Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Cotutelle internationale avec :

Présentée et soutenue par :
Boris Delahaie

Le 13 mars 2015

Titre :

Spéciation, gradients environnementaux et zones hybrides : le cas du Zostérops des Mascareignes

---

École doctorale et discipline ou spécialité :
ED SEVAB : Écologie, biodiversité et évolution

Unité de recherche :
UMR5174 Evolution et Diversité Biologique

Directeur(s) de Thèse :

Christophe Thébaud
Borja Milá

Rapporteurs :

Nicolas Bierne
Jean Secondi

Autre(s) membre(s) du jury :

Pauline Garnier-Géré
Etienne Danchin
Christophe Thébaud
Borja Milá

Avant propos

Cette thèse en Écologie, biodiversité et évolution de l'École doctorale des Sciences écologiques, vétérinaires, agronomiques et bioingénieries (SEVAB) de l'Université Toulouse Paul Sabatier a été financée par une bourse de recherche ministérielle. Ces recherches ont également été financées par un projet de la Fondation pour la Recherche sur la Biodiversité (FRB) et du National Geographic. Ce travail s'est déroulé au sein du laboratoire Évolution et Diversité biologique (UMR 5174 CNRS-UPS). Le manuscrit est composé de trois parties : (i) une introduction générale, (ii) quatre articles scientifiques et (iii) une discussion générale faisant le lien entre les différents résultats obtenus et proposant des pistes de recherches futures.

Remerciements

Mes premiers remerciements s'adressent évidemment à mes deux directeurs de thèse. Merci à tous les deux de m'avoir permis de réaliser cette thèse et de m'avoir fait confiance tout au long de ce travail. Christophe, dès la L3, je te jalousais lorsque tu nous avais présenté tes recherches sous les tropiques, tu m'as finalement permis d'y prendre part et je t'en suis infiniment reconnaissant. Merci aussi pour tous les nombreux conseils, toujours pertinents, que tu m'as donné au cours de ces années. Tes fameux "Non mais Boris, sérieux, c'est quoi la question là ?" me manqueront et j'essaierai de m'en rappeler !

Borja, on s'est un peu moins côtoyé à cause de la distance mais chaque moment partagé avec toi était un plaisir. Merci de m'avoir accueilli par deux fois à Madrid, j'en garderai un très bon souvenir que ça soit au muséum où j'ai pu interagir un peu plus avec toi ou pendant les soirées mouvementées de Lavapiés, c'était super ! J'espère être un jour *totalmente bilingue*, j'y travaille mais ce n'est pas gagné !

J'aimerais remercier l'ensemble de l'équipe ZOBO : Jojo, Yanou, Joss, Thomas, Philipp, Borja et Christophe. Que ça soit à la Réunion, à Madrid, en Angleterre, derrière la paillasse, un écran ou encore un comptoir, c'était tout simplement super. De mes voyages à la Réunion, je garderai de nombreux souvenirs impérissables ! Et notamment de notre réveillon de Noël passé sous cette petite bâche, sous l'averse la plus longue et drue que je n'ai jamais connue. Ce riz-cassoulet-rhum sera sans conteste le repas de Noël dont je me rappellerai le plus longtemps ! Joss, nos pérégrinations dans ces foutus branchages calcinés et nos chasses aux Simulies resteront de très bons souvenirs aussi, c'était l'aventure, la vraie ! Jojo, j'aimerais aussi te remercier tout particulièrement pour ta disponibilité, ta gentillesse et les nombreux conseils que tu m'as donné pendant cette thèse, tu as toujours été très disponible, merci beaucoup ! (en espérant qu'on se retrouve bientôt autour d'une coche !) Et enfin Yanou, merci de n'avoir jamais perdu patience à cause des nombreuses questions naïves que je t'ai posé sur la génomique. Au bureau, ton enthousiasme pour les petites trouvailles quotidiennes a toujours été rafraichissant ! Philipp, tu m'as permis de mettre un pied dans le système Zosterops dès le M1, sans toi je n'aurai probablement pas fait cette thèse, merci ! Thomas, on ne s'est pas vu longtemps mais ces quelques sessions de baguage avec toi étaient géniales !

Merci également aux membres de mon comité de thèse : Pierre-André et Guillaume qui m'ont permis de rester dans le droit chemin (ou remis dedans) par deux fois. Nos discussions m'ont beaucoup apportées ! PAC, ça fait deux essais manqués, j'espère que la troisième fois sera la bonne !

Merci à Benoit Pujol pour ses nombreux conseils sur l'utilisation des outils de génétique quantitative.

Merci à Dominique Strasberg pour son accueil à la Réunion. Merci à Yann et Mickaël pour leur gentillesse et leur hospitalité, les sessions captures de mouches et de lézards avec vous étaient géniales.

Merci à Charline et Ilona qui m'ont fourni mes premières vraies expériences d'encadrement, c'était un vrai plaisir que de travailler avec vous !

Merci à Nicole, Dominique P., Linda, Louise, Caroline et Fred B. pour votre aide avec les tâches administratives dans lesquelles je n'excelle pas vraiment ! Pierre, t'as toujours su me dépatouiller de mes problèmes informatiques au labo et tu as même sauvé mon ordi perso : merci beaucoup ! Amaia, Hélène, Sophie, Céline et Yves merci beaucoup pour votre bonne humeur et votre aide au labo, c'est un plaisir d'aller maniper lorsque vous êtes dans les parages !

Merci beaucoup aux personnes qui ont rendu mes séjours à Madrid agréables : Paloma, Guillermo, Pau, David, Annie, Marta, Sofia, Jorge, Ricardo, Patrick, Joaquin et Isabel.

Je voudrais aussi remercier toutes les personnes qui ont rendu le quotidien agréable au labo : Joss, Guilhou, Zaza, Julien(s), Kiki, Robin, Jérome, Blaisou, Thieuma, Poussin, PJ, ToussTouss, Oliv', Paulo, Lucie, Kévin, Jean-Baptiste, Christophe, Pierrick, Yves, Nico(s) et Cumcum ! D'abord, les collègues de bureau : Léa, déjà 6 ans qu'on se suit à la fac (même si tu as oublié la première année alors que tu étais certainement ma meilleure amie, snif), merci pour tes petites attentions et ta bonne humeur (presque) constante, merci aussi de n'avoir jamais été désagréable même quand tu étais énervée. Ces 3 ans au bureau sont passés à toute vitesse et c'est en grande partie grâce à toi ! Courage, c'est bientôt la fin ! Luc, merci pour ta gentillesse et ta générosité, derrière ton air de guerrier et tes menaces à la machette se cache un petit être sensible ! Arthur, tout va toujours bien avec toi, t'es une « vraie pate », tu rends le quotidien agréable c'était vraiment un plaisir (je ne sais pas à qui je vais piquer des stylos

maintenant). Jojo, Yanou, Léa, Luc et Arthur : j'en ai peur, ça va être très dur de faire aussi bien dans mon prochain bureau !! J'espère que vous me garderez en tant que membre d'honneur du CBur. Ensuite, un merci particulier à Thieuma sans qui les soirées au labo auraient été bien longues et parfois déprimantes. J'aurai sûrement dû mal à retrouver un compère de rugby et de ventriglisse comme toi ! Merci d'avoir toujours su être là mon meilleur copain de Toulouse, on va maintenant pouvoir reprendre ta formation ornitho et ma formation drop ! Kiki, je garderai un souvenir impérissable de tes bus géants de fin de thèse, merci beaucoup de m'avoir aidé à répondre à toutes les questions existentielles que je me suis posé pendant ces trois ans ! Olivia, merci pour ta bonne humeur constante au labo et en dehors ! Blaisou, t'as alimenté mes écouteurs ces deux dernières années, j'espère qu'on continuera à échanger autant ! Pierrick, ça fait du bien d'avoir des gens comme toi dans son entourage, j'espère que je serai un jour un chercheur comme toi, t'es un vrai modèle !

Merci aux autres camarades Toulousains qui m'ont permis de décompresser soirs et weekends : Ju, Toinou, Banban, Jérem, Charlotte !

Merci aussi à tous les copains d'ornitho pour les bons moments partagés et l'épisode décompression annuel à Ouessant. Vous avez su rester sages cette année et je vous en suis reconnaissant. Il va maintenant falloir penser à réserver pour 2015 !

Merci aussi à mes 4 fidèles compagnons d'ornitho et de bien d'autres choses : Coco, Louis, JC et Keke : je ne sais pas ce que je ferai sans vous ! Coco, on a probablement passé plusieurs centaines d'heures au téléphone pendant cette thèse, t'as toujours été à l'écoute de mes petits problèmes, t'es mon confident ! Merci pour tout ! Louis, depuis mon premier Ouessant en passant par nos nuits communes de travail en M2 jusqu'à aujourd'hui, t'as toujours su être là dans toutes les situations, merci ! On va maintenant pouvoir se consacrer à notre groupe de jazz ! Coco, Louis, à nous les gros burgers !! Keke, t'es sûrement le plus mature de la bande, c'est toujours ressourçant et apaisant de passer des moments avec toi, merci ! J'espère qu'on passera un peu plus de temps ensemble dans les années qui viennent ! Jean-Charles, mon garde du corps, merci d'être toujours là en cas de besoin et pour ton enthousiasme à toute épreuve ! Bientôt, à nous les soirées en Camargue ! Et fais gaffe à toi, car maintenant que j'ai fini tu sais ce qui t'attend !

Merci à ma Zaza qui m'aura soutenu jusqu'au bout sans jamais se plaindre de mon stress et de mes habitudes de toqué, tu vas pouvoir souffler et on va pouvoir profiter un peu de nous

maintenant ! Sans toi, cette dernière année de thèse (et pas que) aurait été bien moins simple agréable ! Tu es parfaite, merci pour tout !

Enfin, j'aimerai remercier mes parents sans qui rien n'aurait été possible. Merci de m'avoir soutenu tout au long de mes études sans jamais faillir, de m'avoir toujours laissé libre de mes choix et de m'avoir aidé à réaliser toutes mes envies ! Merci beaucoup ! J'aimerai aussi remercier mon grand père qui y est pour beaucoup dans ma formation naturaliste, c'est en grande partie grâce à toi si j'ai suivi cette voie !

Table des matières

Introduction

La compréhension de l'origine et du maintien de la diversité biologique est l'un des objectifs centraux de l'étude de l'évolution. Les travaux d'Ernst Mayr, dès les années 1940, ont fortement contribué à l'émergence d'un paradigme reposant sur le contexte géographique pour comprendre la spéciation (Mayr 1942). Ainsi, pendant de nombreuses années, seule la spéciation en situation d'isolement géographique complet (*i.e.* allopatrie) était reconnue comme vraisemblable (Mayr 1942, 1963). Elle est particulièrement bien illustrée par les événements de vicariance (*i.e.* apparition d'une barrière géographique au sein de l'aire de répartition d'une espèce) tel que celui provoqué par la fermeture de l'isthme du Panama. Celle-ci a vraisemblablement entrainé la différenciation d'un grand nombre de taxons en interrompant le flux de gènes entre les océans Atlantique et Pacifique. Ainsi, de nombreuses espèces sœurs, des poissons aux échinidés, sont présentes de part et d'autre de cette barrière (Lessios 1998). Si la spéciation a longtemps été étudiée essentiellement en situation d'isolement géographique (Mayr 1942, 1963), il existe cependant de nombreuses situations *in natura* où des populations phénotypiquement divergentes sont observées dans des aires de répartition continues. Ces différences sont regroupées sous le terme de « cline » et désignent les variations spatiales des fréquences alléliques et des traits phénotypiques (*sensu lato* ; Barton & Hewitt 1985). Ces variations peuvent être graduelles (*e.g.* Antoniazza *et al.* 2010) ou abruptes (*e.g.* zones hybrides; *e.g.* Singhal & Moritz 2012). Endler (1977) fut l'un des premiers à souligner la forte fréquence de ces situations dans la nature et à tenter d'expliquer comment des populations en contact et échangeant des gènes, pouvaient présenter et maintenir des différences phénotypiques marquées. A l'encontre de ce que pensaient Mayr (1942, 1963) et Dobzhansky (1970), Endler souligna l'intérêt de ces situations pour étudier les processus de divergence et de spéciation : de tels cas d'étude permettraient de mettre en évidence le rôle de la sélection naturelle pour expliquer la divergence en l'absence de barrière géographique et de mieux comprendre les processus d'isolement reproductif. Ce n'est ensuite que dans les années 1990 que l'importance de la sélection dans l'émergence des différences phénotypiques et de l'isolement reproductif face aux flux de gènes a réellement été reconnue (Mallet 2001). De manière intéressante, ce changement de paradigme tend à faire renaître les idées de Charles Darwin (1859). Darwin soulignait déjà que la divergence en présence de flux de gènes était importante et vraisemblable : « Although I do not doubt that [geographical] isolation is of considerable importance in the production of new species, on the whole I am inclined to believe that largeness of area is of more importance. » (p. 104-106). Il soulignait également l'intérêt des formes intermédiaires : « Those forms which possess in some considerable degree the character of species, but which are so closely similar to some other

forms, or are so closely linked to them by intermediate gradations, that naturalists do not like to rank them as distinct species, are in several respects the most important to us » (p. 57).

De nombreux progrès ont été réalisés dans la compréhension des mécanismes de divergence et de mise en place de l'isolement reproductif. La divergence en présence de flux de gènes sous les effets de la sélection est aujourd'hui admise comme possible et probablement assez courante (Rice & Hostert 1993; Pinho & Hey 2010). Il a de plus été montré que l'adaptation locale pouvait réduire substantiellement les flux de gènes en l'absence de barrière géographique et donc potentiellement promouvoir la spéciation (Nosil *et al.* 2005; Nosil 2008). Cependant, malgré de multiples progrès dans la compréhension de la divergence micro-évolutive, son lien avec la spéciation reste à consolider et de nombreux verrous scientifiques sont encore en places. A quel point la sélection naturelle et les réponses adaptatives des organismes permettent le maintien des différences phénotypiques en dépit des flux de gènes (Doebeli *et al.* 2005) ? Quelle est la fréquence des événements de divergence engendrés par la sélection divergente seule par rapport à celle d'évènements initiés par une phase d'isolement géographique des populations (Pinho & Hey 2010) ? Quelle est la fréquence des événements de différenciation provoqués par des processus non-adaptatifs par rapport à celle d'évènements provoqués par des processus adaptatifs (Nei 2005) ?

Répondre à de telles questions représente un réel défi scientifique, surtout lorsque l'on travaille sur des espèces non-modèles chez lesquelles il est impossible de réaliser des travaux expérimentaux. Dans cette introduction générale, je commencerai par décrire comment la divergence et l'isolement reproductif peuvent se mettre en place entre différentes populations (section 1), puis je me focaliserai sur les causes du maintien de la différenciation au sein des zones de transition entre taxons génétiquement distincts (dites « zones hybrides »), les conséquences évolutives de l'hybridation et leurs rôles dans notre compréhension de la spéciation (section 2). Dans ces deux sections, il ne s'agit pas de dresser un tableau exhaustif des mécanismes à l'origine de l'émergence et du maintien de la biodiversité, seules les situations où les populations divergentes sont au contact les unes des autres seront abordées. Le modèle d'étude, les données collectées ainsi que leurs adéquations avec la problématique seront présentés dans la section 3. Enfin je décrirai les objectifs visés par mon travail de thèse (section 4).

1 - Emergence de la diversité

1.1 - Influence des facteurs neutres, sélectifs et historiques

1.1.1 - Différenciation adaptative

Les variations phénotypiques entre populations distribuées de manière continue sont la plupart du temps interprétées comme le résultat d'adaptation différentielle à l'environnement naturel ou sexuel des individus (Via 2001, 2002; Kirkpatrick & Ravigné 2002; Gavrilets 2003). L'environnement est rarement constant dans l'espace et le temps, on s'attend donc à ce que les populations qui vivent dans ces environnement variables s'adaptent localement aux différentes conditions biotiques et abiotiques qu'elles subissent (Kawecki & Ebert 2004; Leinonen *et al.* 2008). L'adaptation locale aux variations spatiales de la sélection a largement été étudiée et de nombreux agents de sélection ont été mis en évidence. Par exemple, les pressions de sélection induites par le climat, dont la température, semblent être un moteur fort de différenciation, notamment sur les gradients altitudinaux où elles induisent, entre autres, des variations physiologiques et morphologiques (*e.g.* Keller *et al.* 2013). Les pressions induites par les interactions biotiques peuvent également jouer un rôle fort dans la différenciation des populations via la compétition (*e.g.* Grøndahl & Ehlers 2008) ou la prédation (*e.g.*Hoekstra *et al.* 2004; Vignieri *et al.* 2010). Par exemple, les souris du genre *Peromyscus* présentent des variations de la couleur du pelage en fonction du substrat sur lequel elles vivent. Ces variations leur confèrent une meilleure aptitude à la survie en les rendant cryptiques vis-à-vis des prédateurs (Vignieri *et al.* 2010). Enfin, la sélection liée à l'environnement peut également agir sur les signaux de communication servant au choix de partenaire. L'hypothèse du « sensory drive » stipule que les systèmes de communication utilisés pour le choix de partenaire doivent être adaptés à l'environnement local afin de maximiser la transmission des signaux de communication (Boughman 2002). Il a été montré dans plusieurs systèmes d'études que le choix de partenaire était dépendant de l'habitat et favorisait la divergence phénotypique (*e.g.* Uy & Borgia 2000; Seehausen *et al.* 2008; Secondi *et al.* 2014). Dans toutes ces situations, l'écologie est le moteur de la divergence en provoquant des pressions de sélection qui varient spatialement.

Cependant, les effets de la sélection divergente vont dépendre des autres forces évolutives agissant sur les systèmes, principalement du flux de gènes (Kawecki & Ebert 2004). Les échanges d'individus adaptés à des environnements différents tendent à homogénéiser les populations et à éloigner les populations du phénotype optimal localement

(Räsänen & Hendry 2008). La condition pour que les populations puissent se différencier est donc que l'effet diversifiant de la sélection soit assez fort pour s'opposer à l'effet homogénéisateur des flux de gènes (Lenormand 2002; Pinho & Hey 2010; Edelaar & Bolnick 2012). Néanmoins, il ne faut pas interpréter toutes les variations phénotypiques comme des signatures de sélection. En effet, pour que la sélection agisse, le trait doit avoir à la fois une base génétique (*i.e.* être héritable) et un impact sur la fitness des individus. Si le trait n'a pas de base génétique, la variation peut être le résultat de la plasticité phénotypique. La plasticité phénotypique est définie comme l'expression différentielle du phénotype pour un même génotype en fonction de l'environnement dans lequel il se trouve (Via & Lande 1985). Les exemples de plasticité phénotypiques sont nombreux et ont été étudiés tant chez les plantes que chez les animaux (Miner *et al.* 2005).

1.1.2 - Différentiation non-adaptative

La sélection est loin d'être le seul processus permettant la différenciation des populations. Les processus neutres tels que la mutation et la dérive génétique peuvent aussi jouer un rôle. Cependant, la part relative de la différentiation génétique et phénotypique liée aux processus neutres est une question encore débattue (Whitlock & Phillips 2000; O'Hara 2005; Lynch 2007). En particulier, quand la variation phénotypique est graduelle, les différences phénotypiques peuvent être le résultat de différents processus neutres. D'une part, sous l'effet de flux de gènes limités entre les populations, on peut s'attendre à ce que la dérive génétique agisse au sein des populations en provoquant un pattern d'isolement par la distance (Endler 1977). Sous cette hypothèse, plus les populations sont espacées géographiquement, plus elles sont différentes d'un point de vue génétique et phénotypique (Wright 1943). Si les tailles efficaces de populations sont faibles, d'une part les populations répondent plus lentement à la sélection et d'autre part le taux de dérive génétique augmente (Frankham & Weber 2000; England *et al.* 2003). Bien que la divergence due aux processus neutres soit facilement appréhendable d'un point de vue théorique (Endler 1977), les exemples empiriques semblent assez peu nombreux (Leinonen *et al.* 2008). Néanmoins, selon Leinonen *et al.* (2008), ce résultat pourrait venir d'un biais de publication vers les études montrant que la sélection naturelle a eu un rôle dans la différentiation. D'autre part, les expansions d'aire de répartition peuvent aussi former des clines le long de leur route de colonisation (Excoffier & Ray 2008). Les effets fondateurs successifs ayant lieu sur le front de colonisation (e.g. Currat

& Excoffier 2005) peuvent provoquer la formation de patrons de variations clinaux sous l'effet de phénomènes de « surfing » d'allèles (Edmonds *et al.* 2004; Klopfstein *et al.* 2006). Dans cette situation, les allèles émergeant de nouvelles mutations ou faisant déjà partis du bagage génétique de l'espèce surfent sur la vague d'expansion.

Enfin, dans la nature il est courant que la sélection naturelle et la dérive génétique agissent de concert (van Oosterhout *et al.* 2006; Clegg 2010). Par exemple, la sélection naturelle en réduisant les flux de gènes entre les différentes populations peut faciliter l'action de la dérive génétique (Nosil *et al.* 2009).

1.1.3 - Contexte temporel

La distribution des populations dans l'espace n'est pas figée au cours du temps. Dans de nombreux cas, des populations génétiquement et phénotypiquement divergentes, aujourd'hui en contact, ont divergées en allopatrie (Price 2008). Selon Barton *et al.* (2007), la grande majorité des espèces a subi, au moins une fois durant sa divergence, une phase d'allopatrie. Lors de ces phénomènes, la modification de leurs aires de répartition conduit à une remise en contact de populations divergentes, appelée contact secondaire. Un contact secondaire peut être dû à des expansions d'aire, des changements d'habitat ou des introductions liées à l'action de l'homme (*e.g.* Seehausen *et al.* 1997; Goodman *et al.* 1999; Abbott & Comes 2004; Zinner *et al.* 2009). De plus, même si les modèles théoriques de divergence avec flux de gènes paraissent réalistes, il apparaît qu'une phase d'allopatrie favoriserait grandement l'initiation de cette divergence. Pinho et Hey (2010) soulignent d'ailleurs que la fréquence des situations où les populations divergentes aujourd'hui en contact n'ont jamais été séparées géographiquement n'est pas connue. Les processus qui provoquent la divergence en allopatrie sont les mêmes que ceux qui agissent lorsque les populations sont en contact (décrits dans les sections 1.1.1 et 1.1.2). Cependant, l'action de la sélection divergente et de la dérive génétique est favorisée par l'absence de flux de gènes entre les différentes populations (Nosil *et al.* 2009). Les scénarios de contact secondaire peuvent produire des situations où les variations génétiques et phénotypiques sont clinales, sous l'effet de la diffusion neutre entre deux populations auparavant homogènes ; les différences tendent à s'homogénéiser sous l'effet de la reproduction. Si la sélection divergente continue d'agir après la remise en contact, il peut y avoir un maintien des différences et une zone de transition abrupte entre les populations en contact.

1.2 - Divergence et isolement reproductif

L'isolement reproductif (*i.e.* les capacités de populations génétiquement distinctes à se reproduire au même endroit sans mélanger leurs génomes) apparait comme un concept central pour l'étude de la spéciation. Il a d'ailleurs été utilisé pour la définition du concept biologique de l'espèce (Mayr 1970), où les espèces sont définies en tant que groupes d'individus qui se reproduisent librement entre eux mais qui sont reproductivement isolés des individus d'autres espèces. Cette définition a largement été critiquée ces dernières années. Si de nombreux concepts d'espèce ont vu le jour depuis les travaux d'Ersnt Mayr (voir de Queiroz 2007 pour une revue détaillée), d'un point de vue évolutif, la notion d'évolution de l'isolement reproductif apparaît être adaptée pour l'étude des processus de formations d'espèces (Coyne & Orr 2004; Feder *et al.* 2013).

Le rôle de la sélection dans l'évolution de l'isolement reproductif apparait important : il a été montré théoriquement et empiriquement que la plupart des barrières à la reproduction apparaissaient sous l'action de la sélection naturelle ou sexuelle (Schluter 2009; Safran *et al.* 2013). Il semble que l'évolution de l'isolement reproductif soit favorisée lorsque les populations ne sont pas en contact. En effet, en allopatrie, la fixation de différences sous l'effet de la sélection n'est pas ralentie par les flux de gènes (Turelli *et al.* 2001; Calcagno 2007; Feder *et al.* 2013). D'une part, les changements adaptatifs, en provoquant la fixation d'allèles favorables localement et les modifications phénotypiques subséquentes, entraineraient la formation de barrières à la reproduction (Schluter 2009; Sobel *et al.* 2010). D'autre part, dans un contexte de sélection uniforme, différentes mutations peuvent se fixer dans des populations séparées géographiquement et pourtant produire des phénotypes similaires. Dans ce cas, les populations vont fixer des mutations adaptatives différentes à un même locus ou sur des loci différents et dans un ordre différent, cette hypothèse est appelé « mutation order speciation » (Schluter 2009; Nosil & Flaxman 2011). Théoriquement et expérimentalement, il a été montré que ce type de processus pouvait conduire à l'évolution d'un isolement reproductif (Blount *et al.* 2008; Anderson & Harmon 2014). Cependant, l'importance de ce phénomène en conditions naturelles reste peu connue (Nosil & Flaxman 2011). Enfin, l'évolution de l'isolement reproductif par dérive génétique est facilement appréhendable d'un point de vue théorique, mais les exemples empiriques manquent. Selon plusieurs auteurs, la spéciation sous l'action unique de la dérive ne serait pas vraisemblable

(Turelli *et al.* 2001; Coyne & Orr 2004). De plus, des expériences réalisées en laboratoire n'ont pas réussi à faire évoluer l'isolement reproductif sous l'effet de la dérive génétique seule (*e.g.* Rundle 2003).

Les mécanismes d'isolement reproductif peuvent être classés en deux catégories. Premièrement, si les barrières à la reproduction interviennent avant la fécondation, elles sont qualifiées de pré-zygotiques. Ces barrières peuvent être de différentes natures (Fig. 1) :

- écologiques : les populations parentales ne partagent pas les mêmes *preferendum* d'habitat (*e.g.* Ramsey *et al.* 2003) ou les mêmes périodes de reproduction (*i.e.* différences phénologiques ; (Martin & Willis 2007; Ruegg *et al.* 2012). Dans ces situations, les individus n'ont simplement pas l'opportunité de se rencontrer pendant leurs périodes de reproduction ;
- comportementales : les partenaires préfèrent un partenaire leur ressemblant (homogamie) et/ou choisissent leurs partenaires en fonction d'un critère espèce-spécifique tel que le comportement de parade nuptiale (Seehausen *et al.* 1997; Doi *et al.* 2001; McKinnon *et al.* 2004) ;
- morphologiques : les organes reproducteurs présentent des incompatibilités physionomiques et empêchent la copulation (*e.g.* Sota & Vogler 2001) ;
- gamétiques : dans ce cas la copulation a lieu mais la compétition entre gamètes mâles peut n'être favorable qu'aux gamètes similaires aux gamètes femelles. (Bierne *et al.* 2002; Ramsey *et al.* 2003).

Toutes ces barrières représentent un frein à la formation des zygotes en empêchant la fécondation.

Deuxièmement, les barrières à la reproduction peuvent être post-zygotiques. Dans ce cas, la formation des zygotes est possible mais les descendants présentent une valeur sélective (ou fitness) moindre que les phénotypes parentaux (survie ou fécondité réduite). Ces barrières post-zygotiques peuvent être dues à de la sélection exogène ou endogène (Barton & Hewitt 1985). La sélection exogène (parfois appelée extrinsèque) provoque la maladaptation des hybrides aux conditions environnementales locales, qu'elles soient biotiques ou abiotiques (Szymura & Barton 1986; Rice & Hostert 1993; Schluter 2001). La sélection endogène (ou intrinsèque) est, elle, indépendante des conditions environnementales. Elle dépend seulement des caractéristiques intrinsèques des populations. Dans ce cas, on observe généralement des incompatibilités du type « Dobzhansky-Muller » (interactions épistatiques) affectant la valeur sélective de potentiels hybrides (Dobzhansky 1937; Muller 1942; Rieseberg 2001). Dans les

cas d'interactions épistatiques, la plus faible valeur sélective des hybrides résulte de la rupture des liaisons d'un groupe de gènes coadaptés à l'intérieur des génomes parentaux sous l'effet de la recombinaison (Coyne & Orr 2004). Les remaniements chromosomiques peuvent aussi être à l'origine d'un isolement post-reproductif car ils entrainent la stérilité des hybrides (généralement due à des dysfonctionnements de la méiose ; Noor *et al.* 2001). Dans certains cas, on considère aussi l'homogamie comme une barrière post-zygotique à la reproduction. En effet, si les individus intermédiaires ne trouvent pas de partenaire à cause des préférences pour les phénotypes parentaux, on parle de stérilité comportementale (Coyne & Orr 2004).

