **77**: 500–509.
- Grant, B.R. & Schluter, D. 1984. Interspecific competition inferred from patterns of guild structure. In: *Ecological communities: conceptual issues and the evidence* (D. R. Strong, D. S. Simberloff, L. G. Abele, & A. B. Thistle, eds).
- Grant, P.R. 1986. *Ecology and evolution of Darwin's finches*. Princeton University Press.
- Grant, P.R. 1993. Hybridization of Darwin's finches on Isla Daphne Major, Galápagos. *Philos. Trans. R. Soc. Lond. B* **340**: 127–139.
- Grant, P.R. 1994. Population variation and hybridization: comparison of finches from two archipelagos. *Evol. Ecol.* **8**: 598–617.
- Grant, P.R. & Grant, B.R. 2006. Evolution of character displacement in Darwin's finches. *Science* **313**: 224–226.
- Grant, P.R. & Grant, B.R. 1997. Hybridization, sexual imprinting, and mate choice. *Am. Nat.* **149**: 1–28.
- Grant, P.R. & Grant, B.R. 1995. Predicting microevolutionary responses to directional selection on heritable variation. *Ecology* **49**: 241–251.
- Grant, P.R. & Grant, B.R. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**: 707–711.
- Hendry, A.P. & Gonzalez, A. 2008. Whither adaptation? *Biol. Philos.* **23**: 673–699.
- Hendry, A.P., Huber, S.K., De Leon, L.F., Herrel, A. & Podos, J. 2009. Disruptive selection in a bimodal population of Darwin's finches. *Proc. R. Soc. B* **276**: 753–759.
- Huber, S.K., De León, L.F., Hendry, A.P., Bermingham, E. & Podos, J. 2007. Reproductive isolation of sympatric morphs in a population of Darwin's finches. *Proc. R. Soc. B* **274**: 1709–1714.
- Jolly, G.M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* **52**: 225–247.
- Kaplan, J. 2008. The end of the adaptive landscape metaphor? *Biol. Philos.* **23**: 625–638.
- Kéry, M. & Schaub, M. 2012. *Bayesian Population Analysis Using WinBUGS*. Academic Press.

- Kimura, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- Lack, D. 1947. *Darwin's Finches*. Cambridge University Press.
- Lande, R. & Arnold, S.J. 1983. The Measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.
- Legendre, P. & Legendre, L. 2012. *Numerical Ecology*. Elsevier.
- MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. 2005. *Occupancy Estimation and Modeling*. Academic Press.
- Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proc. Natl. Acad. Sci. U.S.A.* **2**: 49–55.
- Martin, C.H. 2016. Context dependence in complex adaptive landscapes: frequency and trait-dependent selection surfaces within an adaptive radiation of Caribbean pupfishes. *Evolution* **70**: 1265–1282.
- Martin, C.H. & Wainwright, P.C. 2013. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* **339**: 208–211.
- McCoy, J.W. 1979. The origin of the “adaptive landscape” concept. *Am. Nat.* **113**: 610–613.
- Mitchell-Olds, T. & Shaw, R.G. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* **41**: 1149–1161.
- Nosil, P. 2012. *Ecological Speciation*. Oxford University Press.
- Nychka, D., Furrer, R., Paige, J. & Sain, S. 2015. *fields: Tools for spatial data*. University Corporation for Atmospheric Research. Boulder, CO, USA.
- Orr, H.A. 2009. Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.* **10**: 531–539.
- Plummer, M. 2003a. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International workshop on distributed statistical computing*.
- Plummer, M. 2003b. rjags: Bayesian Graphical Models using MCMC.
- Price, T.D. 1984. The evolution of sexual size dimorphism in Darwin's finches. *Am. Nat.* **123**: 500–518.
- R Core Team. 2016. *R: a language and environment for statistical computing*. Vienna, Austria.
- Raeymaekers, J.A.M., De León, L.F., Chaves, J., Sharpe, D.M.T., Huber, S.K., Herrel, A., Vanhooydonck, B., Koop, J.A.H., Knutie, S.A., Clayton, D.H., Grant, B.R., Grant, P.R., Podos, J. & Hendry, A.P. n.d. Spatiotemporal variation in contemporary Darwin's finches;

- implications for adaptive radiation. Not published work.
- Ratcliffe, L.M. & Grant, P.R. 1983. Species recognition in Darwin's finches (Geospiza, Gould). I. Discrimination by morphological cues. *Anim. Behav.* **31**: 1139–1153.
- Raup, D.M. 1967. Geometric analysis of shell coiling: coiling in Ammonoids. *J. Paleo.* **41**: 43–65.
- Roches, Des, S., Sollmann, R., Calhoun, K., Rothstein, A.P. & Rosenblum, E.B. 2016. Survival by genotype: patterns at Mc1r are not black and white at the White Sands ecotone. *Mol. Ecol.* **26**: 320–329.
- Ruppert, D., Wand, M.P. & Carroll, R.J. 2003. *Semiparametric regression*. Cambridge University Press.
- Salvin, O. 1876. X. On the avifauna of the Galápagos Archipelago. *Trans. Zool. Soc. London* **9**: 447–510.
- Schluter, D. 2001. Ecology and the origin of species. *Trends Ecol. Evol.* **16**: 372–380.
- Schluter, D. 1984. Feeding correlates of breeding and social organization in two Galápagos finches. *Auk* **101**: 59–68.
- Schluter, D. & Grant, P.R. 1984. Determinants of morphological patterns in communities of Darwin's finches. *Am. Nat.* **123**: 175–196.
- Schluter, D. & Nagel, L.M. 1995. Parallel speciation by natural selection. *Am. Nat.* **146**: 292–301.
- Schluter, D. & Nychka, D. 1994. Exploring fitness surfaces. *Am. Nat.* **143**: 597–616.
- Seber, G.A. 1965. A note on the multiple-recapture census. *Biometrika* **52**: 249–259.
- Shafer, A.B.A. & Wolf, J.B.W. 2013. Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecol. Lett.* **16**: 940–950.
- Siepielski, A.M., Dibattista, J.D. & Carlson, S.M. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* **12**: 1261–1276.
- Simpson, G.G. 1944. *Tempo and Mode in Evolution*. Columbia University Press.
- Simpson, G.G. 1953. *The Major Features of Evolution*. Columbia University Press.
- Smith, J.N.M., Grant, P.R., Grant, B.R., Abbott, I.J. & Abbott, L.K. 1978. Seasonal variation in feeding habits of Darwin's ground finches. *Ecology* **59**: 1137–1150.
- Sober, E. 2001. The two faces of fitness. In: *Thinking about evolution* (R. S Singh, C. B. Krimbas, D. B. Paul, & J. Beatty, eds), pp. 309–321.

Su, Y.-S. & Yajima, M. 2015. R2jags: Using R to Run “JAGS.”

Svensson, E.I. & Calsbeek, R. 2012. *The adaptive landscape in evolutionary biology*. Oxford University Press.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. pp. 356–366. In *Proceedings of the Sixth International Congress of Genetics*.