Les mécanismes d'isolement pré- et post-zygotiques ne sont pas mutuellement exclusifs et il existe de nombreux systèmes où les deux types de mécanismes agissent (*e.g.* Mallet 1989; Bierne *et al.* 2003). Dans les cas où les barrières post-zygotiques sont fortes, il est aussi possible que l'apparition de barrières pré-zygotiques à la reproduction soit favorisée (hypothèse du renforcement ; Servedio & Noor 2003). Bien que difficile à mettre en évidence de façon empirique, cette hypothèse a été démontrée chez plusieurs taxa (*e.g.* Saetre *et al.* 1997; Bímová *et al.* 2011) et pourrait jouer un rôle dans de nombreux systèmes (Servedio & Noor 2003).

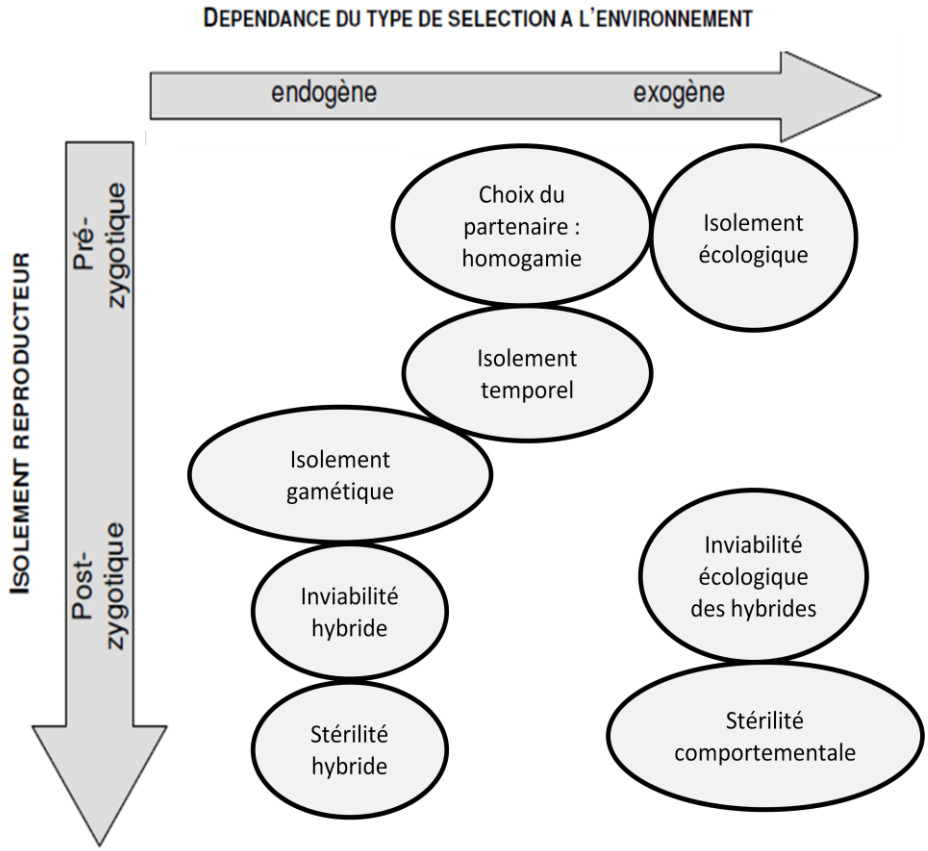


Figure 1 : Mécanismes d'isolement reproducteur classés en fonction du moment où ils agissent (pré- ou post-zygotique) et de leur dépendance à l'environnement. Adapté de Bierne (2001) dans Gay (2006).

2 - Zones hybrides

2.1 - Formation et maintien

Lorsque l'isolement reproductif entre deux taxa est incomplet, la reproduction entre ces taxa est considérée comme de l'hybridation. Au sens large, l'hybridation est définie comme la reproduction entre des individus de deux populations ou de deux groupes de populations distinctes, les individus de ces populations étant distinguables sur la base d'un ou de plusieurs caractères héritables (Harrison 1990, 1993). L'hybridation se produit dans tous les processus de spéciation sauf dans le cas d'une différenciation totalement allopatrique (Abbott *et al.* 2013). Elle est particulièrement remarquable, lorsqu'il y a formation d'une zone

hybride. Une zone hybride est définie comme une zone géographique où deux populations génétiquement distinctes se rencontrent, se reproduisent et produisent une descendance (Harrison 1993).

L'étude de la spéciation est difficile une fois que celle ci est achevée, de la même façon qu'elle est difficile lorsque il n'y a aucun isolement reproductif. Dans les zones hybrides, où l'isolement reproductif est partiel, de nombreux individus au phénotype intermédiaire comparativement à ceux des taxons parentaux sont observés. Cette variabilité permet de mieux appréhender la nature et la mise en place des barrières à la reproduction existantes. Les zones hybrides ont ainsi été qualifiées de « laboratoires naturels » (Hewitt 1988) ou encore de « fenêtres sur les processus évolutifs » (Harrison 1990) au travers desquelles il est possible d'observer les processus de divergence en cours. L'étude des zones hybrides permet d'identifier les gènes et les traits sous sélection divergente ou encore les régions génomiques responsables de l'isolement reproductif (Barton & Hewitt 1985; Harrison 1993; cf. section 2.2).

Les zones hybrides peuvent se former de deux manières principales. Soit il n'y a pas de phase de séparation géographique et les deux taxa diffèrent sous l'effet de la sélection divergente liée à l'environnement (*i.e.* différenciation primaire). Soit la différenciation se fait d'abord lors d'une phase d'allopatrie puis la formation de la zone hybride est liée à un contact secondaire. Même s'il est difficile de distinguer entre les situations de différenciation primaire et de contact secondaire sans connaître précisément l'histoire des populations (Endler 1977; Nosil 2008), il semble néanmoins que le scénario de contact secondaire soit le plus vraisemblable pour expliquer la formation de la majorité des zones hybrides (Hewitt 2000; Barton *et al.* 2007; Price 2008).

L'un des aspects marquants de l'étude des zones hybrides est le maintien de celles-ci pendant des temps très longs (plusieurs milliers de générations ; Harrison 1993). Plusieurs hypothèses existent pour l'expliquer. La première catégorie d'hypothèses est dispersion-dépendante : le maintien de la zone hybride est expliqué par un équilibre entre l'effet homogénéisant des flux de gènes et l'effet diversifiant de la sélection. Au sein de cette classe d'hypothèses, on distingue parfois les zones maintenues par la sélection exogène (« environmental gradient model »; May *et al.* 1975; Endler 1977) de celles maintenues par la sélection endogène (« tension zone model »; Key 1968; Slatkin 1973; Barton & Hewitt 1985). Ces zones sont parfois regroupées sous le terme de « zone de tension » sans que l'on

distingue les deux types de sélection (Barton & Hewitt 1985). Les zones hybrides suivant le modèle dispersion-dépendant peuvent être observées dans différentes configurations géographiques. Les zones de tension maintenues par la sélection endogène sont connue pour se déplacer géographiquement jusqu'à coïncider avec des zones où les densités de populations sont faibles ou avec des zones de transition environnementale (Barton & Hewitt 1985; Bierne *et al.* 2011; cf. Encadré 1). Ces zones hybrides apparaissent donc généralement au sein d'environnements homogènes sur une barrière physique (faible densité de populations) ou sur des écotones (transition environnementale entre deux habitats).

La deuxième catégorie d'hypothèse pour expliquer le maintien des zones hybrides est dispersion-indépendante, elle est appelée hypothèse de supériorité hybride (Moore 1977). Sous cette hypothèse, les phénotypes hybrides sont favorisés par la sélection exogène. Ces zones hybrides sont donc placées au niveau d'écotone et la supériorité des hybrides est bornée au centre de la zone. Ce modèle suppose une dispersion très faible et de forts changements environnementaux. Selon Barton & Hewitt (Barton & Hewitt 1985), les zones hybrides de type « supériorité hybride » sont moins communes que les zones de tension.

2.2 - Conséquences évolutives

Marquées par le concept biologique d'espèce énoncé par Ersnt Mayr, les espèces ont longtemps été pensées comme des entités homogènes. Ainsi, l'hybridation était incompatible avec la notion d'espèce « vraie ». Les espèces étaient définies comme des entités génétiques homogènes qui n'échangeaient pas de gènes avec les populations d'autres espèces. Sous l'effet des flux de gènes entre taxa différenciés, les zones hybrides constituent cependant des zones d'introgression génétique (*i.e.* transfert de matériel génétique d'un taxa à un autre sous l'effet d'événements d'hybridation récurrents ; Mallet 2005)). Sous l'effet de l'hybridation et donc de la recombinaison pendant un grand nombre de générations, les allèles d'une espèce vont diffuser dans le génome de l'autre espèce. Key (1968), après ses travaux sur les zones hybrides, fut l'un des premiers à avancer que les taux d'introgression variaient au sein du génome.

La notion de semi-perméabilité des limites entre espèces (Key 1968; Harrison 1990) a ainsi émergé. Cette notion découle du fait que certaines parties du génome sont échangées librement lors de l'hybridation tandis que d'autres ne sont pas ou peu échangées. La notion

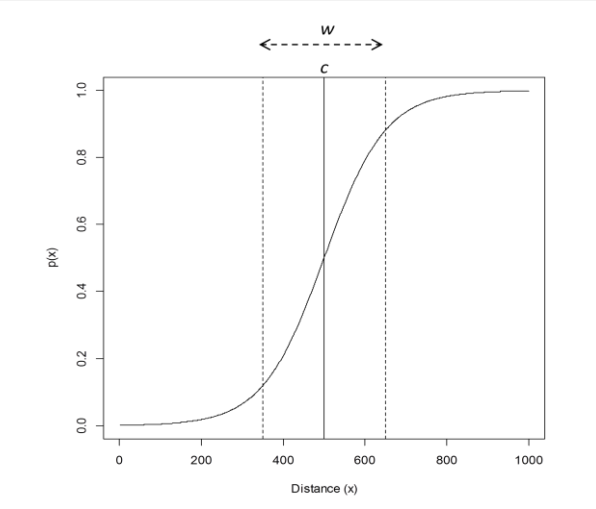
d'isolement génétique doit alors être considérée comme une caractéristique des marqueurs génétique et non du génome en entier (Harrison 1990). Barton et Hewitt (1985, 1989) formalisèrent mathématiquement les raisons de l'hétérogénéité des taux d'introgression au sein du génome, en utilisant des modèles de diffusion issus de la physique (théorie des clines; Encadré 1). Lorsqu'il n'y a pas de contre-sélection des hybrides (*i.e.* absence d'isolement reproductif), les flux de gènes affectent l'ensemble du génome de la même façon. Au contraire, lorsque la sélection (endogène ou exogène) agit contre les hybrides, les flux de gènes sont freinés et variables au sein du génome. Ce sont donc les effets indirects de la sélection qui font varier les taux d'introgression au sein du génome. Ainsi, les allèles des loci impliqués dans l'isolement reproductif ne diffusent pas ou peu au travers des zones hybrides. Les allèles neutres présentent quant à eux des taux d'introgression variables en fonction des liaisons qu'ils possèdent avec des gènes d'adaptation locale (sélection exogène) ou d'isolement reproductif (*e.g.* groupes de gènes coadaptés) (Barton 1979; Barton & Hewitt 1985). Pour diffuser entre deux taxa au travers d'une zone hybride, un allèle neutre spécifique doit recombiner dans l'environnement génétique de l'autre taxon avant que la sélection contre les allèles avec lesquels il était initialement associé ne l'élimine (Barton & Hewitt 1985; Encadré 1).

Encadré 1 : Théorie des clines

Un cline est défini comme un changement de fréquence allélique ou d'un caractère phénotypique avec la distance géographique (Endler 1977). Dans les zones hybrides, les clines sont généralement bien décrit par des courbes de forme sigmoïdes. Bazykin (1969) s'inspira des équations de diffusion-convection utilisées en physique pour décrire le mouvement des particules (*e.g.* gradient thermique) pour mettre au point un modèle à un locus sous sélection. Dans son modèle deux forces s'opposent la migration et la contre-sélection des hybrides. Nicholas Barton (*e.g.* Barton & Hewitt 1985; Szymura & Barton 1986; Barton & Bengtsson 1986) a ensuite développé une série de modèles multi-locus afin de mieux caractériser les différences inter-locus en y introduisant du déséquilibre de liaison. Deux types principaux d'équations sont aujourd'hui utilisés pour décrire les variations au sein d'une zone hybride.

- **logit :** le cline de fréquence (*p*) prend la forme d'une courbe sigmoïde en fonction de la distance (*x*) et est décrit par son centre *c* et sa largeur *w* (inverse de la pente maximale).

$$p_x = \frac{1}{1 + e^{\frac{4}{w}(x-c)}}$$

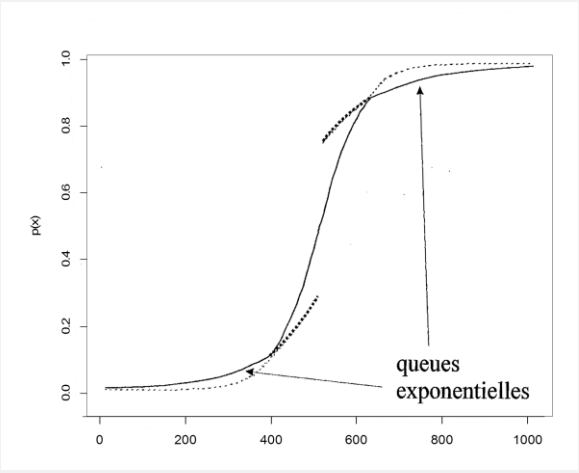


- **cline en trois parties :** dans cette situation, il y a une barrière centrale au flux de gènes créées par *n* locus sous sélection, qui rend le cline discontinu. Il y a ainsi une partie centrale sigmoïde qui se modélise de la même façon qu'un cline logit. De chaque côté, une queue d'introgression exponentielle est ajoutée. Il y a donc 4 paramètres en plus dans le modèle : d_1 et d_2 qui décrivant la position des limites entre la partie centrale et les queues d'introgression, t_1 et t_2 décrivant la pente de chaque queue.

$$\text{si } x \leq c - d_2 \; ; \; p_x = 1 - \left(\frac{1}{1+e^{-wd_2}}\right) e^{\frac{-wt_2(x-c+d_2)}{1+e^{wd_2}}}$$

$$\text{si } c - d_2 < x < c + d_1 \; ; \; p_x = \frac{1}{1 + e^{\frac{4}{w}(x-c)}}$$

$$\text{si } x \geq c - d_1 \; ; \; p_x = \frac{1}{1+e^{-wd_1}} e^{\frac{wt_1(x-c-d_1)}{1+e^{wd_1}}}$$



Tiré de Raufaste, 2001.

Une fois le « meilleur » modèle déterminé, la stratégie d'analyse consiste ensuite à comparer les paramètres d'intérêt entre différents loci ou différents traits. Lorsque deux clines ont la même pente, ils sont dits concordants. S'ils partagent le même centre ils sont alors coïncidents. Le centre nous renseigne sur la position de la zone hybride. Si la zone hybride résulte du contact secondaire et que la migration est isotrope alors la position sera l'endroit du contact initial. Au contraire, si la migration et/ou la densité sont variables dans l'espace, alors la zone hybride se déplacera jusqu'à coïncider avec une zone où la migration et/ou la densité sont faibles (Barton 1979; Barton & Hewitt 1985). La largeur nous renseigne quant à elle sur la sélection globale qui s'exerce sur le génome : plus la largeur est faible plus la sélection est forte. Les caractéristiques des queues d'introgression permettent d'estimer le nombre de locus sous sélection.

Dans bon nombre de cas, on ne peut pas utiliser les modèles de clines en trois parties car il faut avoir un bon échantillonnage à la fois du centre de la zone et de la zone d'introgression. Dans ce cas, il faut se contenter d'utiliser des modèles de type logit ce qui rend l'estimation du nombre de locus sous sélection impossible.

Ainsi c'est le nombre et la position des locus d'adaptation locale et d'isolement reproductif qui détermine l'importance des barrières au flux de gènes (Barton & Hewitt 1985). Plus les locus d'isolement sont nombreux et bien répartis dans le génome plus les allèles neutres ont de chances d'y être liés et de ne pas diffuser au travers de la zone hybride. Au contraire, si peu de gènes sont impliqués dans l'isolement reproductif ou s'ils sont regroupés sur un chromosome alors la majorité du génome neutre s'homogénéisera entre les deux taxons. Il existe aussi des phénomènes d'introgression adaptative : lorsque les allèles sont avantageux pour les deux taxa, leurs taux d'introgression tendra à être plus élevé que le niveau d'introgression basal (*e.g.* Pardo-Diaz *et al.* 2012; Fraïsse *et al.* 2014). Cette hétérogénéité des taux d'introgression est particulièrement bien illustrée par la comparaison de marqueurs génétiques neutres aux variations phénotypiques. Il existe des situations très variées, depuis des zones hybrides où l'introgression est freinée sur la majorité du génome (*e.g.* Baldassarre *et al.* 2014) jusqu'à des zones où les génomes neutres semblent totalement homogénéisés alors que les transitions phénotypiques sont très marquées (*e.g.* Payseur *et al.* 2004; Poelstra *et al.* 2014). Grâce à l'avènement des outils de séquençage haut-débit, les exemples montrant que les taux d'introgression sont variables au sein du génome deviennent de plus en plus nombreux. Par exemple, les différences observées dans une zone entre deux espèces de Corneille (*Corvus corone* et *C. cornix*) sont restreintes à une seule partie du génome (Poelstra *et al.* 2014). Au contraire, les barrières sont réparties sur l'ensemble du génome chez deux sous espèces du Mérion à dos rouge (*Malurus melanocephalus* ; Baldassarre *et al.* 2014). Ces études, en analysant les patterns d'introgression au sein des

génomes, permettent d'identifier les régions génomiques potentiellement impliqués dans l'adaptation locale et la spéciation et d'examiner l'architecture génomique des limites entre espèces (Payseur 2010). L'ensemble de la littérature sur les zones hybrides a ainsi amené certains auteurs à adopter le concept génique d'espèce (Wu 2001) qui met l'accent sur le fait que le flux de gènes est une caractéristique de régions génomiques et non pas des génomes entiers. L'analyse d'un grand nombre de zones hybrides variées, qui représentent autant d'étapes du processus de spéciation, a permis de mieux comprendre comment les barrières à la reproduction se mettent en place (Harrison & Larson 2014).

3 - Système d'étude

3.1 - Le Zostérops des Mascareignes *Zosterops borbonicus*

Le genre *Zosterops*, appartenant à l'ordre des passériformes, est le genre d'oiseaux le plus diversifié au monde avec 87 espèces et plus de 200 sous-espèces reconnues (Gill & Donsker 2014). Le genre possède une vaste aire de répartition puisqu'il occupe îles et continents depuis l'Ouest de l'Afrique subsaharienne jusqu'au milieu de l'Océan pacifique à l'Est. Il a colonisé un grand nombre d'îles de l'océan Indien à l'océan Pacifique et près de la moitié des espèces de *Zosterops* (46%) sont endémiques d'une seule île (Warren *et al.* 2006; Melo *et al.* 2011). Les raisons de cette incroyable diversité semblent être liées à la fois à de remarquables capacités de dispersion à longues distances (Diamond 1974; Glor 2011) et à une grande propension à perdre ces capacités de dispersion une fois une île colonisée (Moyle *et al.* 2009). Au cours de ces deux derniers millions d'années (Warren *et al.* 2006; Melo *et al.* 2011), l'installation sur de nombreuses îles suivi d'évènements de spéciation allopatriques pourraient donc être à l'origine de cette diversité remarquable.

Parmi ces espèces, quatre sont endémiques de l'archipel des Mascareignes dont deux sur l'île de la Réunion (*Z. borbonicus* et *Z. olivaceus*), deux sur l'île Maurice (*Z. mauritianus* et *Z. chloronothos*) et aucune sur l'île Rodrigues. Selon l'hypothèse phylogénétique de Warren *et al.* (2006), il y aurait eu deux événements de colonisation de ces îles : l'un par l'ancêtre commun de *Z. borbonicus et Z. mauritianus* et l'autre par celui de *Z. olivaceus* et *Z. chloronothos*. Cette hypothèse est en accord avec les caractéristiques phénotypiques des deux paires d'espèces (Fig. 2).

Figure 2 : *Zosterops borbonicus* (HBHB) en haut à gauche, *Z. mauritianus* en haut à droite, *Z. olivaceus* en bas à gauche, *Z. chloronothos* en bas à droite.

A la différence des trois autres espèces présentes sur les Mascareignes, *Z. borbonicus* présente une forte variabilité phénotypique et génétique sur l'île de la Réunion (2512 km² ; Encadré 2) (Gill 1973; Milá *et al.* 2010). En effet, les oiseaux présentent une variabilité importante de la couleur du plumage à l'échelle micro-géographique : on trouve 5 variantes de couleur différentes distribuées au travers de l'île (Fig. 3).

Figure 3 : Photo des différentes variantes de couleur. De gauche à droite : LBHB, GHB, BNB, G, HBHB.

Les différentes variantes diffèrent dans la proportion de gris et de brun principalement sur le dos et la tête. Les parties brunes des plumes sont constituées d'un dépôt de phéomélanine dans les barbes et d'eumélanine dans les barbules alors que les plumes grises présentent principalement des dépôts d'eumélanine (Gill 1973). A haute altitude, on trouve des oiseaux à dos bruns et tête brunes (appelés HBHB dans la suite de ce manuscrit) et des oiseaux entièrement gris (G) au sein des mêmes populations. Ces oiseaux sont regroupés sous le terme de « forme de haute altitude »[1]. A basse altitude, on trouve trois variantes (ou formes) réparties de manière parapatrique : à l'Ouest, les oiseaux sont bruns à tête brune (LBHB) ; au Nord et à l'Est, ils sont bruns à tête grise (GHB) ; et au Sud, les oiseaux sont bruns à tête grise avec une nuque brune (BNB). Ces différentes variantes sont séparées par des rivières ou des complexes de coulées de lave : la rivière des Galets entre les oiseaux de l'Ouest et du Nord, le Grand Brûlé (zone d'écoulement des laves du Piton de la Fournaise) entre l'Est et le Sud et la rivière St-Etienne entre le Sud et l'Ouest (Fig. 4).

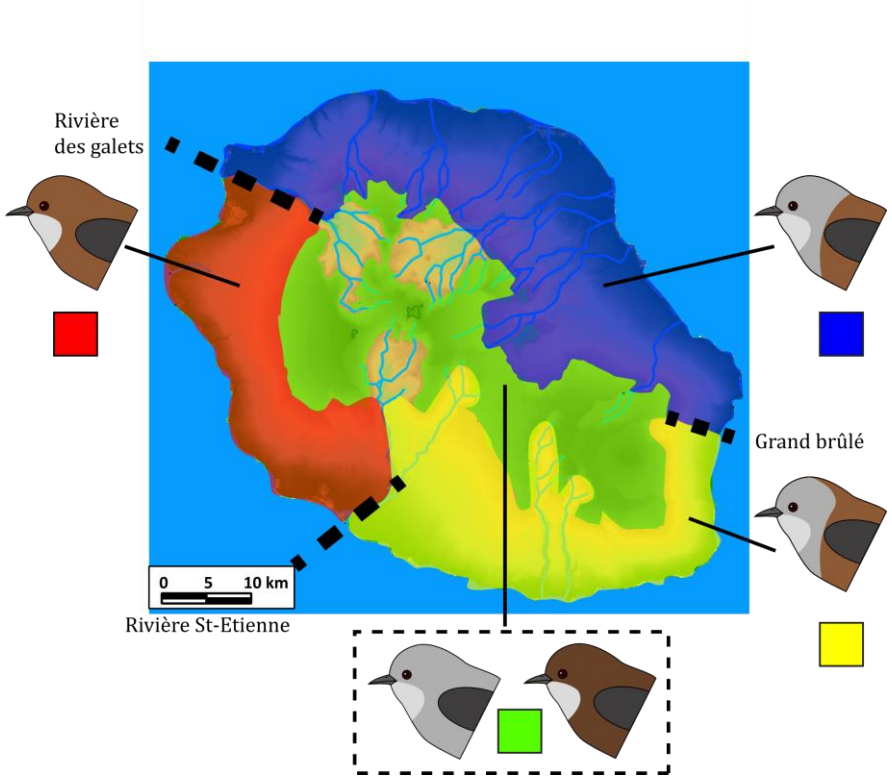


Figure 4 : Aire de répartition des différentes variantes de couleurs. LBHB en rouge, GHB en bleu, BBN en jaune, G et HBHB en vert.

[1] - Jusqu'à récemment (Milá *et al.* 2010), les oiseaux bruns de basse altitude (LBHB) et de haute altitude (HBHB) était considérés comme appartenant à la même forme de couleur. Les travaux de thèse de Yann Bourgeois (Bourgeois 2013), Joris Bertrand (2013) et ceux présentés dans cette thèse ont contribués à établir que cette forme était en fait constituée de deux variantes et que les oiseaux de haute altitude constituait en fait un polymorphisme vrai.

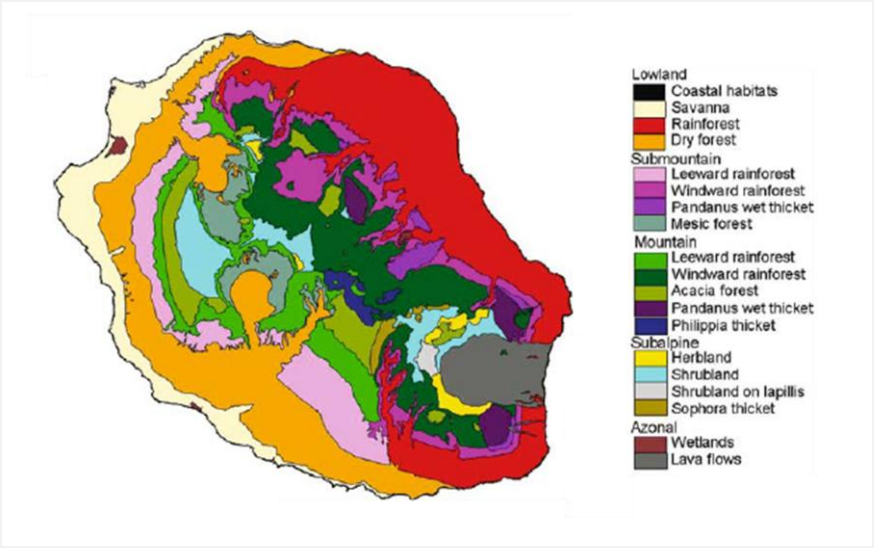
Entre chacune de ses formes, il existe des zones de contact où l'on trouve des phénotypes intermédiaires (Gill 1973). Ces zones de contact sont donc présentes soit sur des barrières physiques à basse altitude, soit sur des gradients altitudinaux forts sans barrière apparente. Les transitions phénotypiques semblent extrêmement abruptes entre les formes de basse altitude alors que les barrières physiques sont de faibles largeurs (Gill 1973). Au contraire, les transitions altitudinales semblent plus graduelles. A l'Est, la transition entre les GHB et la forme de haute altitude a lieu aux alentours de 1500 m d'altitude et semble s'étendre sur plus de 5 km (Gill 1973). A l'Ouest, les oiseaux bruns occupent l'ensemble du gradient altitudinal mais présentent des différences de coloration sur les parties inférieures (poitrine et ventre) avec l'altitude (LBHB et HBHB) (Gill 1973; Chapitre 1). Les oiseaux gris (G) sont observés dès 1000 m d'altitude et sont communs au dessus de 1400 m (Gill 1973; Milá *et al.* 2010). La comparaison entre les patrons de répartitions observés aujourd'hui et ceux décrits par Gill dans les années 1960 montre que cette structuration géographique serait stable depuis quasiment 50 ans. D'autres variations morphologiques sont aussi marquées au travers de l'île puisque les oiseaux de haute altitude apparaissent globalement plus grand que les oiseaux de basse altitude (Gill 1973; Milá *et al.* 2010). L'hypothèse phylogénétique la plus vraisemblable semble indiquer que ces oiseaux partagent un même ancêtre commun et que cette diversification phénotypique ait évolué très récemment (< 430 000 ans) au sein même de l'île (Warren *et al.* 2006; Bourgeois 2013). Ceci constituerait un cas très intéressant de diversification *in-situ*. En effet, selon les travaux récents sur la diversification intra-île et l'échelle spatiale de la spéciation, les exemples de diversification intra-île chez les oiseaux seraient restreints aux îles de grandes tailles (> 10000 km² ; Coyne & Orr 2004; Kisel & Barraclough 2010). Des études plus récentes (Milá *et al.* 2010; Bertrand *et al.* 2014 – Annexe 1) ont quant à elles permis de montrer qu'il existait des différences génétiques au sein même de l'île et que les patrons de différenciation génétique neutre au sein de la Réunion ne suivaient pas strictement la répartition des variantes de couleur. En effet, la structuration génétique (basée sur des marqueurs AFLP) semble plutôt refléter les variations altitudinales (basse *et* haute altitude) et bioclimatiques (Est-Ouest) (Milá *et al.* 2010).

Encadré 2 - La Réunion, une diversité de milieu remarquable

La Réunion est une île relativement jeune (< 3 Ma ; Richer *et al.* 2007 ; Duncan, 2010). Elle est relativement isolée dans l'océan Indien puisqu'elle est située à 4000 km de l'Asie et 1600 km de l'Afrique. L'île la plus proche est Maurice (170 km), Madagascar et Rodrigues se situent quant à elles à plus de 650 km de La Réunion. Le fort taux d'endémisme de sa faune et de sa flore est probablement dû à cet isolement géographique qui tend à limiter les flux de gènes avec les terres les plus proches (Thébaud, 2009).

Elle est caractérisée par un relief particulièrement tourmenté dû à son activité volcanique intense. Le plus haut sommet culmine à 3070 m (Piton des Neiges). Il existe une érosion importante causée, d'une part, par les précipitations tropicales importantes et, d'autre part, par les effondrements d'édifices volcaniques (Richet *et al.* 2007). Cette érosion a ainsi façonnée plusieurs cirques très profonds et de nombreuses vallées encaissées (aussi appelées « ravines »). Les événements volcaniques successifs ainsi que le relief tourmenté de l'île ont probablement joué un rôle dans la démographie et l'isolement des populations au sein de l'île (allopatrie, goulots d'étranglement, etc.).

Ce relief important est à l'origine de la remarquable diversité d'habitats observée sur l'île. D'une part, il génère un gradient de conditions climatiques avec l'altitude et d'autre part, il engendre une forte ségrégation climatique entre l'Est et l'Ouest de l'île. Les températures moyennes sur le littoral sont d'environ 25 °C tandis que les températures moyennes aux sommets des volcans est inférieur à 10 °C. Les précipitations sont très abondantes à l'Est avec plus de 6000 mm/an tandis que l'Ouest reçoit moins de 500 mm/an. Certaines zones d'altitude (env. 1 200 m) à l'Est sont parmi les plus arrosées du monde avec une moyenne annuelle de précipitation de 11000 mm. Ainsi, l'Est de l'île est couvert d'une forêt humide remplacé à très haute altitude par un étage montagnard à végétation éricoïde. A l'Ouest, la zonation altitudinale de la végétation est beaucoup plus marquée puisque se succède 5 grands types de végétation : forêt sèche (< 200 m), forêt semi-sèche (200 à 750 m), forêt humide de basse altitude (750m à 1100m), forêt humide de haute altitude (1100m à 2000m) et lande à végétation éricoïde (> 2000m). Cependant, la majorité de ces habitats ont été modifiés par l'homme et les zones de basse altitude (< 1400m) ne contiennent plus beaucoup de végétation originel.



Tiré de Strasberg *et al*. 2005. Diversité des habitats originels sur la Réunion.

3.2 - Echantillonnage et données

Cette thèse s'intègre dans un projet de recherche débuté en 2007 par Christophe Thébaud, Borja Milá et Philipp Heeb. Dans la continuité des travaux réalisés par Frank Gill dans les années 1960, l'objectif général de ce programme est de comprendre les causes et les bases de la diversification de *Zosterops borbonicus.* Depuis la première étude de Milá *et al.* (2010), qui visait à échantillonner quelques populations au sein de chacune des variantes de couleur de l'île et qui se basait sur 182 individus provenant de 8 localités, l'échantillonnage a largement été augmenté par notre équipe grâce à de nouvelles missions de terrain : à ce jour plus de 1000 individus provenant de 77 localités ont été échantillonnés. Les nouvelles localités ont notamment été placées sur des transects qui traversent les différentes zones de contact entre les formes, sur les gradients altitudinaux à l'Ouest et à l'Est et sur chaque zone de contact de basse altitude, afin d'analyser en détails les zones hybrides qui s'y trouvent. Tous les individus échantillonnés ont été capturés à l'aide de filets japonais. Ils ont alors été bagués avec une bague en aluminium comportant un numéro unique. Sur chaque individu, des prélèvements de sang, de plumes et des mesures morphologiques ont été effectués. Les prélèvements de plumes servent à la quantification de la couleur du plumage via la réalisation de mesures spectrophotométriques. De nouveaux marqueurs moléculaires ont aussi été développé par l'équipe : microsatellites (Bertrand *et al.* 2012) et RAD (Bourgeois *et al.* 2013 – Annexe 2), seuls les marqueurs microsatellites seront utilisés pour ce travail de thèse. Par ailleurs, j'ai aussi eu recours aux données morphologiques collectées par Frank Gill au cours de sa thèse et à des données spectrophotométriques collectées sur les individus conservés au muséum du Michigan (cf. Chapitre 1).

4 - Objectifs de la thèse

Cette thèse a pour objectif de comprendre les processus qui permettent la mise en place et le maintien de la diversité génétique et phénotypique dans des populations en contact, en utilisant *Zosterops borbonicus* comme modèle d'étude. D'une part, je m'intéresserai au rôle de la sélection, des processus neutres et des facteurs historiques dans la mise en place des différentes formes de couleur (Chapitre 1 et 2). D'autre part, je m'intéresserai aux causes du

maintien des différentes zones hybrides présentes entre les formes de couleur et aux effets de l'environnement sur ces celles-ci (Chapitre 3 et 4).

Le premier chapitre traitera du rôle de la sélection naturelle dans l'évolution des différences phénotypiques (morphologie et couleur du plumage) au sein des quatre formes de couleur. Le deuxième chapitre sera consacré à la caractérisation des variations génétiques et phénotypiques sur un gradient altitudinal court mais abrupt d'un point de vue écologique. Dans ce chapitre, l'objectif sera de déterminer les facteurs (neutres, écologiques ou historiques) responsables de la différenciation génétique et phénotypique observée. Dans un troisième chapitre, je m'intéresserai plus particulièrement aux caractéristiques de la zone hybride mise en évidence dans le chapitre 2. A l'aide de modèles de clines, je comparerai les patrons de variations phénotypiques et génétiques pour comprendre les causes du maintien de cette zone hybride. La configuration environnementale de cette zone hybride sera utilisée pour distinguer les traits sous contrôle environnementaux de ceux impliqués dans l'isolement reproducteur. Dans un dernier chapitre, je m'attacherai à caractériser les trois zones hybrides de basse altitude. L'objectif de ce chapitre sera de comprendre comment ces zones sont maintenues à une échelle spatiale extrêmement réduite et sans barrière forte à la dispersion ni structuration environnementale marquée.

References

Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Abbott RJ, Comes HP (2004) Evolution in the Arctic: a phylogeographic analysis of the circumarctic plant, Saxifraga oppositifolia (Purple saxifrage). *New Phytologist*, **161**, 211–224.

Anderson CJR, Harmon L (2014) Ecological and mutation-order speciation in digital organisms. *The American Naturalist*, **183**, 257–268.

Antoniazza S, Burri R, Fumagalli L, Goudet J, Roulin A (2010) Local adaptation maintains clinal variation in melanin-based coloration of European barn owls (Tyto Alba). *Evolution*, **64**, 1944–1954.

Baldassarre DT, White TA, Karubian J, Webster MS (2014) Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. *Evolution*, **68**, 2644–2657.

Barton NH (1979) Gene flow past a cline. *Heredity*, **43**, 333–339.

Barton NH, Bengtsson BO (1986) The barrier to genetic exchange between hybridising populations. *Heredity*, **57**, 357–376.

Barton NH, Briggs DEG, Eisen JA, Goldstein DB, Patel NH (2007) *Evolution*. Cold Spring Harbor, NY.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

Bazykin AD (1969) Hypothetical mechanism of speciaton. *Evolution*, **23**, 685.

Bertrand JAM, Bourgeois YXC, Delahaie B *et al.* (2014) Extremely reduced dispersal and gene flow in an island bird. *Heredity*, **112**, 190–196.

Bertrand JAM, García-Jiménez R, Bourgeois Y *et al.* (2012) Isolation and characterization of twelve polymorphic microsatellite loci for investigating an extreme case of microgeographical variation in an island bird (Zosterops borbonicus). *Conservation Genetics Resources*, **4**, 323–326.

Bierne N (2001) *Barrières au flux génique en milieu marin : sélection et dispersion larvaire dans la zone d'hybridation des moules côtières Mytilus edulis et M. galloprovincialis*. Montpellier 2.

Bierne N, Borsa P, Daguin C *et al.* (2003) Introgression patterns in the mosaic hybrid zone between Mytilus edulis and M. galloprovincialis. *Molecular Ecology*, **12**, 447–461.

Bierne N, David P, Boudry P, Bonhomme F (2002) Assortative Fertilization and Selection at Larval Stage in the Mussels Mytilus Edulis and M. Galloprovincialis. *Evolution*, **56**, 292–298.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044–2072.

Bímová BV, Macholán M, Baird SJE *et al.* (2011) Reinforcement selection acting on the European house mouse hybrid zone. *Molecular Ecology*, **20**, 2403–2424.

Blount ZD, Borland CZ, Lenski RE (2008) Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences*, **105**, 7899–7906.

Boughman JW (2002) How sensory drive can promote speciation. *Trends in Ecology & Evolution*, **17**, 571–577.

Bourgeois Y (2013) Génétique évolutive d'un cas extrême de polymorphisme de la coloration du plumage chez un oiseau insulaire, Zosterops borbonicus (Zosteropideae). Université Toulouse 3 Paul Sabatier, Toulouse.

Bourgeois YXC, Lhuillier E, Cézard T *et al.* (2013) Mass production of SNP markers in a nonmodel passerine bird through RAD sequencing and contig mapping to the zebra finch genome. *Molecular Ecology Resources*, **13**, 899–907.

Calcagno V (2007) Coexistence des espèces, assemblage des communautés et spéciation: rôle de la sélection naturelle: quelques modèles et expériences. Université de Montpellier II.

Clegg S (2010) Evolutionary changes following island colonization in birds: empirical insights into the roles of microevolutionary processes. In: *The theory of island biogeography revisited* (eds Losos JB, Ricklefs RE), pp. 293–326. Princeton, New Jersey.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates Inc.,U.S., Sunderland, Mass.

Currat M, Excoffier L (2005) The effect of the Neolithic expansion on European molecular diversity. *Proceedings of the Royal Society of London B: Biological Sciences*, **272**, 679–688.

Darwin C (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, 1st edn.* London.

Diamond JM (1974) Colonization of exploded volcanic islands by birds: the supertramp strategy. *Science*, **184**, 803–806.

Dobzhansky T (1937) *Genetics and the origin of species*. Columbia University Press, New York.

Dobzhansky T (1970) *Genetics of the evolutionary process*. Columbia University Press, New York.

Doebeli M, Dieckmann U, Metz JAJ, Tautz D, Harrison R (2005) What we have also learned: adaptive speciation is theoretically plausible. *Evolution*, **59**, 691–695.

Doi M, Matsuda M, Tomaru M, Matsubayashi H, Oguma Y (2001) A locus for female discrimination behavior causing sexual isolation in Drosophila. *Proceedings of the National Academy of Sciences*, **98**, 6714–6719.

Edelaar P, Bolnick DI (2012) Non-random gene flow: an underappreciated force in evolution and ecology. *Trends in Ecology & Evolution*, **27**, 659–665.

Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 975–979.

Endler JA (1977) *Geographic variation, speciation, and clines*. Princeton University Press, Princeton, New Jersey.

England PR, Osler GHR, Woodworth LM *et al.* (2003) Effects of intense versus diffuse population bottlenecks on microsatellite genetic diversity and evolutionary potential. *Conservation Genetics*, **4**, 595–604.

Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, **23**, 347–351.

Feder JL, Flaxman SM, Egan SP, Nosil P (2013) Hybridization and the build-up of genomic divergence during speciation. *Journal of Evolutionary Biology*, **26**, 261–266.

Fraïsse C, Roux C, Welch JJ, Bierne N (2014) Gene flow in a mosaic hybrid zone: is local introgression adaptive? *Genetics*, genetics.114.161380.

Frankham R, Weber K (2000) Nature of quantitative genetic variation. In: *Evolutionary genetics: from molecules to morphology* (eds Singh RS, Krimbas CB), pp. 351–368. Cambridge University Press.

Gavrilets S (2003) Perspective: models of speciation: What have we learned in 40 years? *Evolution*, **57**, 2197–2215.

Gay L (2006) *Impact de l'hybridation et de la sélection sur la diversité phénotypique dans deux zones de contact secondaire entre grands goëlands*. Montpellier 2.

Gill FB (1973) Intra-island variation in the Mascarene white-eye Zosterops borbonica. *Ornithological Monographs*, iii–66.

Gill F, Donsker D (2014) IOC Word Bird List (version 4.4).

Glor RE (2011) Remarkable new evidence for island radiation in birds. *Molecular Ecology*, **20**, 4823–4826.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridization: a genetic study of a hybrid zone between red and sika deer (genus Cervus) in Argyll, Scotland. *Genetics*, **152**, 355–371.

Grøndahl E, Ehlers BK (2008) Local adaptation to biotic factors: reciprocal transplants of four species associated with aromatic Thymus pulegioides and T. serpyllum. *Journal of Ecology*, **96**, 981–992.

Harrison RG (1990) Hybrid zones: windows on evolutionary processes. In: *Oxford surveys in evolutionary biology*. (eds Antonovics J, Futuyma D). Oxford, UK.

Harrison RG (1993) *Hybrid zones and the evolutionary process*. Oxford University Press.

Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, **105**, 795–809.

Hewitt GM (1988) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, **3**, 158–167.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hoekstra HE, Krenz JG, Nachman MW (2004) Local adaptation in the rock pocket mouse (Chaetodipus intermedius): natural selection and phylogenetic history of populations. *Heredity*, **94**, 217–228.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Keller I, Alexander JM, Holderegger R, Edwards PJ (2013) Widespread phenotypic and genetic divergence along altitudinal gradients in animals. *Journal of Evolutionary Biology*, **26**, 2527–2543.

Key KHL (1968) The concept of stasipatric speciation. *Systematic Biology*, **17**, 14–22.

Kirkpatrick M, Ravigné V (2002) Speciation by natural and sexual selection: models and experiments. *The American Naturalist*, **159**, S22–S35.

Kisel Y, Barraclough TG (2010) Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, **175**, 316–334.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, **23**, 482–490.

Leinonen T, O'hara RB, Cano JM, Merilä J (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, **21**, 1–17.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Lessios HA (1998) The first stage of speciation as seen in organisms separated by the isthmus of Panama. In: *Endless forms: species and speciation* (eds Howard DJ, Berlocher SH), pp. 186–201.

Lynch M (2007) The evolution of genetic networks by non-adaptive processes. *Nature Reviews Genetics*, **8**, 803–813.

Mallet J (1989) The genetics of warning colour in Peruvian hybrid zones of Heliconius erato and H. melpomene. *Proceedings of the Royal Society of London B: Biological Sciences*, **236**, 163–185.

Mallet J (2001) The speciation revolution. *Journal of Evolutionary Biology*, **14**, 887–888.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, **20**, 229–237.

Martin NH, Willis JH (2007) Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric Mimulus species. *Evolution*, **61**, 68–82.

May RM, Endler JA, McMurtrie RE (1975) Gene Frequency Clines in the Presence of Selection Opposed by Gene Flow. *The American Naturalist*, **109**, 659.

Mayr E (1942) *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press, Cambridge.

Mayr E (1963) *Animal species and evolution*. Harvard University Press, Cambridge.

Mayr E (1970) *Populations, species, and evolution: an abridgment of animal species and evolution*. Belknap Press of Harvard University Press, Cambridge.

McKinnon JS, Mori S, Blackman BK *et al.* (2004) Evidence for ecology's role in speciation. *Nature*, **429**, 294–298.

Melo M, Warren BH, Jones PJ (2011) Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eyes (Aves, Zosteropidae). *Molecular Ecology*, **20**, 4953–4967.

Milá B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: Zosterops borbonicus). *BMC Evolutionary Biology*, **10**, 158.

Miner BG, Sultan SE, Morgan SG, Padilla DK, Relyea RA (2005) Ecological consequences of phenotypic plasticity. *Trends in Ecology & Evolution*, **20**, 685–692.

Moore WS (1977) An evaluation of narrow hybrid zones in vertebrates. *The Quarterly Review of Biology*, **52**, 263–277.

Moyle RG, Filardi CE, Smith CE, Diamond J (2009) Explosive Pleistocene diversification and hemispheric expansion of a "great speciator." *Proceedings of the National Academy of Sciences*, **106**, 1863–1868.

Muller HJ (1942) Isolating mechanisms, evolution, and temperature. *Biological Symposium*, **6**, 71–125.

Nei M (2005) Selectionism and neutralism in molecular evolution. *Molecular biology and evolution*, **22**, 2318–2342.

Noor MAF, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences*, **98**, 12084–12088.

Nosil P (2008) Speciation with gene flow could be common. *Molecular Ecology*, **17**, 2103–2106.

Nosil P, Flaxman SM (2011) Conditions for mutation-order speciation. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 399–407.

Nosil P, Harmon LJ, Seehausen O (2009) Ecological explanations for (incomplete) speciation. *Trends in Ecology & Evolution*, **24**, 145–156.

Nosil P, Vines TH, Funk DJ (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, **59**, 705–719.

O'Hara RB (2005) Comparing the effects of genetic drift and fluctuating selection on genotype frequency changes in the scarlet tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 211–217.

van Balen, B. (2008). Mascarene White-eye (*Zosterops borbonicus*). In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A. & de Juana, E. (eds.) (2014). *Handbook of the Birds of the World Alive*. Lynx Edicions, Barcelona. (retrieved from http://www.hbw.com/node/60224 on 10 December 2014).

van Balen, B. (2008). Mauritius Olive White-eye (*Zosterops chloronothos*). In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A. & de Juana, E. (eds.) (2014). *Handbook of the Birds of the World Alive*. Lynx Edicions, Barcelona. (retrieved from http://www.hbw.com/node/60232 on 10 December 2014).

van Balen, B. (2008). Reunion Olive White-eye (*Zosterops olivaceus*). In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A. & de Juana, E. (eds.) (2014). *Handbook of the Birds of the World Alive*. Lynx Edicions, Barcelona. (retrieved from http://www.hbw.com/node/60231 on 10 December 2014).

Van Oosterhout C, Joyce DA, Cummings SM *et al.* (2006) Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (poecilia reticulata). *Evolution*, **60**, 2562–2574.

Pardo-Diaz C, Salazar C, Baxter SW *et al.* (2012) Adaptive introgression across species boundaries in Heliconius butterflies. *PLoS Genet*, **8**, e1002752.

Payseur BA (2010) Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Molecular Ecology Resources*, **10**, 806–820.

Payseur BA, Krenz JG, Nachman MW (2004) Differential Patterns of Introgression Across the X Chromosome in a Hybrid Zone Between Two Species of House Mice. *Evolution*, **58**, 2064–2078.

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 215–230.

Poelstra JW, Vijay N, Bossu CM *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, **344**, 1410–1414.

Price T (2008) *Speciation in birds*. Greenwood Village, CO.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879–886.

Ramsey J, Bradshaw HD, Schemske DW (2003) Components of reproductive isolation between the monkeyflowers Mimulus lewisii and M. cardinalis (Phrymaceae). *Evolution; International Journal of Organic Evolution*, **57**, 1520–1534.

Räsänen K, Hendry AP (2008) Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, **11**, 624–636.

Rice WR, Hostert EE (1993) Laboratory experiments on speciation: what have we learned in 40 Years? *Evolution*, **47**, 1637.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, **16**, 351–358.

Ruegg K, Anderson EC, Slabbekoorn H (2012) Differences in timing of migration and response to sexual signalling drive asymmetric hybridization across a migratory divide. *Journal of Evolutionary Biology*, **25**, 1741–1750.

Rundle HD (2003) Divergent environments and population bottlenecks fail to generate premating isolation in Drosophila pseudoobscura. *Evolution*, **57**, 2557–2565.

Saetre G-P, Moum T, Bures S *et al.* (1997) A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature*, **387**, 589–592.

Safran RJ, Scordato ESC, Symes LB, Rodríguez RL, Mendelson TC (2013) Contributions of natural and sexual selection to the evolution of premating reproductive isolation: a research agenda. *Trends in Ecology & Evolution*, **28**, 643–650.

Schluter D (2001) Ecology and the origin of species. *Trends in Ecology & Evolution*, **16**, 372–380.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737–741.

Secondi J, Okassa M, Sourice S, Théry M (2014) Habitat-dependent species recognition in hybridizing newts. *Evolutionary Biology*, **41**, 71–80.

Seehausen O, Alphen JJM van, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, **277**, 1808–1811.

Seehausen O, Terai Y, Magalhaes IS *et al.* (2008) Speciation through sensory drive in cichlid fish. *Nature*, **455**, 620–626.

Servedio MR, Noor MAF (2003) The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 339–364.

Singhal S, Moritz C (2012) Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. *Evolution*, **66**, 1474–1489.

Slatkin M (1973) Gene flow and selection in a cline. *Genetics*, **75**, 733–756.

Sobel JM, Chen GF, Watt LR, Schemske DW (2010) The biology of speciation. *Evolution*, **64**, 295–315.

Sota T, Vogler AP (2001) Incongruence of mitochondrial and nuclear gene trees in the Carabid beetles Ohomopterus. *Systematic Biology*, **50**, 39–59.

Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, Bombina bombina and B. variegata, near Cracow in southern Poland. *Evolution*, **40**, 1141.

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology & Evolution*, **16**, 330–343.

Uy JAC, Borgia G (2000) Sexual selection drives rapid divergence in bowerbird display traits. *Evolution*, **54**, 273–278.

Via S (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution*, **16**, 381–390.

Via S (2002) The ecological genetics of speciation. *The American Naturalist*, **159**, S1–S7.

Via S, Lande R (1985) Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. *Evolution*, **39**, 505.

Vignieri SN, Larson JG, Hoekstra HE (2010) The selective advantage of crypsis in mice. *Evolution*, **64**, 2153–2158.

Warren BH, Bermingham E, Prys-Jones RP, Thébaud C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. *Molecular Ecology*, **15**, 3769–3786.

Whitlock MC, Phillips PC (2000) The exquisite corpse: a shifting view of the shifting balance. *Trends in Ecology & Evolution*, **15**, 347–348.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Zinner D, Groeneveld LF, Keller C, Roos C (2009) Mitochondrial phylogeography of baboons (Papio spp.) – Indication for introgressive hybridization? *BMC Evolutionary Biology*, **9**, 83.

Chapitre 1

Rôle de la sélection dans la différenciation phénotypique

Présentation du chapitre

Ce chapitre se concentre sur le rôle de la sélection dans l'évolution des variations phénotypiques observées chez *Zosterops borbonicus*. Frank Gill, dans sa monographie sur l'espèce, proposait que l'évolution des différences de coloration de l'espèce était due à une interaction fine entre sélection divergente et dispersion limitée. Dans les années 1960, il ne disposait cependant pas des outils de génétique quantitative suffisant pour tester cette hypothèse rigoureusement. Afin de quantifier les différences phéntoypiques, nous avons utilisé les données morphologiques récoltées par Frank Gill dans les années 1960 sur 239 individus échantillonnées dans 60 localités d'échantillonnage. Les données de coloration ont été obtenus à partir de mesures spectrophotométriques effectuées par Christophe Thébaud et Philipp Heeb sur un sous échantillon (50 individus) des spécimens conservés au muséum du Michigan. L'article présenté ci-après vise donc à répondre aux trois questions suivantes : i) Quels sont les patrons de variations phénotypiques au sein de l'île ? Est-ce que les différences de coloration sont perceptibles par les oiseaux ? iii) La sélection a-t-elle joué un rôle dans l'évolution des différences phénotypiques chez *Z. borbonicus.*

Contribution : Ce chapitre est constitué d'un article co-rédigé avec Josselin Cornuault, nous nous partageons donc la position de premier auteur. L'idée originale de cet article revient à Josselin. Nous avons néanmoins tous les deux participé à l'ensemble des étapes de la construction de cet article. Josselin a réalisé les analyses de sélection.

Morphological and plumage colour variation in the Réunion grey white-eye (Aves: *Zosterops borbonicus*): assessing the role of selection

Accepted for publication in the Biological Journal of the Linnean Society

Josselin Cornuault[1]*[†], Boris Delahaie[1†], Joris AM Bertrand[1], Yann XC Bourgeois[1], Borja Milá[2], Philipp Heeb[1] & Christophe Thébaud[1]

1. Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174 Centre National de la Recherche Scientifique (CNRS) - Université Paul Sabatier - Ecole Nationale de Formation Agronomique (ENFA), 118 Route de Narbonne, F-31062 Toulouse, France.

2. National Museum of Natural Sciences, Spanish Research Council (CSIC), Madrid 28006, Spain.

[†] These authors have contributed equally to this work.

* Corresponding author:

Josselin Cornuault

Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174 Centre National de la Recherche Scientifique (CNRS) - Université Paul Sabatier - Ecole Nationale de Formation Agronomique (ENFA), 118 Route de Narbonne, F-31062 Toulouse, France.

Tel: +33561556085.

joss.cornuault@gmail.com

Running title: Selection and plumage colour variation in an island bird

Abstract

The Réunion grey white-eye (*Zosterops borbonicus*), a small passerine endemic to the island of Réunion (Mascarene archipelago), constitutes an extraordinary case of phenotypic variation within a bird species, with conspicuous plumage colour differentiation at a microgeographic scale. To understand whether natural selection could explain such variability, we compared patterns of variation in morphological and plumage colour traits within and among populations. To quantify morphological variation, we used measurements obtained by Frank Gill in the 1960s from 239 individuals collected in 60 localities distributed over the entire island of Réunion. To quantify colour variation, we measured the reflectance spectra of plumage patches of 50 males from a subset of Gill's specimens belonging to the five recognized plumage colour variants and used a visual model to project these colours in an avian-appropriate, tetrachromatic, colour space. We found that variants occupy different regions of the avian colour space and that between-variant differences for most plumage patches could be discriminated by the birds. Differences in morphology were also detected, but they were, in overall, smaller than colour differences. Overall, we found that variation in both plumage colour and morphology among variants is greater than would be expected if genetic drift alone was responsible for phenotypic divergence. Since the plumage colour variants correspond to four geographic forms, our results suggest that phenotypic evolution in the Réunion grey white-eye is at least partly explained by divergent selection in different habitats or regions.

Keywords: colouration - geographic variation - Mascarene Islands - morphology - Réunion - selection - white-eye - *Zosterops*

Introduction

Patterns of plumage colour can be strikingly different among closely related bird populations (Hill & McGraw, 2006). This is nowhere as obvious as on islands where populations living in close geographic proximity have often diverged in colour patterns to a greater extent than their mainland counterpart across much vaster distribution ranges (Mayr, 1942; Mayr & Diamond, 2001; Price, 2008). Such patterns of geographic variability in island systems are thought to reflect long isolation times and the role of physical barriers on limiting gene flow (Diamond, Gilpin & Mayr, 1976; Barton, 1996), and are found mostly on large islands or among island populations within archipelagoes (Mayr & Diamond, 2001). However, there are a few striking cases of plumage colour differences that have evolved within islands, and rather remarkably, some are found within relatively small islands (see Coyne & Price (2000) for a list of potential cases). On small islands, historical contingencies (e.g. the vagaries of geological dynamics on volcanic islands) and natural selection arising from altered ecological conditions (e.g. relative scarcity of predators and ecological competitors) have often resulted in a series of unusual adaptations called "island syndrome" (Grant 1998; Blondel 2000; Losos & Ricklefs 2009; Covas 2011). Among these adaptations, reduced dispersal has repeatedly evolved in island birds and it may have facilitated differentiation at small spatial scales (Blondel *et al*., 1999; Komdeur *et al.*, 2004; Moyle *et al.*, 2009; Porlier *et al*., 2012; Bertrand *et al.*, 2014). However, the exact impact of gene flow (or the lack of) on population divergence depends upon several factors, including the spatial context of selection and the strength of divergent selection pressures (Endler, 1973; Lenormand, 2002; Blondel *et al*., 2006). Here, we focused our research on the role of divergent selection in the face of gene flow to explain microgeographic variation in morphology and colouration in an island bird.

We chose *Zosterops borbonicus* (taxonomy following Gill & Donsker, 2014) for our study because this species complex, endemic to the small island of Réunion (2512 km^2), provides an extraordinary case of microgeographical variation in melanic colouration (Gill, 1973), with five distinct colour variants distributed across the island (see Figure 1). This variation, while conspicuous, is relatively complex in terms of melanin pigmentation patterns. Variants differ in the relative extent of grey and brown parts in their plumage. Brown parts involve deposition of phaeomelanin in feather barbs and eumelanin deposition in barbules, while grey parts involve mostly eumelanin deposition. Although colour genes are yet to be discovered in the Réunion grey white-eye (Bourgeois *et al.*, 2012), melanic pigmentation in

birds is genetically determined and differences between variants likely reflect genetic differentiation (Takeuchi *et al.*, 1996; Theron *et al.*, 2001; Mundy *et al.*, 2004). The five variants can be grouped into four geographic forms: three monomorphic forms occupy discrete geographic regions in the lowlands; a fourth form comprises two morphs which occur at high altitudes in complete sympatry (Figure 1; Gill, 1973; Milá *et al.*, 2010) and which, rather unexpectedly, do not show any assortative mating with regards to plumage colour (Gill, 1973; Bourgeois, 2013). Narrow hybrid zones arise where the different forms come into contact, as happens between parapatric lowland forms or between lowland and highland forms (Gill, 1973; Milá *et al.*, 2010).

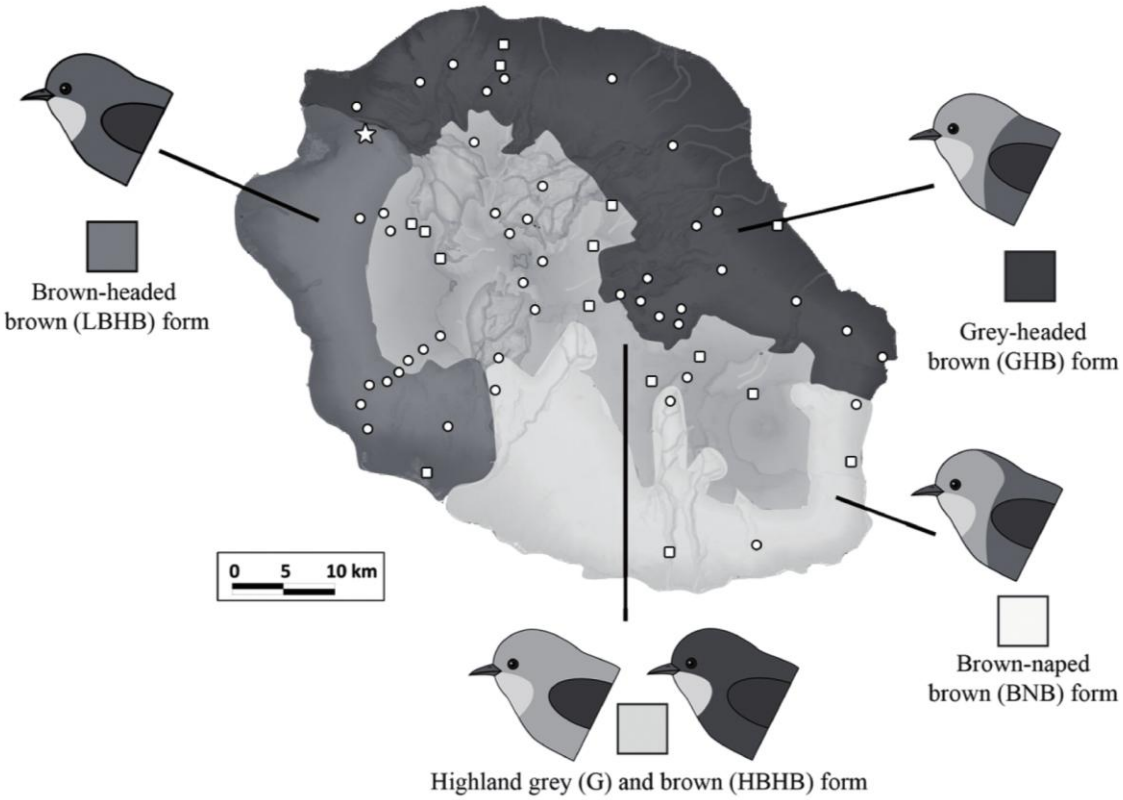


Figure 1: Distribution of the five colour variants of *Zosterops borbonicus* on Réunion.

Map of Réunion with the geographical distribution of the five Réunion grey white-eye plumage colour variants. Note that the three lowland variants correspond to three monomorphic forms that occupy discrete, parapatric, regions in the lowlands while the two highland variants correspond to a polymorphic colour form with a grey and a brown-headed brown morphs. The localities sampled by Gill (1973) and used in this study are shown on the map: Circles, morphological data only; Star, spectrophotometric data only; Squares, both types of data.

In his classic monograph on the Réunion grey white-eye, Frank Gill (1973), building on the observation that the birds are very sedentary with a high degree of site fidelity, suggested that a "delicate balance between gene flow and opposing selection forces must be involved" to explain clinal changes in size and plumage pigmentation with altitude within the different colour forms. However, because these were the early days of quantitative evolutionary genetics (see e.g. Lande, 1979), he was not able to statistically test the hypothesis that selection played a role in the establishment of colour variants and the maintenance of the different forms. In this paper, we revisited Frank Gill's original collections and data and used a quantitative-genetic framework to determine if selection could be responsible for producing the different variants of the Réunion grey white-eye. We first reassessed patterns of variation in morphological and plumage colour traits using modern statistical methods, a spectrophotometer to capture spectral reflectance from different parts of the birds, and an avian visual model to project colour measurements in an avian-appropriate, tetrachromatic, colour space (Goldsmith, 1990; Endler & Mielke, 2005). By using a tetrahedral avian colour space instead of a direct analysis of reflectance spectra, we were able to examine how plumage colour as perceived by the birds themselves varied among the different variants and to discuss its functional significance in relation to colour perception and visual communication (Endler & Mielke, 2005; Stoddard & Prum, 2008). We then tested the hypothesis that divergent natural selection is responsible for geographical variation in plumage colour and morphology among geographic forms. Our approach, similar to the one used by Harmon & Gibson (2006), was to attempt to reject a null hypothesis of neutral evolution by comparing patterns of phenotypic variances and covariances within and among geographic forms (Ackerman & Cheverud, 2002), prior to concluding that selection could explain morphological and plumage colour variation in the Réunion grey white-eye.

Materials and methods

Specimens

All the specimens used in this study were captured and collected by Frank Gill between 1964 and 1967 in 60 localities distributed over the entire island of Réunion. Gill (1973) categorized all specimens into five different variants based on the distribution of brown and grey patches. These variants include a grey variant (G) with grey back and head, a lowland brown-headed brown variant (LBHB) with light brown back and head, a highland

brown-headed brown variant (HBHB) with dark brown back and head, a grey-headed brown variant (GHB) with a brown back and a grey head, and a brown-naped brown variant (BNB) with brown back and nape and a grey crown. Since all specimens were sexed, we decided to restrict our analyses to male specimens so as to control for a putative sexual dimorphism, even though it was considered absent by Gill (1973). In total, we considered 239 specimens for morphometric measurements, with 19 to 64 (mean: 47.8) individuals per variant. Spectrophotometric measurements were obtained directly from specimens caught by Frank Gill and now deposited at the University of Michigan Museum of Zoology. Specimens used in this study are listed by variant and collecting sites in Table S1.

Morphological measurements

To quantify morphological variation, we used for each specimen the morphological measurements that were performed and kindly provided for this study by Frank Gill: wing length (arc) was measured to the nearest 0.5 millimetre; tail length was obtained by inserting a pair of dividers to the base of the middle pair of rectrices and measuring to the tip of the longest rectrix; bill length was taken as the distance from the tip of the upper mandible to the nasofrontal hinge; tarsus length was measured from the back side of the middle of the tibiotarsal-tarsometatarsal joint to the lower edge of the lowest undivided scute on the front of the junction of the metatarsus with the base of the middle toe.

Spectrophotometric measurements and plumage colour analysis

To quantify variation in plumage colour, we captured spectral reflectance from a total of 50 specimens, representing ten male specimens for each of the five colour variants. For that purpose, we used a USB 2000 spectrophotometer connected to a PH-2000 light source via a bifurcate optical fibre probe (Ocean Optics, Dunedin, FL, USA). All measurements were done with the probe positioned at a standardised three-millimetre distance from the measured surface and an angle of 90°. OOIBase32 software (Ocean Optics, Dunedin, FL, USA) was used for the integration of spectrophotometric data. We recorded reflectance of eight plumage patches (rump, back, flanks, crown, throat, nape, breast and belly). For each specimen, the reflectance of each plumage patch was measured three times in the wavelength range between 300 and 700 nm and averaged prior to analysis.

We then used Goldsmith's (1990) tetrahedral colour space for analysing colour spectra, since this colour space provides a quantitative map of avian colour perception that is easy to

calculate and quantitatively precise, without requiring measurements of ambient light variation from the appropriate environments (Stoddart & Prum, 2008). By integrating reflectance spectra and the cone sensitivity spectra, it provides a powerful framework to study visual communication through the mapping of the signaller phenotype onto the realized sensory phenotype of the receiver. Colours are represented by vectors, whose components are the relative stimulations of the four types of avian retinal cones: ultraviolet-sensitive (*u*), short-wavelength-sensitive (*s*), medium-wavelength-sensitive (*m*) and long-wavelength-sensitive (*l*). The stimulation of each type of cone was calculated with equation 1 in Stoddard & Prum (2008), using spectral sensitivity functions measured in the blue tit (Hart *et al.*, 2000), assuming that these functions are broadly similar in white-eyes. Following Goldsmith (1990) and Stoddard & Prum (2008), we did not take into account the influence that ambient light (the irradiance spectrum) could have on colour perception by birds and treated the irradiance as a constant across all wavelengths. This implies that colours are analysed as perceived under a white light. The $\{u\ s\ m\ l\}$ values were then scaled to sum to 1 and converted into spherical coordinates$\{\theta, \phi, r\}$, with θ and ϕ the angles defining the direction of the colour vector (the hue) and r the length of the colour vector (the chroma). ϕ is the UV component of hue, while θ pertains to other hue components (Endler & Mielke, 2005). Following Stoddard & Prum (2008), we chose to use achieved chroma (r_a) which corresponds to the chroma of a colour relative to the maximum chroma given its hue, which is more informative than chroma itself. We calculated normalized brilliance using the definition of Stoddard & Prum (2008) as a measure of colours' brightness (where 0 corresponds to pure white and 1 corresponds to pure black).

To quantify colour differences between individuals for a given plumage patch, we evaluated disparity in the different components of colour (hue, chroma, brilliance). Hue disparity measures the difference in the hue of any two colours and was calculated using a corrected version of equation A3 in Stoddard & Prum (2008). Achieved chroma disparity and normalized brilliance disparity were measured as the absolute value of the difference between the achieved chroma and the normalized brilliance of two colours, respectively. We also obtained values of colour span, as defined by Stoddard & Prum (2008), i.e. an integrative measure of the difference between any two colours (i.e. it integrates differences in hue and chroma).

Colour variables described above characterize the colour of a given plumage patch or the difference in the colour of two plumage patches. To obtain colour measures of entire plumages, we used an integrative approach by which we derived a series of variables

summarising the colours of all plumage patches for a given individual. We calculated the average, maximum and variance of normalized brilliance and achieved chroma across all plumage patches so that we obtained measures of how bright and colourful an individual could be for other individuals. We also calculated the average colour span and average hue disparity (within individual and between patches) as measures of the within-plumage variance in colour.

Determining whether birds are able to discriminate between pairs of colours provides crucial information about the functional significance of plumage colour variation in relation to colour perception and visual communication. To do so, we used Vorobyev & Osorio's (1998) model of colour perception to calculate ΔS, the chromatic contrast between two colours, in units of just-noticeable-differences (JNDs), where 1 JND is the threshold value of discrimination between two colours. We calculated ΔS with equation 5 in Vorobyev & Osorio (1998), using cone densities experimentally estimated in the blue tit (1:2:2:4) (Hart *et al.,* 2000) and the noise-to-signal ratio estimated in the Red-billed Leiothrix (0.1, Vorobyev *et al.,* 1998).

All the calculations done in this section have been done with the R-package {pavo} (Maia *et al.,* 2013). Further detailed explanations about all the calculations mentioned above and how to interpret them in the context of bird vision can be found in Stoddard & Prum (2008).

Analyses

Phenotypic variation

Principal Component Analysis (PCA) was used to visualise differences in plumage colour and morphology among variants. Non-parametric multivariate analysis of variance (npMANOVA; Anderson, 2001) was then used to test for pairwise differences between variants. We used pairwise tests to determine which particular pairs of variants were significantly different, and applied Holm's correction (Holm, 1979) to all *p*-values to correct the accrued type I error resulting from multiple comparisons. This test makes use of distance matrices as dependant variables and can be univariate or multivariate depending on whether the distance matrix is derived from a single or several variables. We hereafter refer to npANOVA for univariate tests and to npMANOVA for multivariate tests and all tests were carried out with 100,000 permutations. These tests were carried out with the function *adonis* in the R-package {vegan} (Oksanen *et al.,* 2008). PCA and npMANOVA were carried out with the morphological measurements and variables obtained from spectrophotometric data

(θ, ϕ, $r_{achieved}$, normalized brilliance for eight plumage patches). In each case, the elements of the dependent distance matrix were calculated as the multivariate Euclidean distance between each pair of birds after scaling the data so that all variables have mean 0 and variance 1. Variant was used as the independent variable.

We tested for differences among variants in each individual morphological trait and in each single plumage colour patch using npANOVAs. For each morphological trait, the dependent variable was a between-individual Euclidean distance matrix. For each plumage patch the dependent variable was a distance matrix whose elements consisted of colour disparity measures between each pair of individual birds (i.e. normalized brilliance disparity, achieved chroma disparity, hue disparity, and colour span in Euclidean distance and bird JND). In all cases, variant was used as the independent variable. This test was repeated for the four morphological traits and the eight plumage patches for which spectrophotometric measurements were taken.

We also used npANOVAs to test for differences among variants in entire-plumage colour variables (i.e. average normalized brilliance, maximum normalized brilliance, variance of normalized brilliance, average achieved chroma, maximum achieved chroma, variance of achieved chroma, average colour span and average hue disparity). For each variable, the elements of the dependent distance matrix were calculated as pairwise Euclidean distances between the values of the variable, and variant was used as the independent variable.

Role of selection

To determine if selection could be responsible for producing phenotypic variation in the Réunion grey white-eye, we compared within- and between-form patterns of variance and covariance for morphological (length of bill, wing, tail and tarsus) and plumage colour traits (θ, ϕ, $r_{achieved}$, normalized brilliance for eight plumage patches) (Ackermann & Cheverud, 2002). The null hypothesis assumes that only genetic drift is responsible for the observed differences among geographic forms. In this case, within- and between-form variances are expected to be proportional (Ackermann & Cheverud, 2002; Marroig & Cheverud, 2004). Within- and between-form variances were calculated along each principal component (PC) of the correlation matrix (data were scaled so that all variables have mean 0 and variance 1 before analysis). The within-form variance of a PC is its eigenvalue and the between-form variance is the variance of the forms' mean scores along the PC (Harmon & Gibson, 2006). If between- (B) and within-form (W) variances are proportional (as expected under drift), the

regression of ln(B) over ln(W) yields a slope of 1. A slope greater than 1 indicates that the first PC axes are more variable between geographic forms than expected under drift (Harmon & Gibson, 2006). Conversely, a slope lower than 1 occurs when the last PC axes are more variable than expected under drift. Such deviations from the null expectation suggests that other processes such as divergent selection have been at work (Ackermann & Cheverud, 2002). We tested whether the slope of the log-log regression of B versus W variances differed from 1 with a linear regression of ln(B) minus ln(W) over ln(W). We used the Shapiro-Wilk test (Royston, 1982) to verify that the distribution of the model residuals was consistent with a normal distribution.

All statistical analyses were carried out with R software v.2.15.2 (R Development Core Team, 2012).

Results

Morphological differences among variants

PCA loadings for morphological traits can be found in Table S2. The first two PCA axes accounted for 56% and 25% of the total variance in measurements, respectively. Factors that contributed most heavily to the first axis were tarsus length, wing length, and tail length, thus mostly illustrating variation in body size along that axis (Table S2). The second axis was mostly affected by bill length (Table S2). Morphological differences among variants are visible on the PCA summary plot (Figure 2A). These differences were found to be significant for all pairs of variants except the comparison between the two highland morphs (HBHB and G) and the comparison between two of the lowland variants (GHB and BNB) (Table 1). In spite of much overlap between the different variants on the first plane of the PCA summary plot (Figure 2A), highland and lowland variants differ strikingly along the first PCA axis (representing body size), with higher values for highland variants and lower values for lowland variants, revealing that highland birds are on average larger than lowland ones. Trait-by-trait analyses further indicated that all morphological traits (wing, tail, bill and tarsus lengths) differed significantly among variants (Table 2), emphasizing a greater morphological variation among variants than within variants in *Z. borbonicus*.

Table 1: Multivariate analysis of between-variant differences in morphometry and colour in *Zosterops borbonicus*. These data give the results of the npMANOVAs that test for differences between each pair of variants in morphometry and colour. Significant adjusted *P*-values under a 5% error threshold are in bold. R^2 values represent the proportion of variance explained by the factor 'Variant' for each comparison.

		Morphometry				Spectrophotometric colour data			
Variant 1	Variant 2	$N_{Variant1}$	$N_{Variant2}$	R^2	*p*	$N_{Variant1}$	$N_{Variant2}$	R^2	*p*
G	LBHB	46	50	0.15	**<0.001**	10	10	0.29	**<0.001**
HBHB	G	64	46	0.01	0.3217	10	10	0.32	**<0.001**
GHB	G	60	46	0.33	**<0.001**	10	10	0.21	**<0.001**
BNB	G	19	46	0.21	**<0.001**	10	10	0.34	**<0.001**
HBHB	LBHB	64	50	0.20	**<0.001**	10	10	0.26	**<0.001**
GHB	LBHB	60	50	0.17	**<0.001**	10	10	0.18	**<0.001**
BNB	LBHB	19	50	0.14	**<0.001**	10	10	0.12	**0.03**
GHB	HBHB	60	64	0.37	**<0.001**	10	10	0.35	**<0.001**
BNB	HBHB	19	64	0.22	**<0.001**	10	10	0.26	**<0.001**
BNB	GHB	19	60	0.04	0.051	10	10	0.24	**<0.001**

Colour differences between variants

PCA loadings for colour variables can be found in Table S3. We found clear colour differences among variants (Figure 2B). In contrast to morphology, npMANOVAs in colour traits showed that all between-variant comparisons were significant (Table 1). However, after excluding the pairs of variants not differing by their morphology (HBHB/G and GHB/BNB, Table 1), similar R^2 values were found for morphology (average 0.22) and colour (average 0.25), revealing a similar extent of differentiation for morphology and colour at least among some of the variants. The mean spectra for each variant and each patch are given in Figure S1. All spectra were mostly flat, reflecting that colours composing the plumage of *Z. borbonicus* (white, brown or grey) are mainly achromatic.

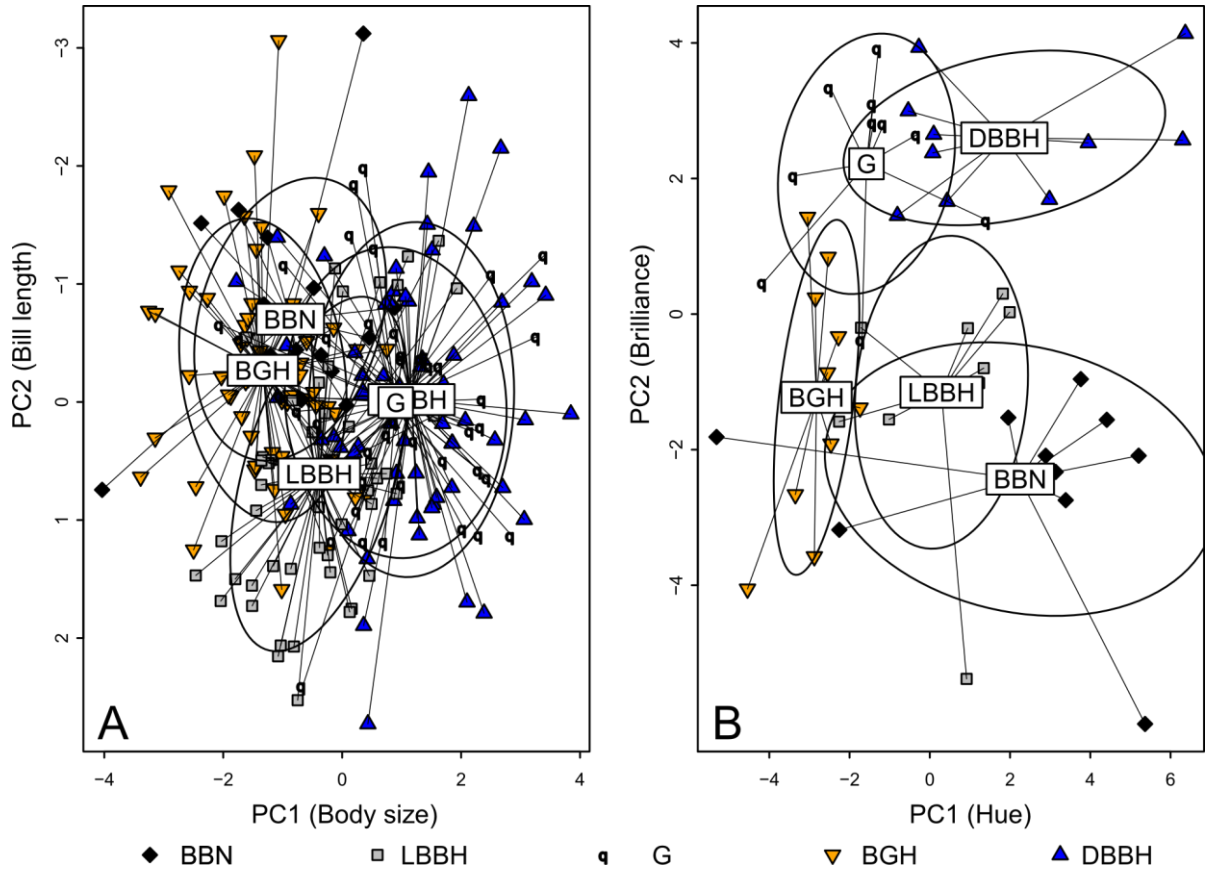


Figure 2: Principal Component Analysis of phenotypic variation in *Zosterops borbonicus*.

A. Morphology, cumulative variance of the first two axes: 81%. B. Spectrophotometric data, cumulative variance of the first two axes: 45%. Each variant's ellipse is oriented so that its focal axis is parallel to the first principal component of the variant point cloud in the PCA space. Ellipses are only graphical tools helping summarize the relative dispersion of the point clouds of each variant. The relative sizes of ellipses are interpretable, while their absolute sizes are arbitrary. Segments connect each data point to the centroid of the point cloud of the variant it belongs to.

Patch-per-patch analyses showed that all colour variables (normalized brilliance, achieved chroma, hue and colour span) differed significantly among variants for all patches except for the rump, the throat, and the hue of the breast (Table 2). The relative position of the variants in the tetrahedral colour space (Figure 3) looks similar to what is perceived by the human eye: the G variant is the only variant with a grey back and stands alone in the back tetrahedron; the GHB and G variants have a grey nape and cluster together in the nape tetrahedron; the LBHB and HBHB variants have a brown crown and cluster together in the crown tetrahedron; the G variant has generally grey flanks with more or less prominent traces of brown, as opposed to

the markedly rufous-brown flanks of other variants and stands alone in the flank tetrahedron (Figure 3). Breast colour of the HBHB variant is particularly divergent from the others (Figure 3). Belly colour is also highly variable to the bird's eye (Table 2), but Figure 3 does not show clear differences between the variants for this patch. The colour span in avian JNDs is on average greater than 1 for both within- and between-variant comparisons, indicating that between-individual colour differences are perceivable by birds under a white illuminant, for every patch measured. In addition, for all patches except rump and throat, colour span in avian JNDs is significantly greater for inter-variant comparisons than for intra-variant comparisons, indicating that birds see colour differences among variants for all these patches (Table 2).

All whole-plumage variables (average normalized brilliance, maximum normalized brilliance, variance of normalized brilliance, average achieved chroma, maximum achieved chroma, variance of achieved chroma, average colour span and average hue disparity) significantly differed among variants (Table 3). The average and maximum brilliance of the two highland variants (G and HBHB) is on average lower than for lowland variants. These differences are low, as the normalized brilliance averaged across all plumage patches varies between 11.1 and 16.2% (means per variant). Average and maximum achieved chroma are lowest for the G variant, reflecting the mainly achromatic nature of the plumage of this all-grey variant. Average colour span, hue disparity and the variance of brilliance and chroma of the G variant are also among the lowest values, indicating that it is the most uniform variant in terms of colour.

Table 2: Between-variant differences in morphology and colour of eight plumage patches of *Zosterops borbonicus*. Measures of morphological difference were calculated for each traits and each pair of individuals. Measures of colour disparity were calculated for each plumage patch and each pair of individuals. The within- and among-variant averages are given, showing for which patches and which variables the plumage of *Z. borbonicus* is more different among variants than within. Colour span in JNDs indicates to what extent birds are able to discriminate between pairs of colours. When colour span is greater than 1 JND (underlined), the difference is perceivable to the bird's eye. Results of npANOVAs are given with *P*-values significant under the 5% error threshold in bold.

				npANOVAs	
		Average Within-Variant	Average Between-Variant	R^2	*p*
Morphology:					
	Wing length difference (mm)	1.65	2.23	0.20	**<0.001**
	Tail length difference (mm)	1.58	2.07	0.21	**<0.001**
	Bill length difference (mm)	0.62	0.73	0.14	**<0.001**
	Tarsus length difference (mm)	0.75	0.92	0.16	**<0.001**
Rump:					
	Normalized brilliance disparity (%)	4.51×10^{-2}	4.89×10^{-2}	0.14	0.079
	Achieved chroma disparity (%)	6.85×10^{-2}	6.73×10^{-2}	0.1	0.283
	Hue disparity (radian)	2.79×10^{-1}	2.89×10^{-1}	0.12	0.131
	Colour span (Euclidean distance)	2.34×10^{-2}	2.31×10^{-2}	0.08	0.477
	Colour span (bird JND)	1.46	1.44	0.07	0.546
Back:					
	Normalized brilliance disparity (%)	1.20×10^{-2}	1.54×10^{-2}	0.24	**<0.001**
	Achieved chroma disparity (%)	7.01×10^{-2}	1.12×10^{-1}	0.39	**<0.001**
	Hue disparity (radian)	1.28×10^{-1}	1.54×10^{-1}	0.24	**0.002**
	Colour span (Euclidean distance)	2.48×10^{-2}	4.53×10^{-2}	0.61	**<0.001**
	Colour span (bird JND)	1.65	3.05	0.62	**<0.001**
Flank:					
	Normalized brilliance disparity (%)	2.32×10^{-2}	3.52×10^{-2}	0.38	**<0.001**
	Achieved chroma disparity (%)	7.65×10^{-2}	1.19×10^{-1}	0.39	**<0.001**
	Hue disparity (radian)	6.29×10^{-2}	1.42×10^{-1}	0.62	**<0.001**
	Colour span (Euclidean distance)	2.90×10^{-2}	5.66×10^{-2}	0.63	**<0.001**
	Colour span (bird JND)	2.1	4.03	0.62	**<0.001**
Crown:					
	Normalized brilliance disparity (%)	1.10×10^{-2}	1.61×10^{-2}	0.27	**<0.001**
	Achieved chroma disparity (%)	9.74×10^{-2}	1.72×10^{-1}	0.44	**<0.001**
	Hue disparity (radian)	2.06×10^{-1}	2.63×10^{-1}	0.32	**<0.001**
	Colour span (Euclidean distance)	3.23×10^{-2}	5.80×10^{-2}	0.53	**<0.001**
	Colour span (bird JND)	2.24	3.93	0.52	**<0.001**

Throat:					
	Normalized brilliance disparity (%)	4.42×10^{-2}	8.50×10^{-2}	0.52	**<0.001**
	Achieved chroma disparity (%)	7.07×10^{-2}	7.20×10^{-2}	0.09	0.371
	Hue disparity (radian)	2.48×10^{-1}	2.81×10^{-1}	0.12	0.122
	Colour span (Euclidean distance)	2.22×10^{-2}	2.29×10^{-2}	0.09	0.369
	Colour span (bird JND)	1.37	1.4	0.08	0.408
Nape:					
	Normalized brilliance disparity (%)	1.13×10^{-2}	1.65×10^{-2}	0.34	**<0.001**
	Achieved chroma disparity (%)	6.11×10^{-2}	1.38×10^{-1}	0.67	**<0.001**
	Hue disparity (radian)	1.32×10^{-1}	1.65×10^{-1}	0.37	**<0.001**
	Colour span (Euclidean distance)	2.24×10^{-2}	5.23×10^{-2}	0.77	**<0.001**
	Colour span (bird JND)	1.49	3.51	0.77	**<0.001**
Breast:					
	Normalized brilliance disparity (%)	3.15×10^{-2}	5.28×10^{-2}	0.44	**<0.001**
	Achieved chroma disparity (%)	5.87×10^{-2}	7.04×10^{-2}	0.18	**0.01**
	Hue disparity (radian)	2.72×10^{-1}	2.85×10^{-1}	0.14	0.07
	Colour span (Euclidean distance)	2.02×10^{-2}	2.51×10^{-2}	0.29	**<0.001**
	Colour span (bird JND)	1.26	1.61	0.32	**<0.001**
Belly:					
	Normalized brilliance disparity (%)	3.40×10^{-2}	6.12×10^{-2}	0.48	**<0.001**
	Achieved chroma disparity (%)	5.48×10^{-2}	6.17×10^{-2}	0.19	**0.012**
	Hue disparity (radian)	2.04×10^{-1}	2.45×10^{-1}	0.25	**<0.001**
	Colour span (Euclidean distance)	1.95×10^{-2}	2.13×10^{-2}	0.20	**0.005**
	Colour span (bird JND)	1.23	1.33	0.18	**0.006**

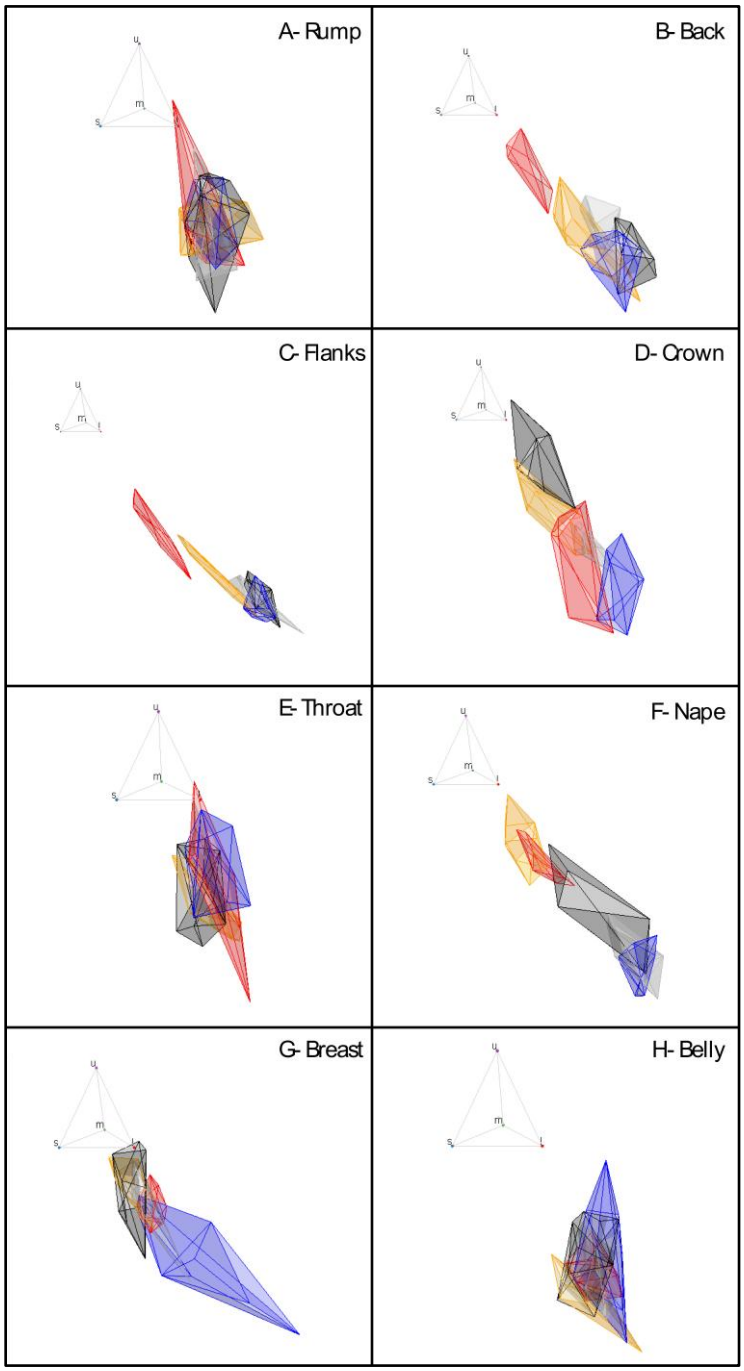


Figure 3. Projection of the colours of *Zosterops borbonicus* individuals into Goldsmith's tetrahedral colour space. Axes correspond to the percent stimulation of the different types of retinal cones: u, UV-sensitive; s, small-wavelength-sensitive, m, medium-wavelength-sensitive, l, long-wavelength-sensitive. Each volume represents minimum convex polygons constructed around each colour variant for a given plumage patch. Colours correspond to the different variants: blue, HBHB; red, G; grey, LBHB; black, BNB; orange, GHB. As all colour points were relatively close to the achromatic centre of the tetrahedron, representing the entire tetrahedron axes (from 0 to 100% cone stimulation) would produce a very small, unintelligible, point cloud near the achromatic centre. Consequently, the axes for each panel extend from 0 to 2.5% cone stimulation. The larger the axes in a panel, the smaller the dispersion of the point cloud.

Evidence for selection

For all regression analyses, Shapiro-Wilk tests indicated that residuals did not significantly deviate from a normal distribution. The log-log regression slope of between- against within-form phenotypic variance is 2.01 (SE 0.15) for morphological measurements and 1.50 (SE 0.08) for plumage colour data (Figure 4). The slopes of linear regressions of ln(B) minus ln(W) over ln(W) were significantly greater than 1 for both morphology ($p_{slope \neq 1}$ = 0.02; Figure 4A) and plumage colour ($p_{slope \neq 1}$ < 0.001; Figure 4B), allowing us to reject the null expectation that within- and between-form variance patterns are proportional. Thus, along the most variable PCA axes, more variation exists among the different geographic forms of *Z. borbonicus* than would be expected under drift for both morphological and plumage colour traits (Tables S2 and S3).

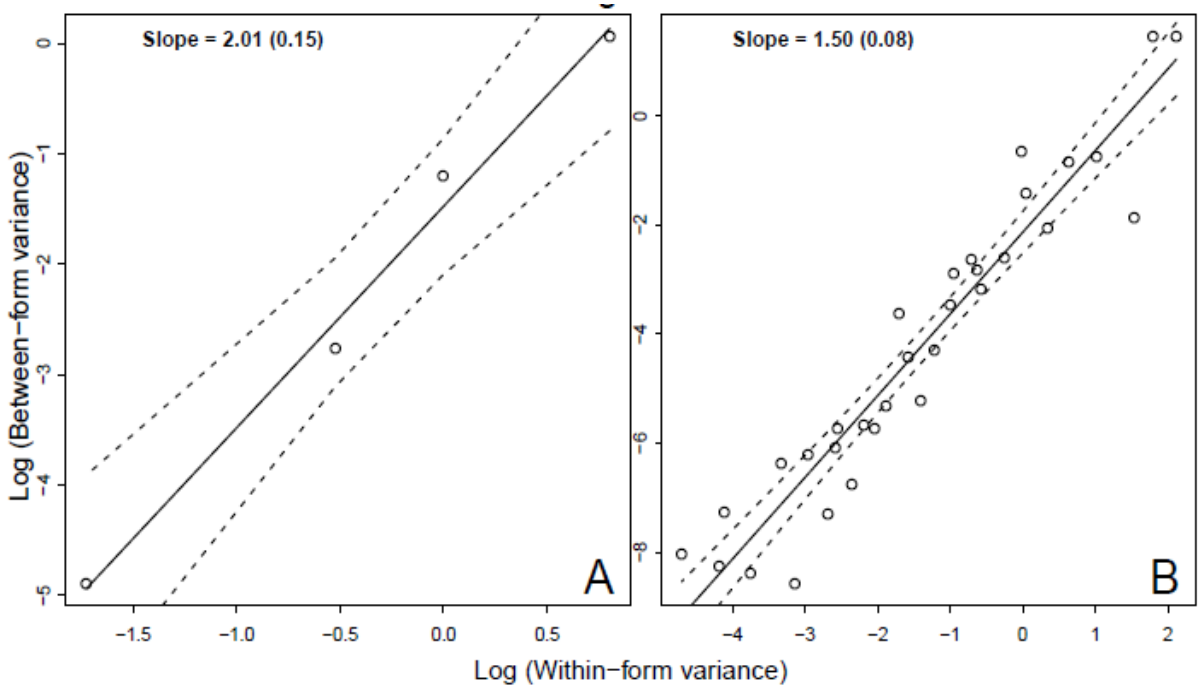


Figure 4: Regression of between- against within-form variances along each PC axis. Solid lines show the regression line; Dashed lines show the 95 % confidence interval. Slopes of regression lines are indicated with standard errors in parentheses. A, morphology; B, plumage colour.

Table 3. Between-variant differences in the entire-plumage colour of *Zosterops borbonicus*. The means of colour variables that characterize the entire plumage are given per variant and for the whole species, with standard deviations in parentheses. These data give information about the intensity (average and maximum achieved chroma) and the brightness (average and maximum brilliance) of the colours composing *Z. borbonicus* plumages. Average colour span and hue disparity reflect the heterogeneity of the colours within each plumage, variance of achieved chroma reflects the variation in the intensity of colours and variance of brilliance reflects the variation in the brightness of colours. For each variable, results of npANOVAs are given, illustrating that all these entire-plumage variables differ among variants, with generally high r-squared.

	Average Brilliance	Variance of Brilliance	Maximum Brilliance	Average Colour Span	Average Hue Disparity	Average Achieved Chroma	Variance of Achieved Chroma	Maximum Achieved Chroma
LBHB	1.62×10^{-1} (2.37×10^{-2})	8.87×10^{-3} (3.53×10^{-3})	2.84×10^{-1} (4.21×10^{-2})	6.14×10^{-2} (6.85×10^{-3})	2.06×10^{-1} (5.04×10^{-2})	2.56×10^{-1} (2.98×10^{-2})	1.67×10^{-2} (4.86×10^{-3})	4.87×10^{-1} (4.87×10^{-2})
HBHB	1.11×10^{-1} (1.07×10^{-2})	3.39×10^{-3} (1.36×10^{-3})	1.96×10^{-1} (2.77×10^{-2})	6.86×10^{-2} (6.20×10^{-3})	1.78×10^{-1} (4.92×10^{-2})	2.94×10^{-1} (3.55×10^{-2})	2.35×10^{-2} (6.62×10^{-3})	5.28×10^{-1} (5.69×10^{-2})
BNB	1.56×10^{-1} (1.54×10^{-2})	9.06×10^{-3} (1.96×10^{-3})	2.95×10^{-1} (4.61×10^{-2})	6.64×10^{-2} (5.15×10^{-3})	3.12×10^{-1} (9.93×10^{-2})	2.24×10^{-1} (4.48×10^{-2})	1.76×10^{-2} (2.90×10^{-3})	4.86×10^{-1} (5.36×10^{-2})
GHB	1.59×10^{-1} (1.64×10^{-2})	6.19×10^{-3} (1.59×10^{-3})	2.68×10^{-1} (2.70×10^{-2})	5.03×10^{-2} (8.50×10^{-3})	2.14×10^{-1} (4.96×10^{-2})	2.26×10^{-1} (2.56×10^{-2})	1.55×10^{-2} (5.95×10^{-3})	4.57×10^{-1} (8.53×10^{-2})
G	1.31×10^{-1} (1.79×10^{-2})	3.60×10^{-3} (1.37×10^{-3})	2.28×10^{-1} (3.82×10^{-2})	3.72×10^{-2} (9.00×10^{-3})	1.75×10^{-1} (9.26×10^{-2})	2.02×10^{-1} (2.70×10^{-2})	1.23×10^{-2} (7.47×10^{-3})	4.16×10^{-1} (9.62×10^{-2})
Z. borbonicus	1.44×10^{-1} (2.61×10^{-2})	6.22×10^{-3} (3.20×10^{-3})	2.54×10^{-1} (5.15×10^{-2})	5.68×10^{-2} (1.37×10^{-2})	2.17×10^{-1} (8.54×10^{-2})	2.41×10^{-1} (4.52×10^{-2})	1.71×10^{-2} (6.66×10^{-3})	4.75×10^{-1} (7.75×10^{-2})
npANOVA R^2	0.59	0.60	0.52	0.74	0.35	0.50	0.31	0.23
npANOVA P	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**

Discussion

Significance of plumage colour variation

Melanic pigmentation is known to be mainly genetically determined and thus heritable in birds (Theron *et al.*, 2001; Mundy *et al.*, 2004). Besides, adaptive correlates of melanin variation patterns in birds are varied and numerous (Price & Bontrager, 2001), providing grounds for an action of selection on plumage colour traits. Melanin-based plumage colouration is often described by the distribution of eumelanin and phaeomelanin pigments throughout the bird's plumage, and such description provided the basis for defining the five distinct colour variants of *Z. borbonicus* (Gill, 1973). While melanin-pigment contents of feathers provide useful proxies for colour phenotype variation (see e.g. McGraw, Safran & Wakamatsu, 2005), they do not directly reflect the colour space as perceived by birds.

Here, the use of an avian visual model for quantifying colour differences allowed us to establish that the between-variant differences in plumage colour can be perceived by the birds and that this perception extends to two plumage patches (breast and belly) that were formerly not considered as variable among the variants, based on human vision (Gill, 1973). Moreover, subtle differences between the two all-brown variants (LBHB and HBHB) appear very pronounced when considering bird vision, something that had never been anticipated on the basis of pigment characteristics. The colour space analysis provides a realistic representation of what is seen by birds and presents multiple advantages over simple spectral descriptions for studying the functional significance and evolution of plumage colour (Stoddard & Prum, 2008). We were able to show that not only can the birds perceive the colour differences between variants but both intra- and inter-variant differences in plumage colour can also be discriminated. Thus, our results highlight the biological significance of the plumage colour differences observed in *Z. borbonicus*.

Since colouration signals are commonly used as a mate choice criterion in birds (Roulin & Bize, 2007), colour differences such as those detected in this study, independently of their evolutionary origin, could potentially act as pre-mating barriers between the different *Z. borbonicus* forms, e.g. through assortative mating (e.g. Baker & Baker, 1990; Uy, Moyle & Filardi, 2009). This may be especially relevant to lowland variants, as they correspond to three geographic forms meeting at very narrow hybrid zones across which there is no apparent change in the environment (Gill, 1973). However, this clearly does not apply to the grey and brown variants, which are true morphs, occurring at high altitudes in complete sympatry and not showing any assortative mating with regards to plumage colour (Gill, 1973; Bourgeois, 2013). Testing whether birds from the different lowland forms show a strong preference for

partners from the same colour and colour patterns will require further field studies of mating patterns and behavioural interactions.

Role of selection

Gill (1973) proposed that the degree of differentiation between the different *Z. borbonicus* forms results from the combined effects of reduced dispersal and small-scale spatial variation in selective factors. However, he did not specifically test for the role of selection in shaping the evolution of phenotypic differences. Our results, based on a quantitative analysis of multivariate phenotypic evolution, demonstrate the role of selection as an important driver of morphological and plumage colour variation in *Z. borbonicus*. Since, to the best of our knowledge, the species is not dimorphic with respect to size and melanin-based plumage colouration (Gill, 1973; Boris Delahaie, *unpublished data*), sexual selection is probably weak relative to natural selection counteracting divergence for these traits between males and females (Andersson 1994) and unlikely to explain phenotypic diversification across variants. Therefore, we restrict our discussion to the possible mechanisms and causes of natural selection acting on morphology and plumage colouration in *Z. borbonicus* populations.

We found that birds were larger in the highlands (Figure 2A), with size-related traits best explaining the differences in morphology between lowland and highland forms (mostly wing and tail lengths; Table 2). Our test for selection on morphological traits reveals that selection has likely influenced this altitudinal increase in size-related traits. This interpretation is, of course, correct only if we assume a genetic basis to body size variation among populations of *Z. borbonicus*, something we currently do not know. Altitudinal increase in size-related traits is often interpreted as an adaptive response for reducing heat loss in the colder environment of high altitude, following the early suggestion of Bergmann (1847) (but see McNab, 2010). However, such increase could also be driven by other adaptive mechanisms related to climate (e.g. increased ability of larger individuals to withstand starvation in the extreme, highly variable, environment of high altitude) or to other factors (e.g. food availability, parasites and pathogens) (McNab, 2010; Hille & Cooper, 2014). To further assess the adaptive nature of size changes and to clarify the role of the potential agents of selection, future quantitative genetic analyses combining capture-recapture data with pedigree information will be necessary.

Variants greatly differ in colour and colour patterns, and a number of possible agents of selection may have shaped the evolution of these differences (Price & Bontrager, 2001).

First, it is well-known that melanic pigmentation plays a role in thermoregulation (Riley, 1997; Ward *et al.*, 2002). Therefore, it is possible that variation in temperature along altitudinal gradients (*Z. borbonicus* occurs from 0 to > 2400m a.s.l.) generate differential selection pressures, not only on body size, but also on colour traits. This seems unlikely as a general explanation of colour differentiation in *Z. borbonicus*, because the two extreme colour variants (all-grey and all-brown morphs) both occur at high elevations.

Second, the density and height of vegetation cover are highly heterogeneous on Réunion (Strasberg *et al.*, 2005; Thébaud *et al.*, 2009). The resulting variation in lighting conditions may have differentially selected plumage colour for crypsis or optimal foraging in the different environments (Gomez & Théry, 2004), although we do not have data to formally test these hypotheses. *Z. borbonicus* has few predators on Réunion, and most are recent arrivals on the island (Gill, 1973). The only predator against which crypsis could have evolved is the Réunion harrier (*Circus maillardi*). However, this species is a very generalist predator, only preying occasionally on small birds like white-eyes (Safford & Hawkins, 2013). Thus, selection for crypsis seems unlikely to have influenced plumage colour diversification in *Z. borbonicus*.

Third, habitat heterogeneity is also generally associated with different species community assemblages, notably parasitic communities (Blondel *et al*., 2006; Vanbergen *et al.*, 2006). For example, it has been demonstrated that the composition of blood parasite communities in *Z. borbonicus* varies according to climate within the island of Réunion (Cornuault *et al.*, 2013). Such spatial patterns in bird parasites might exert spatially-structured selection pressures on colour as the expression of melanic pigments has been correlated to pathogen resistance (Mackintosh, 2001; Wilson *et al.*, 2001; Ducrest, Keller & Roulin, 2008). While this hypothesis may seem plausible, testing whether plumage colour differences reflects distinct parasite or pathogen communities will be a necessary step before any conclusion can be reached on the role of parasites on plumage colour diversification in *Z. borbonicus*.

Conclusions

Genetic drift and founder events have long been considered as the main processes of evolution on islands (Barton, 1996) but the role of divergent selection as a strong driver of phenotypic diversification on islands has recently gained support (Blondel *et al*., 1999; Clegg *et al.*, 2008; Harmon *et al.*, 2008; Losos & Ricklefs, 2009). Despite the unambiguous influence of selection in shaping phenotypic variation in *Z. borbonicus*, our results do not

preclude a joint effect of neutral and demographic processes. Dispersal and gene flow appear to be limited in *Z. borbonicus*, with low levels of historical and/or contemporary gene flow among populations, unless very close geographically (<10km) (Bertrand *et al.*, 2014). The key implication is that it is the combination of reduced dispersal and divergent selection that seem to explain why *Z. borbonicus* was able to differentiate into multiple geographic forms within Réunion. Therefore, our results echo Frank Gill's prophetic suggestion that "a delicate balance between gene flow and opposing selection forces must be involved" [to explain the divergence of the different phenotypic forms] (Gill, 1973). We conclude that divergent natural selection and reduced dispersal are the dominant mechanisms explaining phenotypic divergence in morphological and plumage colour traits at a very small spatial scale in *Z. borbonicus*, one of the most extraordinary case of microgeographical variation in plumage colouration in birds. Given that hybrid zones arise where the different forms come into contact, as happens between parapatric lowland forms or between lowland and highland forms, it seems possible that *Z. borbonicus* may also represent one of the few examples in support of the divergence-with-gene-flow model (Maynard-Smith, 1966; Felsenstein, 1976; Rice & Hostert, 1993). Investigation of the history and degree of isolation of the different *Z. borbonicus* forms, by making use of the new genomic tools that are becoming available (Bourgeois *et al*., 2012), should quickly provide extensive new data to test the conclusions of our study.

Acknowledgments

We thank Frank Gill for kindly providing morphometric data, Janet Hinshaw for providing access to the collections and facilitating our visit to the University of Michigan Museum of Zoology, Robert Payne for permitting spectrophotometric measurements on the museum specimens, Luke Harmon for his assistance with the analyses, and Sergine Ponsard and Juli Broggi for useful comments during the elaboration of this work. We also thank Jacques Blondel and one anonymous reviewer for valuable comments on an earlier version of this manuscript. JB, YB, BD and JC were supported by MESR (Ministère de l'Enseignement Supérieur et de la Recherche) PhD scholarships. The research was supported by Agence Française pour le Développement grants to CT, the Institut Français de la Biodiversité (IFB), the Fondation pour la Recherche sur la Biodiversité (FRB) through its Centre for Synthesis

and Analysis of Biodiversity (CESAB), and the 'Laboratoire d'Excellence' TULIP (ANR-10-LABX-41).

References

Ackermann RR, Cheverud JM. 2002. Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. *American Journal of Physical Anthropology* **117:** 260-271.

Anderson MB. 1994. *Sexual selection*. Princeton, NJ: Princeton University Press.

Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26:** 32-46.

Baker MC, Baker AEM. 1990. Reproductive behavior of female buntings: isolating mechanisms in a hybridizing pair of species. *Evolution* **44:** 332-338.

Barton NH. 1996. Natural selection and random genetic drift as causes of evolution on islands. *Philosophical Transactions of the Royal Society B* **351:** 785-795.

Bergmann C. 1847. Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien* **3:** 595-708.

Bertrand JAM, Bourgeois YXC, Delahaie B, Duval T, García-Jiménez R, Cornuault J, Heeb P, Milá B, Pujol B, Thébaud C. 2014. Extremely reduced dispersal and gene flow in an island bird. *Heredity* **112:** 190-196.

Blondel J, Dias PC, Perret P, Maistre M, Lambrechts MM. 1999. Selection-based biodiversity at a small spatial scale in a low-dispersing insular bird. *Science* **285**: 1399-1401.

Blondel J. 2000. Evolution and ecology of birds on islands: trends and prospects. *Vie et Milieu* **50:** 205-220.

Blondel J, Thomas DW, Charmantier A, Perret P, Bourgault P, Lambrechts MM. 2006. A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *Bioscience* **56**: 661-673.

Bourgeois YXC. 2013. Génétique évolutive d'un cas extrême de polymorphisme de la coloration du plumage chez un oiseau insulaire, *Zosterops borbonicus* (Zosteropideae). Unpublished D. Phil. Thesis, University Paul Sabatier.

Bourgeois YXC, Bertrand JAM, Thébaud C, Milá B. 2012. Investigating the role of the Melanocortin-1 receptor gene in an extreme case of microgeographical variation in the pattern of melanin-based plumage pigmentation. *PLoS One* **7:** e50906.

Clegg SM, Frentiu FD, Kikkawa J, Tavecchia G, Owens IPF. 2008. 400 years of phenotypic change in an island bird: heterogeneity of selection over three microevolutionary timescales. *Evolution* **62:** 2393-2410.

Cornuault J, Khimoun A, Harrigan RJ, Bourgeois YXC, Milá B, Thébaud C, Heeb P. 2013. The role of ecology in the geographical separation of blood parasites infecting an insular bird. *Journal of Biogeography* **40:** 1313-1323.

Covas R. 2011. Evolution of reproductive life histories in island birds worldwide. *Proceedings of the Royal Society of London B* **279:** 1531-1537.

Coyne JA, Price TD. 2000. Little evidence for sympatric speciation in island birds. *Evolution* **54:** 2166-2171.

Diamond JM, Gilpin ME, Mayr E. 1976. Species-distance relation for birds of the Solomon Archipelago, and the paradox of great speciators. *Proceedings of the National Acadamy of Sciences USA* **73:** 2160-2164.

Ducrest A-L, Keller L, Roulin A. 2008. Pleiotropy in the melanocortin system, colouration and behavioural syndromes. *Trends in Ecology and Evolution* **23:** 502-510.

Endler JA. 1973. *Geographic Variation, Speciation, and Clines.* Princeton, NJ: Princeton University Press.

Endler JA, Mielke PWJ. 2005. Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society* **86:** 405-431.

Felsenstein J. 1976. The theoretical population genetics of variable selection and migration. *Annual Review of Genetics* **10:** 253-280.

Gill FB. 1973. Intra-island variation in the Mascarene white-eye *Zosterops borbonica*. *Ornithological Monographs* **12:** 1-66.

Gill FB, Donsker D. 2014. *IOC World Bird List (v 4.3)*. Available at: http://www.worldbirdnames.org/ioc-lists/crossref/

Goldsmith TH. 1990. Optimization, constraint, and history in the evolution of eyes. *Quarterly Review of Biology* **65:** 281-322.

Gomez D, Théry M. 2004. Influence of ambient light on the evolution of colour signals: comparative analysis of a Neotropical rainforest bird community. *Ecology Letters* **7:** 279-284.

Grant PR. 1998. *Evolution on Islands.* New York: Oxford University Press.

Harmon LJ, Gibson R. 2006. Multivariate phenotypic evolution among island and mainland populations of the ornate day gecko, *Phelsuma ornata*. *Evolution* **60:** 2622-2632.

Harmon LJ, Melville J, Larson A, Losos JB. 2008. The role of geography and ecological opportunity in the diversification of day geckos (*Phelsuma*)**.** *Systematic Biology* **57:** 562-573.

Hart NS, Partridge JC, Cuthill IC, Bennett ATD. 2000. Visual pigments, oil droplets, ocular media and cone photoreceptor distribution in two species of passerine bird: the blue tit (*Parus caeruleus* L.) and the blackbird (*Turdus merula* L.). *Journal of Comparative Physiology A* **186:** 375-387.

Hill GE, McGraw KJ. 2006. *Bird Colouration, Volume 2: Function and Evolution*. Cambridge, MA: Harvard University Press.

Hille SM., Cooper CB. 2014. Elevational trends in life histories: revising the pace-of-life framework. *Biological Reviews* DOI: 10.1111/brv.12106

Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6:** 65-70.

Komdeur J., Richardson DS, Piersma T., Kraaijeveld K. 2004. Why Seychelles warblers fail to recolonise nearby islands: were they selected for reduced flight performance? *Ibis* **146:** 298-302.

Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33:** 402-416.

Lenormand T. 2002. Gene flow and the limits to natural selection. *TRENDS in Ecology & Evolution* **17:** 183-189.

Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* **457:** 830-836.

Mackintosh JA. 2001. The antimicrobial properties of melanocytes, melanosomes and melanin and the evolution of black skin. *Journal of Theoretical Biology* **211:** 101-113.

Maia R, Eliason CH, Bitton P-P, Doucet SM, Shawkey MD. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution* **4:** 906-913.

Marroig G, Cheverud J. 2004. Cranial evolution in sakis (Pithecia, Platyrrhini) I: Interspecific differentiation and allometric patterns. *American Journal of Physical Anthropology* **125:** 266-278.

Maynard-Smith J. 1966. Sympatric speciation. *American Naturalist* **100:** 637-650.

Mayr E. 1942. *Systematics & the Origin of Species.* New York, NY: Columbia University Press.

Mayr E, Diamond JM. 2001. *The Birds of Northern Melanesia. Speciation, ecology and biogeography*. Ney York, NY: Cambridge University Press.

McGraw KJ, Safran RJ, Wakamatsu K. 2005. How feather colour reflects its melanin content. *Functional Ecology* **19:** 816-821.

McNab BK. 2010. Geographic and temporal correlations of mammalian size reconsidered: a resource rule. *Oecologia* **164**:13-23.

Milá B, Warren BH, Heeb P, Thébaud C. 2010. The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology* **10:** 158.

Moyle RG, Filardi CE, Smith CE, Diamond J. 2009. Explosive Pleistocene diversification and hemispheric expansion of a 'great speciator'. *Proceedings of the National Acadamy of Sciences USA* **106:** 1863-1868.

Mundy NI, Badcock NS, Hart T, Scribner K, Janssen K, Nadeau NJ. 2004. Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science* **303:** 1870-1873.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2008. *Vegan: community ecology package*. Available at: http://cc.oulu.fi/~jarioksa/softhelp/vegan.html

Porlier M, Garant D, Perret P, Charmantier A. 2012. Habitat-linked population genetic differentiation in the blue tit *Cyanestes caeruleus*. *Journal of Heredity* **103**: 781-791.

Price T. 2008. *Speciation in birds.* Boulder, CO: Roberts and Co.

Price T, Bontrager A. 2001. Evolutionary genetics: The evolution of plumage patterns. *Current Biology* **11:** R405-R408.

R Development Core Team. 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rice WR, Hostert EE. 1993. Laboratory experiments on speciation: What have we learned in 40 years? *Evolution*, **47:** 1637-1653.

Riley PA. 1997. Melanin. *International Journal of Biochemistry & Cell Biology* **29:** 1235-1239.

Roulin A, Bize P. 2007. Sexual selection in genetic colour-polymorphic species: a review of experimental studies and perspectives. *Journal of Ethology* **25:** 99-105.

Royston JP. 1982. An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society C* **31:** 115-124.

Safford R, Hawkins F. 2013. *The Birds of Africa: Volume VIII: The Malagasy Region: Madagascar, Seychelles, Comoros, Mascarenes.* London, ULK: A&C Black.

Stoddard MC, Prum RO. 2008. Evolution of avian plumage colour in a tetrahedral colour space: A phylogenetic analysis of new world buntings. *American Naturalist* **171:** 755-776.

Strasberg D, Rouget M, Richardson DM, Baret S, Dupont J, Cowling RM. 2005. An assessment of habitat diversity and transformation on La Reunion Island (Mascarene Islands, Indian Ocean) as a basis for identifying broad-scale conservation priorities. *Biodiversity and Conservation* **14:** 3015-3032.

Takeuchi S, Suzuki H, Yabuuchi M, Takahashi S. 1996. A possible involvement of melanocortin 1-receptor in regulating feather colour pigmentation in the chicken. *Biochimica Et Biophysica Acta-Gene Structure and Expression* **1308:** 164-168.

Thébaud C, Warren BH, Cheke AC, Strasberg D. 2009. Mascarene Islands, biology. In: Gillespie RG, Clague DA, eds. *Encyclopedia of islands*. Berkeley, CA: University of California Press, 612-619.

Theron E, Hawkins K, Bermingham E, Ricklefs RE, Mundy NI. 2001. The molecular basis of an avian plumage polymorphism in the wild: A melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. *Current Biology* **11:** 550-557.

Uy JAC, Moyle RG, Filardi CE. 2009. Plumage and song differences mediate species recognition between incipient flycatcher species of the solomon islands. *Evolution* **163:** 153-164.

Vanbergen AJ, Hails RS, Watt AD, Jones TH. 2006. Consequences for host-parasitoid interactions of grazing-dependent habitat heterogeneity. *Journal of Animal Ecology* **75:** 789-801.

Vorobyev M, Osorio D. 1998. Receptor noise as a determinant of colour thresholds. *Proceedings of the Royal Society of London B* **265:** 351-358.

Vorobyev M, Osorio D, Bennett ATD, Marshall NJ, Cuthill IC. 1998. Tetrachromacy, oil droplets and bird plumage colours. *Journal of Comparative Physiology A* **183:** 621-633.

Ward JM, Blount JD, Ruxton GD, Houston DC. 2002. The adaptive significance of dark plumage for birds in desert environments. *Ardea* **90:** 311-323.

Wilson K, Cotter SC, Reeson AF, Pell JK. 2001. Melanism and disease resistance in insects. *Ecology Letters* **4:** 637-649.

Supporting information

Table S1: Museum specimens used in the analyses.

University of Michigan Museum of Zoology Catalog Number	Collector Number	Locality (corresponding to Gill 1973)	Colour Morph	Used for Morphological Analyses	Used for Colour Analyses
213298	FBG 400	9	GHB	Yes	No
213299	FBG 401	9	GHB	Yes	No
213493	FBG 414	31	GHB	Yes	No
213495	FBG 416	31	GHB	Yes	No
213496	FBG 417	31	GHB	Yes	No
213497	FBG 418	31	GHB	Yes	No
213498	FBG 419	31	GHB	Yes	No
213328	FBG 570	10	GHB	Yes	No
210372	FBG 571	10	GHB	Yes	No
213330	FBG 572	10	GHB	Yes	No
213331	FBG 573	10	GHB	Yes	No
213334	FBG 576	10	GHB	Yes	No
213567	FBG 646	45	GHB	Yes	No
213572	FBG 651	45	GHB	Yes	No
213573	FBG 652	45	GHB	Yes	No
213574	FBG 653	45	GHB	Yes	No
213576	FBG 655	45	GHB	Yes	No
213579	FBG 660	46	GHB	Yes	No
213580	FBG 661	46	GHB	Yes	No
213582	FBG 663	46	GHB	Yes	No
213586	FBG 667	46	GHB	Yes	No
213587	FBG 668	46	GHB	Yes	No
213588	FBG 669	46	GHB	Yes	No
213315	FBG 732	14	GHB	Yes	No
213314	FBG 733	14	GHB	Yes	No
213312	FBG 735	14	GHB	Yes	No
213308	FBG 739	14	GHB	Yes	No
213322	FBG 847	12	GHB	Yes	No
213324	FBG 849	12	GHB	Yes	No
213326	FBG 851	12	GHB	Yes	No
213327	FBG 852	12	GHB	Yes	No
213589	FBG 931	85	GHB	Yes	No
213594	FBG 936	85	GHB	Yes	No
213248	FBG 942	2	GHB	Yes	No
213251	FBG 945	2	GHB	Yes	No
213223	FBG 966	1	GHB	Yes	No
213226	FBG 969	1	GHB	Yes	No
213227	FBG 970	1	GHB	Yes	No
213228	FBG 971	1	GHB	Yes	No
213288	FBG 980	11	GHB	Yes	Yes
213290	FBG 982	11	GHB	Yes	No
213295	FBG 987	11	GHB	Yes	Yes
213296	FBG 988	11	GHB	Yes	Yes
213297	FBG 989	11	GHB	Yes	No
213929	FBG 1080	96	GHB	Yes	Yes
213934	FBG 1085	96	GHB	Yes	No
213936	FBG 1087	96	GHB	Yes	No
213937	FBG 1088	96	GHB	Yes	No
213938	FBG 1089	96	GHB	Yes	Yes
213451	FBG 1104	99	GHB	Yes	No
213455	FBG 1108	99	GHB	Yes	No
213435	FBG 1174	81	GHB	Yes	Yes
213442	FBG 1181	81	GHB	Yes	Yes
213426	FBG 1207	25	GHB	Yes	No

210080	GI 919	9	GHB	Yes	No
NA	GI 884	10	GHB	Yes	No
NA	GI 885	10	GHB	Yes	No
NA	GI 886	10	GHB	Yes	No
NA	GI 896	12	GHB	Yes	No
210085	GI 900	12	GHB	Yes	No
213617	FBG 672	47	BNB	Yes	No
213599	FBG 702	48	BNB	Yes	Yes
213605	FBG 708	48	BNB	Yes	Yes
213606	FBG 709	48	BNB	Yes	Yes
213608	FBG 711	48	BNB	Yes	Yes
213609	FBG 712	48	BNB	Yes	Yes
213611	FBG 714	47	BNB	Yes	No
213613	FBG 716	47	BNB	Yes	No
213615	FBG 718	47	BNB	Yes	No
213510	FBG 740	34	BNB	Yes	Yes
213511	FBG 741	34	BNB	Yes	Yes
213512	FBG 742	34	BNB	Yes	Yes
213515	FBG 745	34	BNB	Yes	Yes
213516	FBG 746	34	BNB	Yes	No
213518	FBG 748	34	BNB	Yes	Yes
213519	FBG 749	34	BNB	Yes	No
213897	FBG 887	86	BNB	Yes	No
213899	FBG 889	86	BNB	Yes	No
213901	FBG 891	86	BNB	Yes	No
214659	FBG 438	30	HBHB	Yes	No
214661	FBG 440	30	HBHB	Yes	Yes
213862	FBG 466	73	HBHB	Yes	No
213773	FBG 492	63	HBHB	Yes	No
213779	FBG 500	63	HBHB	Yes	No
213780	FBG 501	63	HBHB	Yes	No
213781	FBG 502	63	HBHB	Yes	No
213783	FBG 504	63	HBHB	Yes	No
213785	FBG 506	63	HBHB	Yes	No
213787	FBG 508	63	HBHB	Yes	No
213798	FBG 509	64	HBHB	Yes	No
213800	FBG 511	64	HBHB	Yes	No
213801	FBG 512	64	HBHB	Yes	No
213803	FBG 514	64	HBHB	Yes	No
213405	FBG 602	23	HBHB	Yes	Yes
213408	FBG 605	23	HBHB	Yes	Yes
213337	FBG 623	18	HBHB	Yes	No
213344	FBG 630	18	HBHB	Yes	No
213367	FBG 637	19	HBHB	Yes	No
213371	FBG 641	19	HBHB	Yes	No
213372	FBG 642	19	HBHB	Yes	No
213540	FBG 763	39	HBHB	Yes	Yes
213546	FBG 769	39	HBHB	Yes	Yes
213522	FBG 789	8	HBHB	Yes	No
213527	FBG 794	8	HBHB	Yes	No
213550	FBG 804	41	HBHB	Yes	Yes
213552	FBG 806	41	HBHB	Yes	Yes
213555	FBG 809	41	HBHB	Yes	Yes
213688	FBG 818	58	HBHB	Yes	No
213790	FBG 947	63	HBHB	Yes	No
213564	FBG 955	41	HBHB	Yes	No
213503	FBG 1010	40	HBHB	Yes	No
213506	FBG 1013	40	HBHB	Yes	No
213507	FBG 1014	40	HBHB	Yes	No
213853	FBG 1023	72	HBHB	Yes	No
213855	FBG 1025	72	HBHB	Yes	No
213856	FBG 1026	72	HBHB	Yes	No
213857	FBG 1027	72	HBHB	Yes	No
213866	FBG 1030	73	HBHB	Yes	No

213874	FBG 1042	74	HBHB	Yes	No
213875	FBG 1043	74	HBHB	Yes	No
213878	FBG 1046	74	HBHB	Yes	No
213819	FBG 1056	98	HBHB	Yes	No
213822	FBG 1060	98	HBHB	Yes	No
213750	FBG 1066	62	HBHB	Yes	No
213751	FBG 1067	62	HBHB	Yes	No
213752	FBG 1068	62	HBHB	Yes	No
213754	FBG 1070	62	HBHB	Yes	No
213758	FBG 1074	62	HBHB	Yes	No
213759	FBG 1075	62	HBHB	Yes	No
213793	FBG 1098	63	HBHB	Yes	No
213695	FBG 1131	58	HBHB	Yes	No
213697	FBG 1133	58	HBHB	Yes	No
213700	FBG 1136	58	HBHB	Yes	No
213711	FBG 1149	58	HBHB	Yes	No
213712	FBG 1150	58	HBHB	Yes	No
213717	FBG 1155	58	HBHB	Yes	No
213476	FBG 1190	29	HBHB	Yes	No
213761	FBG 1219	62	HBHB	Yes	No
213764	FBG 1222	62	HBHB	Yes	No
213765	FBG 1223	62	HBHB	Yes	No
213722	FBG 1256	58	HBHB	Yes	No
210089	GI 975	58	HBHB	Yes	No
NA	GI 989	58	HBHB	Yes	No
213465	FBG 411	27	LBHB	Yes	No
213485	FBG 443	36	LBHB	Yes	No
213663	FBG 449	53	LBHB	Yes	No
213664	FBG 450	53	LBHB	Yes	No
213665	FBG 451	53	LBHB	Yes	No
213638	FBG 457	54	LBHB	Yes	No
213633	FBG 750	44	LBHB	Yes	No
213623	FBG 752	44	LBHB	Yes	Yes
213628	FBG 756	44	LBHB	Yes	Yes
213627	FBG 757	44	LBHB	Yes	Yes
213630	FBG 759	44	LBHB	Yes	Yes
213631	FBG 760	44	LBHB	Yes	No
213632	FBG 761	44	LBHB	Yes	No
213634	FBG 762	44	LBHB	Yes	No
213316	FBG 781	49	LBHB	Yes	No
213537	FBG 783	49	LBHB	Yes	No
213538	FBG 784	49	LBHB	Yes	No
213650	FBG 822	51	LBHB	Yes	No
213651	FBG 823	51	LBHB	Yes	No
213654	FBG 826	51	LBHB	Yes	No
213658	FBG 830	51	LBHB	Yes	No
213558	FBG 864	41	LBHB	Yes	No
213671	FBG 923	55	LBHB	Yes	No
213672	FBG 924	55	LBHB	Yes	No
213927	FBG 964	94	LBHB	Yes	No
213928	FBG 965	94	LBHB	Yes	No
213806	FBG 991	95	LBHB	Yes	No
213808	FBG 993	95	LBHB	Yes	No
210374	FBG 995	95	LBHB	Yes	No
213813	FBG 998	95	LBHB	Yes	No
213738	FBG 999	61	LBHB	Yes	No
213739	FBG 1000	61	LBHB	Yes	No
213740	FBG 1001	61	LBHB	Yes	No
213741	FBG 1002	61	LBHB	Yes	No
213742	FBG 1003	61	LBHB	Yes	No
213744	FBG 1005	61	LBHB	Yes	No
213745	FBG 1006	61	LBHB	Yes	No
213894	FBG 1051	97	LBHB	Yes	No
213886	FBG 1092	97	LBHB	Yes	No

213887	FBG 1093	97	LBHB	Yes	No
213891	FBG 1097	97	LBHB	Yes	No
213457	FBG 1110	27	LBHB	Yes	No
213462	FBG 1115	27	LBHB	Yes	No
213490	FBG 1117	33	LBHB	Yes	No
213681	FBG 1160	56	LBHB	Yes	No
213815	FBG 1216	95	LBHB	Yes	No
213685	FBG 1262	56	LBHB	Yes	No
210076	GI 924	51	LBHB	Yes	No
NA	GI 926	51	LBHB	Yes	No
NA	GI 927	51	LBHB	Yes	No
213469	FBG 407	29	G	Yes	No
213494	FBG 415	31	G	Yes	No
213481	FBG 430	35	G	Yes	No
213859	FBG 461	73	G	Yes	Yes
213864	FBG 469	73	G	Yes	Yes
213844	FBG 471	71	G	Yes	No
213840	FBG 482	70	G	Yes	Yes
213775	FBG 496	63	G	Yes	No
213778	FBG 499	63	G	Yes	No
213784	FBG 505	63	G	Yes	No
213802	FBG 513	64	G	Yes	No
213281	FBG 529	6	G	Yes	No
213284	FBG 532	6	G	Yes	No
213286	FBG 535	6	G	Yes	No
213730	FBG 547	60	G	Yes	No
213731	FBG 548	60	G	Yes	No
213359	FBG 587	16	G	Yes	No
213403	FBG 600	23	G	Yes	No
213404	FBG 601	23	G	Yes	No
213412	FBG 609	23	G	Yes	No
213396	FBG 613	24	G	Yes	No
213363	FBG 633	19	G	Yes	No
213365	FBG 635	19	G	Yes	No
213375	FBG 684	20	G	Yes	No
213543	FBG 766	39	G	Yes	Yes
213528	FBG 774	49	G	Yes	No
213524	FBG 791	8	G	Yes	No
213525	FBG 792	8	G	Yes	No
213690	FBG 820	58	G	Yes	No
213559	FBG 833	41	G	Yes	No
213905	FBG 907	87	G	Yes	No
213668	FBG 920	55	G	Yes	No
213669	FBG 921	55	G	Yes	No
213851	FBG 1021	72	G	Yes	No
213869	FBG 1033	73	G	Yes	No
213870	FBG 1034	73	G	Yes	No
213872	FBG 1036	73	G	Yes	No
213755	FBG 1071	62	G	Yes	No
213706	FBG 1144	58	G	Yes	No
213715	FBG 1153	58	G	Yes	No
NA	FBG 1193	29	G	Yes	No
213796	FBG 1270	63	G	Yes	No
NA	GI 512	41	G	Yes	No
210109	GI 519	42	G	Yes	Yes
210088	GI 971	58	G	Yes	No
210090	GI 973	58	G	Yes	No
213626	FBG 755	44	LBHB	No	Yes
210106	GI 1009	41	G	No	Yes
213545	FBG 768	39	G	No	Yes
213438	FBG 1177	81	GHB	No	Yes
213440	FBG 1179	81	GHB	No	Yes
213445	FBG 1184	81	GHB	No	Yes
213547	FBG 770	39	HBHB	No	Yes

213829	FBG 676	68	LBHB	No	Yes
213830	FBG 677	68	LBHB	No	Yes
213833	FBG 680	68	LBHB	No	Yes
213834	FBG 681	68	LBHB	No	Yes
213826	FBG 558	68	LBHB	No	Yes
213847	FBG 478	71	G	No	Yes
213865	FBG 1029	73	G	No	Yes
213400	FBG 617	24	HBHB	No	Yes
213260	FBG 1200	NA	HBHB	No	Yes

Table S2: PC loadings for morphological traits.

Variables	PC1	PC2	PC3	PC4
Wing length	0,92	0,01	0,25	0,30
Tail length	0,91	0,00	0,29	-0,29
Tarsus length	0,67	0,44	-0,60	-0,01
Bill length	-0,33	0,90	0,29	0,00
Variance Explained (%)	0,56	0,25	0,15	0,04

Table S3: PC loadings for plumage colour traits.

	Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20	PC21	PC22	PC23	PC24	PC25	PC26	PC27	PC28	PC29	PC30	PC31	PC32
Breast	Brilliance	-0,15	-0,88	0,11	0,16	0,05	-0,14	-0,06	0,13	0,01	0,09	-0,14	0,01	0,03	0,07	-0,16	0,11	-0,08	-0,03	0,04	-0,06	0,01	0,07	-0,04	0,05	-0,03	0,12	0,03	0,04	0,04	0,01	0,00	0,01
	θ	-0,39	0,21	-0,09	0,55	0,18	-0,34	0,31	-0,15	-0,06	0,26	-0,23	0,14	-0,22	-0,05	-0,14	-0,05	0,06	0,00	-0,03	-0,04	-0,02	-0,02	0,07	-0,06	0,00	-0,01	0,02	-0,01	-0,01	0,00	0,01	0,00
	φ	0,47	0,26	0,52	0,32	-0,10	-0,09	-0,26	-0,07	-0,28	0,20	0,13	-0,14	0,03	0,04	0,11	0,18	0,16	-0,05	-0,01	-0,06	-0,09	0,07	0,04	-0,01	0,00	0,02	-0,03	-0,03	0,00	0,00	0,01	0,00
	r_a	0,11	0,52	-0,42	0,14	0,03	-0,32	0,14	-0,26	0,45	-0,17	-0,03	-0,14	0,12	-0,10	0,12	0,09	-0,02	-0,06	-0,05	-0,07	0,02	0,07	0,02	-0,06	-0,02	0,03	0,01	0,02	0,01	-0,02	0,00	0,00
Rump	Brilliance	-0,14	-0,66	0,08	-0,27	-0,27	-0,36	0,22	-0,06	0,05	0,08	-0,22	-0,32	-0,03	0,04	0,10	-0,04	0,04	-0,03	0,12	0,05	0,02	0,01	-0,05	0,01	0,07	-0,02	-0,01	-0,05	0,00	0,02	-0,01	-0,01
	θ	-0,70	0,28	0,04	0,14	-0,07	-0,43	0,12	-0,16	-0,14	0,07	0,14	0,05	0,10	0,20	0,17	-0,11	-0,03	0,08	-0,14	-0,01	0,05	-0,04	-0,02	0,09	-0,02	0,04	-0,01	0,00	0,03	0,01	-0,01	0,00
	φ	0,43	0,10	0,63	0,44	0,03	0,25	0,20	0,02	0,16	0,05	-0,02	0,06	0,02	0,15	0,11	0,02	-0,15	0,12	0,07	-0,02	-0,06	0,02	0,01	-0,01	0,04	-0,01	0,07	-0,05	0,00	-0,04	-0,02	0,01
	r_a	-0,32	0,00	-0,43	-0,26	0,06	-0,62	-0,21	0,21	-0,04	-0,02	0,32	0,04	-0,07	0,12	-0,08	0,05	-0,11	0,00	0,08	0,01	-0,02	0,00	0,03	-0,07	0,05	-0,01	0,02	-0,01	-0,02	-0,02	0,00	0,00
Back	Brilliance	-0,55	-0,21	0,47	-0,17	-0,45	0,02	0,00	0,10	0,24	0,24	0,11	0,13	0,08	-0,16	0,01	0,02	0,01	0,00	-0,04	0,04	-0,03	-0,01	0,01	0,02	0,07	0,00	0,03	0,03	0,02	-0,02	0,04	-0,04
	θ	-0,72	-0,02	-0,12	0,34	-0,34	0,24	-0,08	-0,29	0,06	0,06	0,08	0,13	-0,03	0,00	0,08	-0,01	-0,03	-0,08	0,09	0,10	0,08	-0,01	0,02	-0,05	0,03	0,08	-0,03	-0,03	-0,02	0,01	0,02	0,03
	φ	0,73	-0,12	0,38	0,23	0,17	-0,19	0,15	0,25	0,15	-0,10	0,06	0,15	0,09	-0,01	0,02	-0,07	0,06	0,10	-0,02	-0,01	0,03	0,07	-0,05	-0,01	0,01	0,04	-0,02	-0,02	-0,07	0,04	0,03	-0,01
	r_a	0,04	-0,21	-0,66	0,41	0,14	0,16	-0,34	-0,32	-0,04	-0,06	-0,11	0,01	-0,05	0,18	-0,02	0,04	-0,03	0,00	-0,05	0,00	0,03	0,04	-0,07	0,03	0,07	-0,03	0,02	0,00	-0,03	-0,01	0,03	-0,03
Flank	Brilliance	-0,48	-0,08	0,69	0,01	0,24	-0,04	-0,30	-0,10	0,28	0,01	0,04	0,03	-0,07	0,05	0,00	-0,04	-0,11	-0,10	-0,04	0,02	-0,03	0,05	0,02	-0,02	-0,05	-0,06	0,00	-0,03	0,03	0,06	-0,01	-0,01
	θ	-0,74	0,39	0,46	0,00	0,12	0,17	0,04	-0,02	-0,04	0,02	-0,02	-0,05	-0,11	0,05	0,03	-0,03	-0,05	0,01	0,04	0,03	-0,01	0,05	0,00	-0,04	-0,03	0,07	-0,04	0,06	-0,05	-0,01	-0,05	-0,04
	φ	0,68	-0,44	-0,37	0,29	-0,15	-0,03	-0,11	-0,09	0,03	-0,10	0,07	0,13	0,10	-0,07	-0,07	-0,05	0,11	-0,03	-0,03	0,06	0,00	-0,02	0,04	-0,03	0,02	0,05	0,03	-0,05	0,02	0,01	-0,07	-0,03
	r_a	0,10	-0,36	-0,68	0,28	-0,43	-0,01	-0,07	-0,04	0,13	0,11	0,07	0,14	-0,07	-0,01	0,10	0,03	0,02	0,14	0,14	-0,11	-0,04	-0,01	-0,02	0,01	-0,07	-0,05	-0,03	0,04	0,01	0,03	-0,02	-0,01
Crown	Brilliance	-0,33	-0,84	0,16	0,02	-0,10	-0,05	0,07	-0,13	-0,07	-0,02	0,08	-0,08	-0,07	-0,10	-0,04	0,13	0,01	0,15	-0,08	0,13	0,07	0,06	-0,01	-0,03	-0,11	-0,03	0,02	-0,03	-0,01	-0,02	0,00	0,00
	θ	-0,65	0,33	0,18	0,48	0,04	-0,13	-0,20	0,11	-0,09	-0,20	-0,08	-0,10	0,05	-0,15	0,03	-0,01	0,02	0,05	0,12	0,03	0,06	0,02	0,12	0,11	0,01	-0,02	0,06	0,02	-0,01	0,01	0,00	0,00
	φ	0,70	-0,37	0,26	0,08	0,13	0,14	0,02	-0,06	0,14	0,03	0,27	-0,22	-0,25	-0,01	0,00	-0,05	0,01	-0,04	0,03	-0,10	0,11	-0,10	0,03	0,04	-0,03	0,03	0,02	-0,01	0,00	-0,01	0,01	-0,02
	r_a	-0,13	0,77	-0,13	0,05	-0,16	0,00	-0,32	0,31	0,03	0,06	-0,22	-0,07	-0,05	-0,13	0,03	0,09	-0,10	0,11	-0,05	-0,04	0,06	-0,07	-0,04	-0,04	-0,01	0,04	-0,03	-0,07	0,01	0,00	0,00	-0,01
Throat	Brilliance	0,19	-0,87	0,05	0,23	-0,01	0,10	-0,04	0,17	-0,06	0,00	-0,06	-0,13	0,07	0,11	0,09	0,00	-0,04	0,07	-0,11	-0,01	0,06	-0,03	0,11	-0,11	0,06	-0,02	-0,02	0,06	0,01	0,02	0,00	0,00
	θ	-0,34	-0,55	-0,18	0,41	0,32	0,02	0,34	0,10	-0,11	-0,16	0,10	0,06	-0,05	-0,15	0,08	0,17	-0,13	-0,08	-0,02	0,00	-0,06	-0,04	0,00	0,05	0,04	-0,01	-0,07	-0,02	0,01	0,00	0,00	-0,01
	φ	0,45	0,41	0,57	0,27	-0,21	-0,14	0,10	0,01	0,09	0,01	-0,03	0,09	0,12	0,16	-0,18	0,15	0,01	-0,07	0,05	0,02	0,15	-0,04	-0,01	0,02	-0,01	-0,06	-0,06	0,01	0,00	-0,01	-0,01	0,00
	r_a	-0,33	-0,11	-0,34	0,07	0,66	0,12	-0,04	-0,11	0,05	0,37	0,10	-0,14	0,32	-0,11	-0,07	-0,02	-0,03	0,05	0,06	-0,01	0,03	-0,02	-0,01	0,00	0,00	-0,01	-0,01	-0,01	-0,02	0,00	-0,01	0,00
Nape	Brilliance	-0,59	-0,46	0,49	-0,17	-0,10	-0,08	-0,19	-0,19	0,11	-0,04	-0,02	0,06	-0,05	-0,13	-0,05	0,05	0,03	0,02	-0,10	-0,13	0,00	-0,04	-0,03	0,03	0,07	-0,01	0,01	0,01	-0,05	0,00	-0,04	0,03
	θ	-0,54	0,22	0,07	0,59	-0,31	0,09	0,03	0,04	0,02	-0,14	0,18	-0,23	0,00	-0,03	-0,23	-0,17	0,00	0,07	-0,01	-0,04	-0,05	0,06	-0,04	-0,02	0,04	-0,01	-0,05	-0,01	0,03	-0,01	0,00	0,01
	φ	0,64	0,00	0,26	0,33	0,40	-0,26	-0,25	0,05	0,02	0,08	0,01	0,07	-0,16	-0,15	0,12	-0,08	0,04	0,00	0,02	0,10	0,04	0,02	-0,09	-0,01	0,04	-0,01	-0,02	0,05	0,04	-0,03	-0,01	0,02
	r_a	0,38	0,15	-0,55	0,41	-0,33	0,04	0,05	0,34	0,08	0,24	0,04	-0,13	-0,05	-0,01	0,01	-0,03	-0,08	-0,14	-0,11	0,07	-0,01	0,04	-0,01	0,06	-0,02	-0,02	0,03	0,02	-0,05	0,01	-0,02	0,01
Belly	Brilliance	-0,16	-0,77	0,24	0,32	-0,07	-0,17	-0,25	0,11	0,03	-0,09	-0,15	0,02	0,14	0,02	0,05	-0,14	-0,01	-0,09	0,02	-0,02	-0,07	-0,09	0,01	-0,01	-0,09	-0,01	-0,03	-0,02	-0,03	-0,05	0,02	0,00
	θ	-0,83	0,12	0,16	0,35	0,02	0,07	0,11	0,15	-0,14	-0,09	0,09	-0,04	0,06	0,00	0,07	0,06	0,10	-0,08	0,02	-0,03	0,02	-0,07	-0,13	-0,09	-0,02	-0,02	0,09	0,02	0,00	0,02	0,00	0,00
	φ	0,77	0,24	0,27	0,16	-0,11	-0,20	-0,04	-0,31	0,00	-0,03	0,00	-0,15	0,02	-0,03	-0,09	0,06	-0,10	0,07	0,00	0,11	-0,13	-0,13	-0,02	0,00	0,00	0,03	0,01	0,05	-0,01	0,03	0,01	0,00
	r_a	-0,70	-0,01	-0,26	0,08	0,29	0,09	-0,04	0,21	0,41	-0,02	-0,01	-0,07	-0,09	0,21	-0,01	0,10	0,21	0,06	-0,02	0,09	-0,07	-0,06	0,02	0,04	0,01	0,02	-0,02	0,00	0,01	0,00	0,00	0,01
	Explained Variance (%)	0,26	0,19	0,14	0,09	0,06	0,04	0,03	0,03	0,02	0,02	0,02	0,02	0,01	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

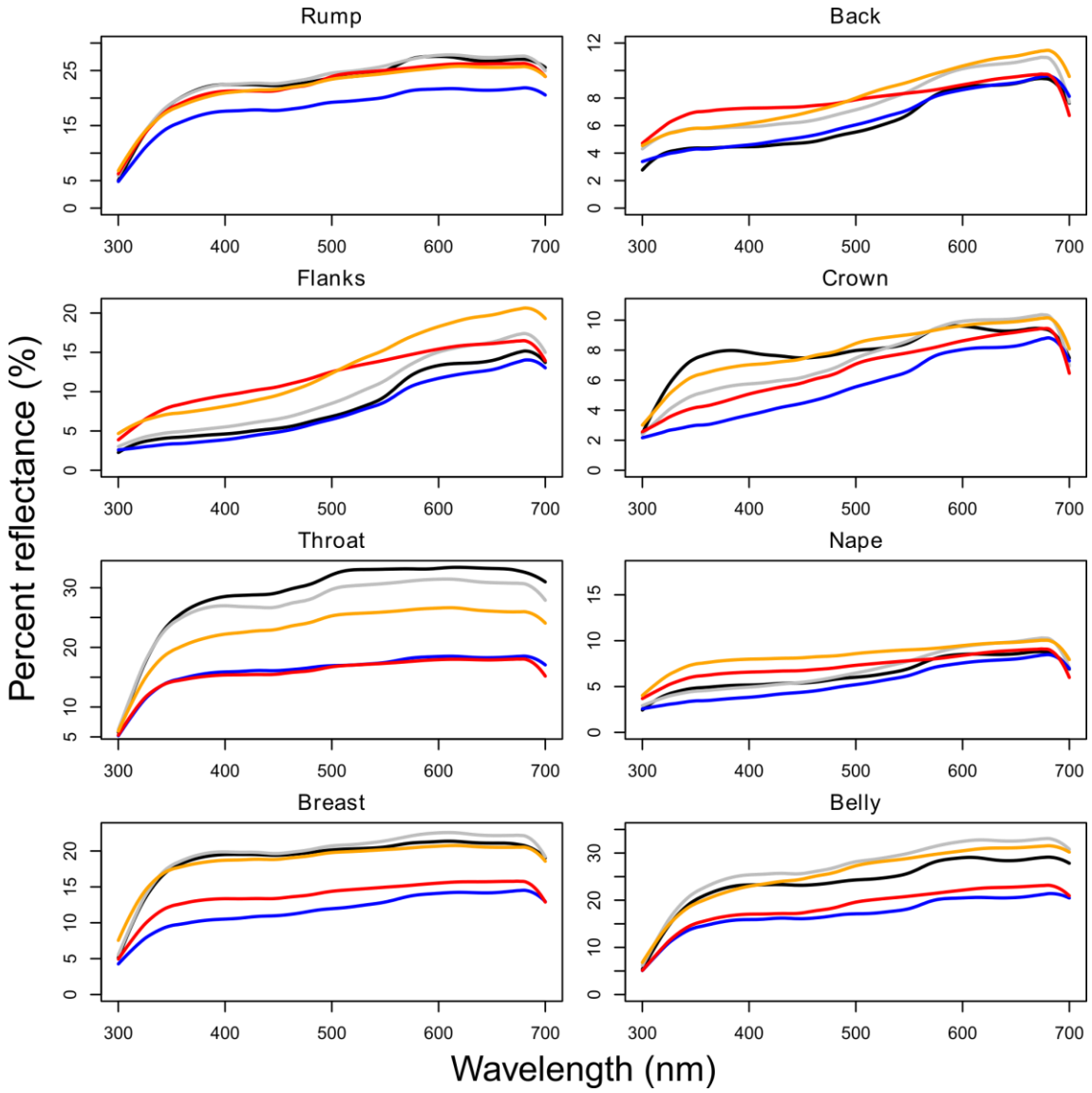


Figure S1: Mean spectra per plumage patch and per variant of *Zosterops borbonicus*.

Colours correspond to the different variants: blue, HBHB; red, G; grey, LBHB; black, BNB; orange, GHB.

Chapitre 2

Gradients altitudinaux et différenciation des populations

Présentation du chapitre

Les gradients altitudinaux sont généralement caractérisés par des changements forts de nombreuses caractéristiques abiotiques (*e.g.* température, pression atmosphérique, radiation UV, etc.) et biotiques (quantité de compétiteurs, prédateurs, etc.) de l'environnement. Ces zones sont donc souvent considérées comme des laboratoires naturels pour étudier l'effet des changements climatiques sur les populations. Le présent chapitre traite de l'étude des processus qui ont mené à la divergence morphologique des populations de *Z. borbonicus* sur les gradients altitudinaux de l'Ouest de l'île. L'article présenté dans ce chapitre vise à répondre à la question suivante : les patrons de variations génétiques et phénotypiques observés sont-ils le résultat de l'adaptation locale aux conditions environnementales, d'un contact secondaire entre des populations ayant auparavant divergé en allopatrie ou d'une action combinée des deux types de processus.

Contribution : Ce chapitre est constitué d'un article co-rédigé avec Joris Bertrand, nous nous partageons donc la position de premier auteur. Nous avons tous les deux participé à l'ensemble des étapes de la construction de cette article, depuis le terrain jusqu'à la rédaction. Joris a produit la majeure partie des données microsatellites.

The role of history and local adaptation on population differentiation along an elevational gradient in an island passerine bird

In preparation

Joris A. M. Bertrand[1*], Boris Delahaie[1*], Yann X. C. Bourgeois[1], Thomas Duval[2], Ricardo García-Jiménez[3], Josselin Cornuault[1], Benoit Pujol[1], Christophe Thébaud[1] and Borja Milá[3]

1. Laboratoire Évolution et Diversité Biologique, UMR 5174, Université Paul Sabatier, Toulouse 3 - Centre National de la Recherche Scientifique (CNRS) - École Nationale de Formation Agronomique (ENFA), 118 route de Narbonne, F-31062 Toulouse Cedex 9, France.

2. BP 438, 98822 Poindimié, Nouvelle-Calédonie

3. National Museum of Natural Sciences, Spanish Research Council (CSIC), José Gutiérrez Abascal, 2, 28006 Madrid, Spain.

* Contributed equally to this work and both should be therefore considered as first authors.

Correspondence: Joris Bertrand, Institute of Oceanography, room 111, National Taiwan University, No.1, Sec.4, Roosevelt Rd, Taipei 10617, Taiwan. Email : jorisbertrand@gmail.com

Running title: Elevational gradients and population differentiation

Abstract

Environmental gradients provide the opportunity to study how populations respond to spatial changes in ecological conditions at a scale at which gene flow should prevent differentiation. However, the historical dynamics of these systems can be complex and requires the comparison of both phenotypic and genetic variation to be rigorously unravelled. Here, we test for evidence that secondary contact and environmentally based selection underlie clinal phenotypic variation across populations of an island passerine bird (*Zosterops borbonicus*) that are distributed along an elevational gradient on the high island of Réunion (Mascarene archipelago). Using multilocus microsatellites screened in 401 individuals sampled in 18 populations distributed over the entire gradient, we found that genetic differentiation occurred at two spatial levels: (i) between two main groups arranged according to elevation, and (ii) within each of these two groups. The genetic break found at mid-elevation along the gradient as well as the occurrence of putative hybrids in a very narrow zone contrasts with the smoother variation in morphology and provide strong support for secondary contact between differentiated populations. Comparison of neutral genetic differentiation (F_{ST}) and phenotypic differentiation (P_{ST}) showed that P_{ST} largely exceeds F_{ST} at several morphological traits confirming that divergent selection might influence phenotypic variation among populations. Overall, our results suggest that local adaptation shapes phenotypic differentiation irrespective of population history, and may explain clinal patterns of variation in phenotype along elevational gradients even in cases when two genetically distinct populations come into geographic contact and form hybrid zones.

Keywords: Indian Ocean, Mascarene archipelago, *Zosterops borbonicus*, Réunion Grey White-eye, P_{ST}-F_{ST} comparisons, Microsatellites.

Introduction

Variation patterns in phenotypic and genetic diversity among populations along elevational gradients have been the subject of many recent studies examining the roles of natural selection and gene flow in driving population divergence at relatively small geographical scales (see Keller et al. 2013 and references therein). Populations of different elevational origin often show phenotypic differentiation for a variety of traits, including morphological (*e.g.* Price 1991; Bears et al. 2008, Milá *et al.* 2009, Pitchers et al. 2012) and life history traits (*e.g.* Angert and Schemske 2005; Luquet et al. 2015; review by Hille and Cooper 2014). In species with a wide elevational distribution, populations may exhibit changes in traits with elevation in response to varying climate-related selection pressures (see Körner et al., 2007), with differentiation taking place between populations in contiguous areas (*e.g.* McCormack and Smith 2008; Gonzalo-Turpin and Hazard 2009; Muir et al. 2013). Such effect is particularly likely on steep elevational gradients where environmental transitions occur at spatial scales that can be small relative to the species' dispersal distance. In such environments, the level of gene flow is not only dependent on dispersal capacity (*i.e.* expected gene flow) but also on habitat preference and the individual fitness of immigrants in the new environment (*i.e.* realized gene flow). In this case, active habitat choice combined with low performance of immigrants can promote local adaptation of populations even at small spatial scales (Kawecki and Ebert 2004; Morjan and Rieseberg 2004; Leinonen et al. 2008). However, phenotypic differences associated with the changing environment along elevation gradients can also develop in more complex evolutionary situations where previously isolated populations come into secondary contact and form hybrid zones, in spite of the relatively small geographical scale of most elevational gradients (Endler 1977; Barton and Hewitt 1985; 1989).

Secondary intergradation between populations living in two different environments often leads to sharp clines in adaptive traits (Endler 1977). However, this might not necessarily be the case when previously isolated populations meet at intermediate elevations along gradients in which climate- or habitat-related selection pressures change consistently with altitude (*e.g.* Cheviron and Brumfield 2009; Dubay and Witt 2014). In such situations, gradual changes in the optimal phenotype along the elevation gradient are expected for fitness-related traits (Endler 1986), with little or no sharp transition between the two populations where they meet. This can make contact zones hard or even impossible to detect without genetic data. This might be especially likely where there has been a long history of

contact between populations and ample opportunities for adaptation to local environmental conditions. Thus, in species with a wide elevational distribution, unless a clear expansion from either the lowlands or the highlands throughout the gradient has been documented, it seems important to account for population structure and history when discussing the causes for variation patterns in adaptive traits among populations living at different elevations.

The aim of this study was to examine the causes underlying variation in morphological traits among populations along a steep but regular elevational gradient in relation to the evolutionary context using the Réunion Grey White-eye (*Zosterops borbonicus;* taxonomy following Gill and Donsker 2014), a passerine bird endemic to the high island of Réunion (Mascarene archipelago, South-western Indian Ocean). On the western slopes of Réunion, *Z. borbonicus* is distributed from sea level to over 2500 m above sea level (asl) and shows a striking pattern of plumage color variation along the elevational gradient over very short geographic distances (15 km). Populations below 1350 m asl contain brown birds only and those above 1350 m asl contain a mixture of brown and grey birds (Gill 1973; see also Cornuault et al in press - Chapitre 1). Milá et al. (2010) have shown recently that lowland (brown) and highland (brown or grey) individuals differ in their morphology and belong to two distinct genetic clusters, an intriguing result consistent with a secondary contact between differentiated geographic forms but at odds with the apparent widespread distribution of brown birds along the gradient.

Here, we use genetic and morphological data from population samples collected along the entire elevational gradient to determine if the small-scale and contrasting patterns of phenotypic and genetic variation detected in previous studies could reflect adaptation to local conditions along the gradient, secondary contact, or combined action of both processes. We predicted that secondary contact and subsequent barriers to gene flow should result in a highland/lowland genetic structure with two main genetic clusters whereas an upward or downward expansion by a single ancestral population along a continuously varying environmental gradient should be accompanied by genetic differentiation at a very small spatial scale since these birds show an extremely reduced propensity to disperse (see Bertrand et al. 2014). We also examined whether the patterns of morphological variation were consistent with a predominant role of local adaptation by comparing the levels of neutral genetic differentiation (F_{ST}) to those of phenotypic differentiation (assessed by P_{ST}; Leinonen et al. 2006) for several morphological traits. According to this framework, if selection drives morphological variation among populations, P_{ST} should significantly exceed F_{ST}. The validity

of P_{ST}/F_{ST} comparisons relies however upon several assumptions that are often difficult to meet in nature (Leinonen et al. 2008; Pujol et al. 2008; Brommer, 2011; Brommer et al. 2014). To alleviate any concern, we used the approach, proposed by Brommer (2011) which consists in evaluating global P_{ST} at several traits and comparing these values to the F_{ST} of the most differentiated neutral microsatellite neutral locus.

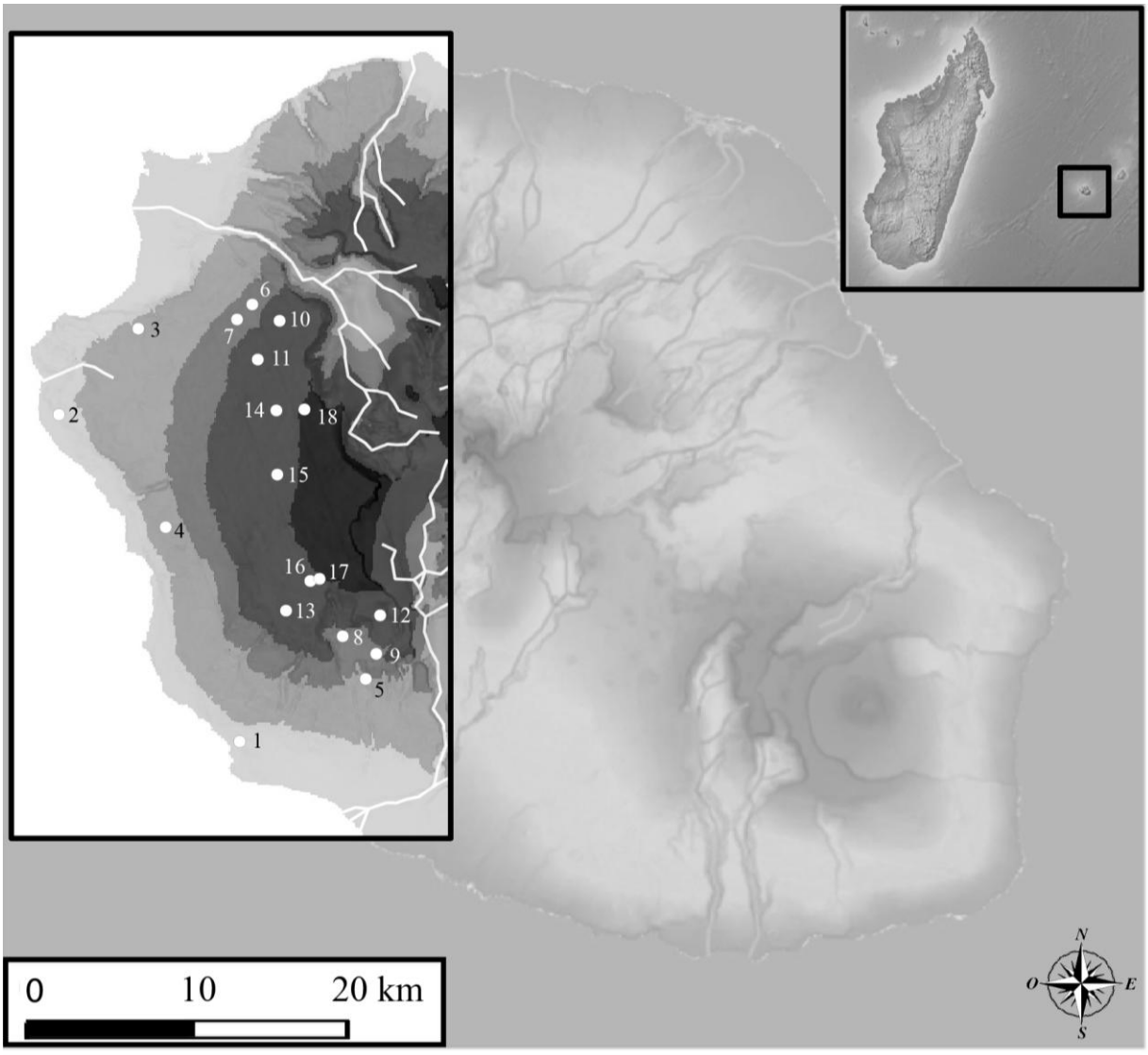


Figure 1: Map showing the location of Réunion Island as well as the geographical position of the 18 sampling localities on the island (number 1 to 18). Nuances of grey depict the transition between the five ecological zones (described by Strasberg *et al.*, 2005) being respectively from light to dark: lowland savannah (0-200 m), semi-dry sclerophyllous forest (200-750 m), lowland rainforest (from 750-1100 m) mountain rainforest (1100-2000 m) and subalpine scrubland (> 2000 m).

Materials & Methods

Population samples

We sampled a total of 401 individuals from 18 localities distributed along an elevational gradient located on the western slopes of Réunion and encompassing four distinct vegetation zones (from dry lowland forest to subalpine scrub) (Fig. 1 and Table 1). As ecological and geographic variations are expected to be highly correlated along environmental gradients (*i.e.* eco-spatial autocorrelation), we sampled at least two localities, as distant as possible from each other, in each ecological zone (as defined in Strasberg et al. 2005 and Thébaud *et al.* 2009). Pairwise geographic distances between localities varied from 0.5 km to 26.4 km with no obvious physical barriers to gene flow between them. Birds were caught using mist nets, ringed, and measured. They were released unharmed after a blood sample was collected by gently puncturing the subbrachial vein and stored in lysis buffer until freezing at 20°C. All manipulations were conducted under a ringing permit issued by the Centre de Recherches sur la Biologie des Populations d'Oiseaux–Museum d'Histoire Naturelle (Paris).

Table 1 Microsatellite diversity for the 18 localities across the 11 *loci*, sample size (*n*), average number of alleles per locus (*A*), Allelic richness (A_R) Observed and expected heterozygosity (H_O and H_E), and inbreeding coefficient (F_{IS}). Bold values indicate significant departures from Hardy-Weinberg Equilibrium. *H_E : Unbiased Expected heterozigosity computed in Genodive as calculated in Nei, 1987.

1	Étang Salé	-21.26	55.34	40	Dry Lowland Forest	32	9.73	5.91	0.77	0.78	0.01
2	Ermitage	-21.07	55.23	44	Dry Lowland Forest	16	7.18	5.75	0.77	0.81	0.04
3	Petit Bernica	-21.03	55.28	285	Semi-Dry Lowland Forest	10	6.46	5.62	0.83	0.76	**-0.09**
4	St-Leu	-21.14	55.30	526	Semi-Dry Lowland Forest	21	8.09	5.77	0.81	0.81	0.00
5	Canot	-21.22	55.41	687	Semi-Dry Lowland Forest	10	6.27	5.59	0,80	0.78	-0.03
6	Feoga 2	-21.01	55.35	925	Semi-Dry Lowland Forest	11	7.36	6.19	0.76	0,80	0.04
7	Feoga 1	-21.02	55.34	953	Lowland Rainforest	24	8.55	5.95	0.81	0.78	-0.04
8	Bon Acceuil	-21.20	55.40	994	Lowland Rainforest	8	6.09	5.86	0.79	0.76	-0.03
9	Makes 1050	-21.21	55.42	1047	Lowland Rainforest	10	7.00	5.98	0.75	0.79	0.04
10	Alcide 2	-21.02	55.36	1329	Cloud Forest	30	9.00	5.96	0.78	0.78	0.01
11	Alcide 1	-21.04	55.35	1331	Cloud Forest	8	6.36	6.13	0.78	0.81	0.03
12	Makes 1400	-21.19	55.42	1403	Cloud Forest	25	7.82	5.58	0.79	0.79	0.00
13	PK7	-21.19	55.37	1449	Cloud Forest	20	8.73	6.04	0.77	0.81	0.04
14	Tamarin 2	-21.07	55.36	1738	Cloud Forest	31	8.91	5.53	0.75	0.76	0.01
15	Tamarin 1	-21.11	55.36	1778	Cloud Forest	22	7.91	5.61	0.76	0.78	0.03
16	Sentier Tamarins	-21.17	55.38	1887	Cloud Forest	32	9.36	5.92	0.79	0.79	0.01
17	Tévelave	-21.17	55.39	1985	Subalpine Shrubland	27	8.09	5.65	0.73	0.79	**0.07**
18	Maïdo	-21.07	55.38	2062	Subalpine Shrubland	64	9.91	5.63	0.78	0.77	-0.02
	Mean					22	8.82	5.70	0.76	0.78	0.02

Morphological data

We measured six other morphological traits with either a ruler (wing length; to the nearest 0.5 mm) or dial calipers (all other traits; to the nearest 0.1 mm): wing length (chord of unflattened wing from the carpal joint to the tip of the longest primary, tail length (from the uropygial gland to the tip of the longest rectrix), tarsus length (from the intertarsal joint to the most distal undivided scute on the tarsometatarsus), bill length (from the anterior end of the nares to the tip of the upper mandible), bill width and depth (both measured at the anterior end of the nares). Tarsus length is a good proxy for overall body size (*e.g.* Senar and Pascual 1997), wing and tail length are traits connected to flight performance (*e.g.* Bears et al. 2008 and references therein) and body mass also reflects overall body size but is more dependent to condition and nutrient reserves (*e.g.* Balbontin et al. 2012). Bill characteristics are mainly related to foraging, song and heat regulation (*e.g.* Edelaar et al. 2012; Caro et al., 2013; Greenberg et al., 2012). To control for allometric covariance on wing length, tail length and bill characteristics, we extracted residuals from linear models using these traits as response variables and tarsus length as a fixed effect to account for overall body size. These residuals were used for all further univariate analyses.

Molecular markers

We extracted genomic DNA from blood samples using DNeasy Blood & Tissue kits (Qiagen, Venlo, Netherlands). All 401 individuals were genotyped at 12 polymorphic microsatellite loci previously isolated in the study species (Bertrand et al. 2012). PCR amplifications were performed in three 10 μL multiplexes (see Table S1), each containing ~ 5-30 ng of DNA, 0.2 mM dNTPs, 0.5 μM of each primer and 0.25 U *Taq* polymerase in 1X manufacturer's buffer (2 mM $MgCl_2$). PCR thermal profiles were as follows: initial denaturation at 94°C for 3 min, followed by 35 cycles at 94°C for 30 s, locus-specific annealing temperature (see Table S1) for 30 s, 72°C for 30 s and a final elongation step at 72°C for 10 min. Fluorescently labelled PCR products were mixed with formamide. Fragment analysis was carried out on an ABI PRISM 3730 DNA analyzer (Applied Biosystems, Foster City, California, USA) with GeneScan-500(LIZ)^{TM} size standard. Genotyping profiles were scored using GENEMAPPER v.4.0^{TM} (Applied Biosystems, Foster City, California, USA).

Data analyses

Environmental variation and morphological differentiation along the elevational gradient

To visualise environmental and morphological variation all along the gradient, we first performed a Principal Component Analysis (PCA) and plotted the scores derived from the first component (PC1) against elevation. For environmental data, we obtained from the French Meteorological Office (Météo-France, Toulouse) 19 bioclimatic variables related to temperature and rainfall (see Table S2 for details). In addition, we extracted the Normalized Difference Vegetation Index (NDVI) values (for the year 2009) from Moderate Resolution Imaging Spectroradiometer (MODIS) layers at each of our sampling locations as a proxy of standing biomass and vegetation cover (see details in Cornuault et al. 2013). A simple linear regression was then used to evaluate the degree of the correlation between PCA scores and elevation and its statistical significance. PCA were computed with the R-package {ade4} (Chessel et al. 2004). To examine phenotypic differentiation along the altitudinal gradient, we investigated the correlation between these variables and elevation and assessed its statistical significance with Kendall's rank correlation tests.

Population structure along the elevational gradient

We first performed a series of tests to validate the reliability of our microsatellite data set and assess within-population variation. The presence of null alleles was investigated using MICRO-CHECKER v.2.2.3 (van Oousterhout et al. 2004). Since the probability of null alleles across populations was low for all loci except *Z16*, we excluded this locus from all subsequent analyses. We used FSTAT v2.9.3.2 (Goudet 2001) to test for linkage disequilibrium between each pair of loci and to estimate the allelic richness corrected for sample size (A_R). We used GENODIVE v2.0 (Meirmans and Van Tienderen 2004) to characterize within-population genetic variation by calculating the mean number of alleles per locus (A) along with expected and observed heterozygosities (H_E and H_O). Inbreeding coefficients (F_{IS}) and deviation from Hardy-Weinberg equilibrium were also estimated with GENODIVE v2.0.

Patterns of neutral genetic differentiation and structure were investigated according to three complementary methodological approaches (F_{ST}, clustering analyses and AMOVA).

Here, the rationale was first to test for and quantify overall and among-population differentiation using F_{ST} and second to highlight potential barriers to gene flow between populations using a clustering method. To further validate the presence of barrier to gene flow with an independent method, we tested differentiation between the clusters obtained.

FSTAT v2.9.3.2 (Goudet 2001) was used to calculate global F_{ST} values over all loci and for each locus one by one (θ_{ST}: Weir and Cockerham 1984). The 95% confidence interval was obtained through bootstrapping (10000 bootstrap replicates) for overall F_{ST} values. Standard errors for F_{ST} per locus were obtained by jacknifing over populations. AMOVA-F_{ST}s were also performed using GENODIVE v2.0 (Meirmans and Van Tienderen 2004) to compute pairwise genetic differentiation between localities and significance levels were obtained using 10 000 permutations.

Neutral genetic structure was assessed using Bayesian clustering performed by the program STRUCTURE 2.3 (Pritchard et al. 2000; Falush et al. 2003). The number of clusters (K) was varied from 1 to 19. For each value of K, twenty MCMC runs were performed, each consisting of 500 000 generations after a burn-in period of 100 000 generations. The optimal value of K was evaluated by examining the change in mean likelihood values: L(K), as well as with the ΔK method (Evanno et al. 2005). These results have been computed and visualized in the online interface STRUCTURE HARVESTER (Earl and vonHoldt 2012). The optimal consensus of the 20 replicates was then determined with the *Greedy* algorithm implemented in CLUMPP (Jakobsson and Rosenberg 2007). The analysis was run using the admixture model with correlated allele frequencies that estimates, for each individual, the admixture proportion belonging to each cluster. We used locality as a prior for clusters (LOCPRIOR option), which allows detecting low levels of genetic structure without affecting the estimation of K (Hubisz et al. 2009).

Genetic structure was studied at different hierarchical levels by examining the partition of neutral genetic variation in an analysis of molecular variance (AMOVA) framework (Excoffier 1992; Michalakis and Excoffier 1996). This analysis consisted in investigating the proportion of genetic variance found within individual ($\approx F_{IT}$), among individuals within populations ($\approx F_{IS}$) and among populations ($\approx F_{ST}$). A further nested AMOVA provided tests for genetic differentiation among populations within groups ($\approx F_{SC}$) and between groups of populations ($\approx F_{CT}$) to assess highlands/lowlands differentiation. For the latter analysis, groups were defined based on assignation proportion based on clustering

analyses results (see below) and excluded apparent significantly admixed populations (whose assignation proportions range from 20 to 70 %).

Assessing the role of local adaptation in promoting morphological differentiation

To assess the role of local adaptation in this system, we compared neutral genetic differentiation (F_{ST}) with morphological differentiation. To do so, we quantified phenotypic differentiation for each morphological trait by P_{ST} (Leinonen et al. 2006). We estimated P_{ST} according to the equation proposed by Brommer (2011):

$$P_{ST} = \frac{\frac{C}{h^2}.\sigma_B^2}{\frac{C}{h^2}.\sigma_B^2 + 2.\sigma_B^2}$$

with ${\sigma_B}^2$ and ${\sigma_W}^2$ the phenotypic variance between and within populations respectively, h^2 is the heritability of the trait under study (i.e. the proportion of phenotypic variance due to additive genetic effects). The scalar c expresses the proportion of the between-population variance due to genetic effects across populations. Under controlled conditions (*e.g.* reciprocal transplants and common garden experiments), phenotypic differences are expected to be entirely due to additive genetic effects *i.e.* $c/h^2 = 1$ and P_{ST} to be equivalent to Q_{ST}, the analogue to F_{ST} for a given quantitative trait (Wright 1951; Spitze 1993). However, when individuals are measured in wild populations, non-additive effects, environmental factors or genotype-environment interactions may strongly affect the estimation of P_{ST} (Leinonen et al. 2008; Pujol et al. 2008; Brommer 2011; Brommer et al. 2014). The parameters h^2 and c may allow taking into account any non-additive genetic component in the estimation of P_{ST}. However, since h^2 and c are often cumbersome to assess in the wild too, we followed the approach described by Brommer (2011) to check the robustness of our comparisons by evaluating the sensitivity of the P_{ST} estimates to the variation of the proportion of environmental versus additive genetic effects across populations (*i.e.* c/h^2 ratio; in particular when $0 < c < h^2$) rather than just comparing P_{ST} to F_{ST} values. Although we did not assess the

heritability of these traits in our system, all of them have been shown to be heritable in several species (Charmantier et al. 2004; Teplitsky et al. 2009; Merilä et al. 2001).

Global P_{ST} values were calculated for each of the seven morphological traits. Using the R-package {MCMCglmm} (Hadfield 2010), we fitted Bayesian generalized linear mixed models for each trait in order to extract σ_B^2 and σ_W^2. Age and sex identity were included as fixed effects to correct any potential effect of these factors on the measurements. To avoid loosing power in our analyses, we assigned mean age for individuals of unknown age. Population was included as a random effect because it allows obtaining directly its variance components. The posterior modes of population-specific variances (σ_B^2) and residual variances (σ_W^2) were used to estimate the mode of P_{ST}. Confidence intervals were calculated by using the function HPDinterval of R-package {lme4} (Bates et al. 2014) which corresponds to the 95% highest posterior density of the distribution. Bayesian mixed models have been shown to give more precise and less biased estimates of P_{ST} than other methods (O'Hara and Merilä 2005; Leinonen et al. 2008). These models have the advantage of directly providing the variance of the P_{ST} as the complete distribution is inferred from the posteriors. In contrast, in maximum likelihood methods, the variance is estimated *a posteriori* through approximation or resampling methods. We used uninformative priors for the fixed and random effects (inverse-Wishart with nu = 0 and alpha mu = 1). The function uses Markov Chain Monte Carlo (MCMC). We ran three independent chains of 100 000 iterations of which the first 20 000 were discarded as burn-in. Convergence was assessed by computing the Brooks-Gelman-Rubin statistics for each parameter (Brooks and Gelman 1998). We checked for autocorrelation in the posteriors by calculating the effective size for each parameter.

P_{ST} values were then compared to F_{ST}. There are three possible outcomes for such comparisons: *(i)* P_{ST}-F_{ST} = 0 means that level of differentiation displayed by phenotypic traits equals neutral genetic differentiation. This pattern is consistent with neutral processes (genetic drift) as the main mechanism explaining divergence; *(ii)* P_{ST}-F_{ST} > 0 means that level of differentiation displayed by phenotypic traits is greater than expected in absence of any selective pressure (and environmental variation). It suggests that directional selection might be involved in promoting phenotypic divergence among populations; *(iii)* P_{ST}-F_{ST} < 0 means that level of differentiation displayed by phenotypic traits is smaller than expected under the assumption of neutrality. It suggests that other types of selective pressures such as stabilizing selection might restrict population phenotypic divergence.

For each comparison, the critical c/h^2 value was extracted as the value for which the lower 95% confidence interval of P_{ST} equalled the upper 95% confidence interval of F_{ST} (according to Kekkonen et al. 2012). Critical c/h^2 (noted c/h^{2*}) value denotes the robustness of the comparison between P_{ST} and F_{ST} (Brommer, 2011). It was estimated as follows:

$$c/h^{2^*} = \frac{\sigma^2_{W(upper)} . F_{ST}}{\sigma^2_{B(lower)} . (1 - F_{ST})}$$

with $\sigma_W{}^2_{(upper)}$ being the upper 95% phenotypic variance within population estimate and $\sigma_B{}^2_{(lower)}$ being the lower 95% phenotypic variance between populations estimate. As we used univariate phenotypic differentiation indexes (*i.e.* P_{ST} estimated for each trait separately), we compared them to the upper 95% F_{ST} estimate of the more differentiated locus in our study (*Z31*) (according to Whitlock 2008). This critical value gives an indication on the robustness of the P_{ST}-F_{ST} comparison. Although it is difficult to tell what would constitute a robust value of critical c/h^2, it seems that P_{ST} should be greater than F_{ST} over a major part of $c/h^2 < 1$ (Brommer 2011).

Results

We found that environmental variation was substantial along the gradient and varied gradually with elevation with no obvious step at mid-elevation or elsewhere. Overall body size, as estimated by PCA analysis, also appeared to be associated with elevation along the gradient (Fig. 2A), consistent with environmental variation. In contrast, neutral genetic structure exhibited a clear break at mid-elevation (see Fig. 3 and details in section below), consistent with secondary contact between two differentiated entities.

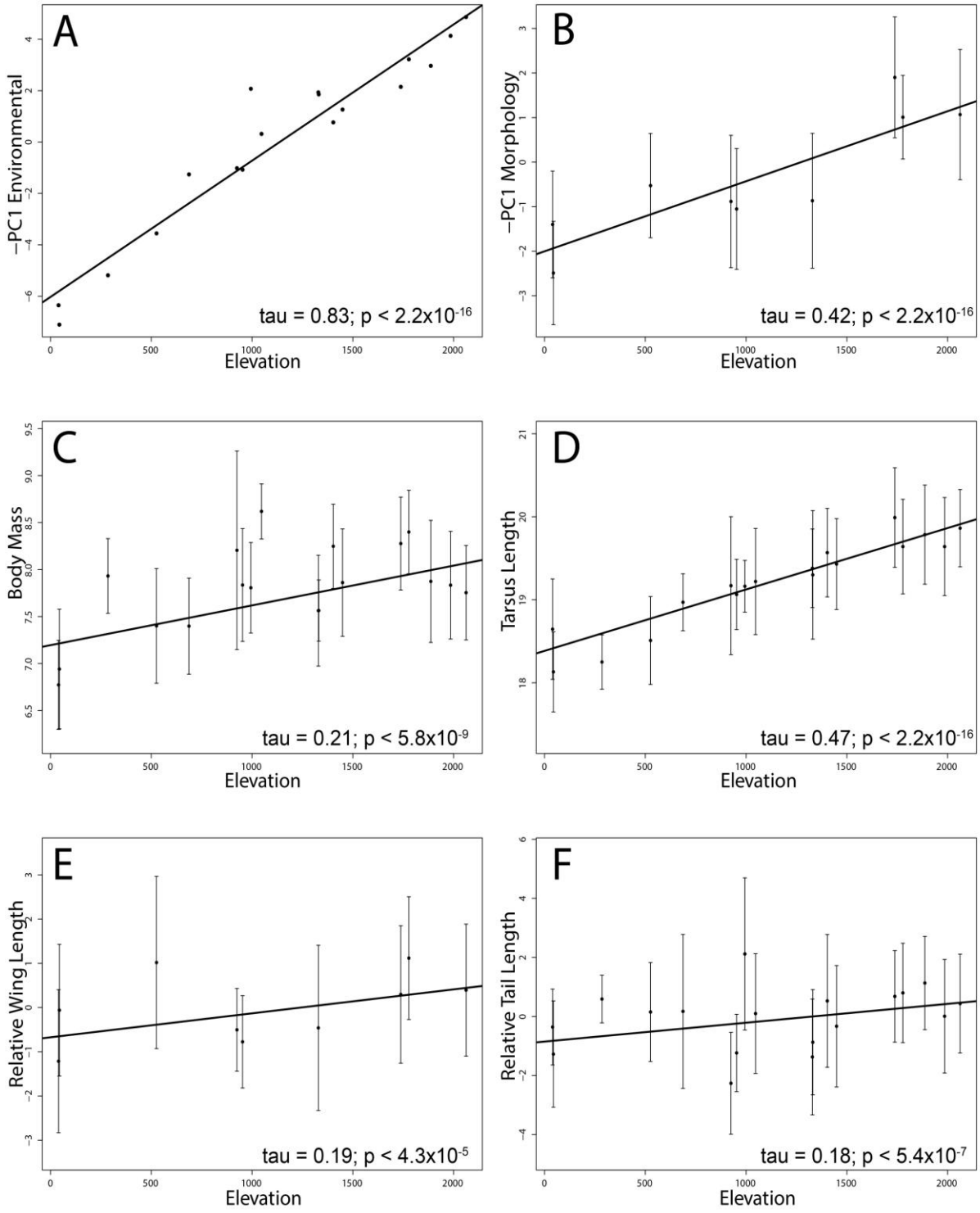


Figure 2: Environmental and morphological variation along the elevational gradient. **A)** PC1 scores from PCA on environmental and **B)** morphological variation (PC1 scores) against elevation**.** Elevational variation at different morphological traits: from to **C)** to **F)**.

Morphological variation along the elevational gradient

PC1 axis from the PCA done on the 19 environmental variables represented 62 % of the total variance and showed a strong positive correlation with elevation (tau = 0.83; $P = 2.^{7e-8}$, Fig 2A, Table S2 et S3). This confirmed that the pattern of environmental variation was gradual along the elevational gradient, with mean annual temperature decreasing by 13.6°C and rainfall increasing by over 1500 mm along the gradient from dry and hot lowlands to wet and cold highlands. PC1 axis was correlated with all morphometric variables and reflects variation in overall body size. Thus, populations showed a gradual increase in overall body size along the gradient, as depicted by the linear relationship between PC1 scores and elevation (tau = 0.42; p < 2.2e^{-}16, Fig. 2B). Trait-by-trait analyses further indicated that body mass tarsus length showed a consistent, gradual, increase with elevation (Fig 2C and 2D). Wing and tail lengths adjusted for body size (tarsus length) also showed a consistent, gradual, increase with elevation, indicating that along the gradient, for a given body size, individuals from increasingly higher elevations had proportionately longer wings and tails (Fig. 2E and 2F). Bill related traits (length, width and depth) did not show such a significant trend (tau = -0.03; 0.03 and -0.05 and *P*-values = 0.32; 0.44 and 0.13; Fig. S1)

Population genetic structure along the elevational gradient

None of the 11 microsatellite loci (out of the 12 genotyped: *Z16* was discarded) exhibited significant deviation from Hardy-Weinberg equilibrium. No significant linkage disequilibrium was found across all pairs of loci. All the localities sampled presented similar levels of within-population polymorphism (with A ranging from 6.27 to 9.91 and A_R ranging from 5.53 to 6.19, Table 1). Microsatellite loci presented differences in their relative level of polymorphism with number of alleles per locus ranging from 7 (for Z1 and Z2) to 39 (for Z15) (Table S1). This important allelic polymorphism was associated with high mean heterozygosities (H_O = 0.76 and H_E = 0.78). The mean inbreeding coefficient (F_{IS}) could not be statistically differentiated from zero in all but two localities (n°17: Tévelave, F_{IS} = 0.07 and n°3: Petit Bernica, F_{IS} = - 0.09) indicating no major significant deviation from panmixia (Table 1).

Overall F_{ST} value suggested a low but significant pattern of among-population neutral genetic differentiation (θ_{ST} = 0.034; 95% CI: 0.027 - 0.041). Almost all pairwise F_{ST}

comparisons between localities were significant and are consistent with a pattern of global neutral genetic differentiation (Table S4). Evanno's criterion (ΔK) derived from the STRUCTURE analysis suggested an optimal number of genetic clusters equal to two (L(K) = -17195 ΔK = 12.7) (Fig. 3 and S2). These two clusters are consistent with an elevational partition of the genetic structure. The first genetic cluster includes the sampling localities found at low elevation (< 1000 m) in which individuals exhibit posterior probabilities of assignment to the highland cluster < 20 % (with the exception of the locality n°4: St-Leu; at this site birds were sampled in a botanical garden, where vegetation and the ecological conditions were not strictly representative of the surrounding habitats (Milá, *pers. obs.*)). The second genetic cluster includes sampling stations located at high elevation (> 1500 m) in which individuals exhibit posterior probabilities of assignment to the highland cluster 70 %. At mid elevations (from 1000 to 1500 m asl) we observed a narrow transition zone with localities showing intermediate values of assignment probability to the highland cluster (between 20 % and 70 %) suggesting the existence of a cryptic hybrid zone around 1100-1400 m asl. A further examination of other values of *delta K* revealed secondary peaks potentially associated with levels of sub-structuring at K = 5 (L(K) = -16873, ΔK = 1.7) and predominately K = 9 (L(K) = -16741 ΔK = 3.0) (Fig. S2).

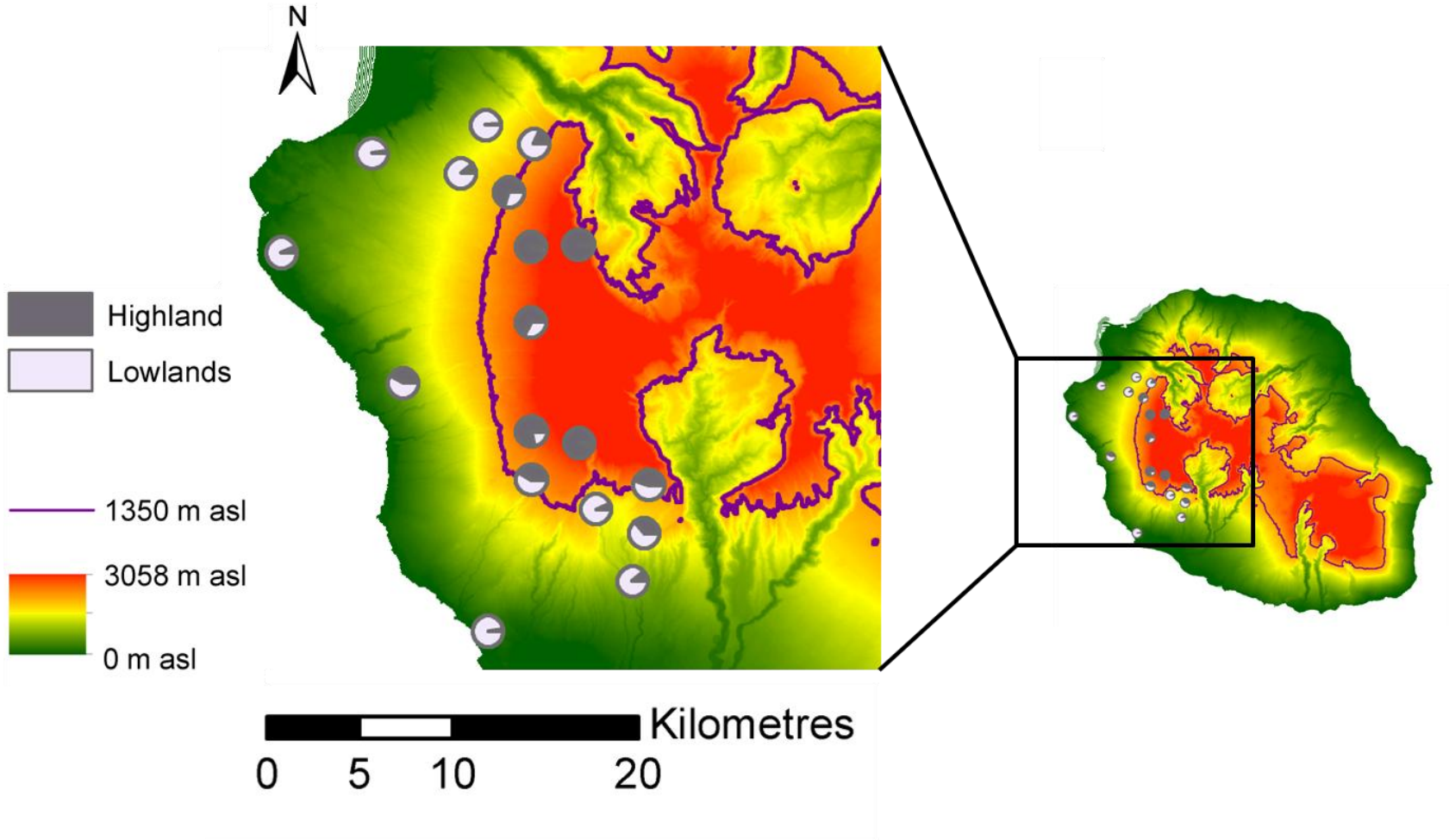


Figure 3: Spatial structure of microsatellite data according to the Bayesian clustering analysis performed with STRUCTURE. Pie diagrams represent the percent posterior probability of assignment (averaged across all individuals) to a given cluster at each locality. Dark grey corresponds to the highland cluster and light grey to the lowland one. The gradient between green and red corresponds to elevation variation from sea level (green) to the highlands (red).The 1350 m asl isocline is highlighted by the purple line.

As expected for highly polymorphic loci such as microsatellites, the within-individual variance component in the AMOVA (≈ F_{IT}) was high (> 96 %; Table 2). The among-individual level (≈ F_{IS}) confirms the absence of departure for Hardy-Weinberg equilibrium with a proportion of variance statistically negligible at this level (P ≥ 0.11). The among-population variance components were consistent with the Weir and Cockerham value mentioned above (F_{ST} = 0.031 and F_{SC} = 0.025 with and without grouping by cluster respectively) but statistically significant in the two combinations we tested (P = 0.00). The nested-AMOVA analyses also showed that the groups (here, the clusters) could explain a significant part of the total genetic variance (F_{CT} = 0.006, P = 0.02) thus confirming that highland and lowland birds were genetically different.

Table 2 AMOVA results: On the whole dataset, without any grouping and with a grouping corresponding to clusters found with STRUCTURE analysis (at K = 2) as well as on the highland cluster alone and on lowland cluster alone. The results show percentage of genetic variance at several hierarchical levels with associated *P*-values. Significant values are in bold.

No grouping	Nested in	F-stat	% Variance	P-value
Within Individuals		F_{IT}	0.964	*NA*
Among Individuals	Population	F_{IS}	0.005	0.219
Among Populations		F_{ST}	**0.031**	**0.000**
Among Groups				
Genetic cluster			% Variance	P-value
Within Individuals		F_{IT}	0.96	*NA*
Among Individuals	Population	F_{IS}	0.009	0.127
Among Populations	Genetic cluster	F_{SC}	**0.025**	**0.000**
Among Groups		F_{CT}	**0.006**	**0.021**
Highland cluster (alone)				
	Nested in	F-stat	% Variance	P-value
Within Individuals		F_{IT}	0.968	--
Among Individuals	Population	F_{IS}	0.013	0.112
Among Populations		F_{ST}	**0.020**	**0.000**
Lowland cluster (alone)				
	Nested in	F-stat	% Variance	P-value
Within Individuals		F_{IT}	0.966	--
Among Individuals	Population	F_{IS}	0.004	0.361
Among Populations		F_{ST}	**0.030**	**0.000**

Role of local adaptation in driving morphological differentiation along the elevational gradient

The most differentiated locus was *Z31* with a Weir & Cockerham F_{ST} value of 0.049 (+/- 0.014). Thus, we observed that phenotypic differentiation (P_{ST}) was significantly higher than neutral genetic differentiation (F_{ST}) for three of the seven traits considered: body mass, tarsus length and bill width adjusted for body size with P_{ST} = 0.24, 0.30, and 0.18 respectively (Table 3; Fig. 4.A, 4.B and 4.F). For these traits, the lower limit of the P_{ST} 95 % confidence interval was higher than the upper one of the more differentiated locus. The lowest critical c/h^2 value (0.38) was found for tarsus length (see Table 3, Fig. 4.B) denoting a greater robustness of the comparison between P_{ST} and F_{ST} for this trait. These results suggest that phenotypes are more divergent than expected under neutrality at these traits even if some across-population or non-heritable genetic variance is influencing differentiation, providing support to the idea that local adaptation drives population differentiation for these traits along the gradient.

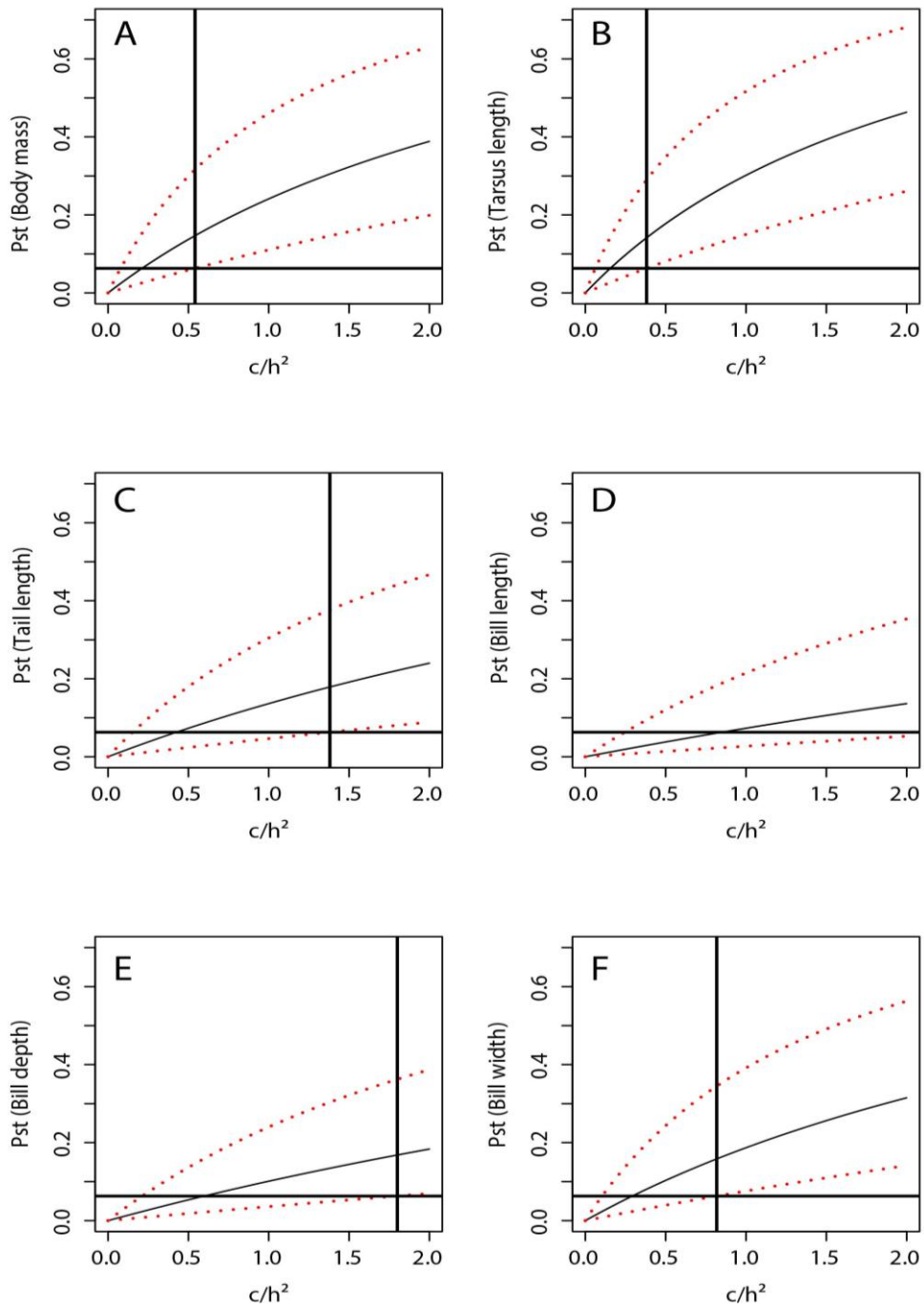


Figure 4 Evolution of P_{ST} for different values of c/h^2 at six morphological traits (wing length was discarded). Horizontal black lines represent the higher F_{ST} value observed at the most differentiated locus (*Z31*) and vertical black line represents the critical value of c/h^2 (noted c/h^{2}*). Dashed lines represent the 95 % confidence interval for P_{ST} values.

Table 3: P_{ST} values for each of the five morphometric traits (wing, tarsus and bill length, bill depth width), associated Confidence Interval (at 95 %) and critical c/h^2 values (c/h^2*). Bold values indicate traits for which the lower 95% CI P_{ST} value is greater than higher F_{ST} value at the most differentiated locus (*Z31*).

Trait	P_{ST}	Lower CI	Upper CI	Critical c/h^2 (c/h^2*)
Body Mass	0.24	0.11	0.46	0.54
Tarsus length	0.30	0.15	0.51	0.38
Relative Wing Length	-	-	-	-
Relative Tail Length	0.14	0.05	0.30	1.38
Relative Bill Length	0.07	0.03	0.21	2.41
Relative Bill Depth	0.10	0.04	0.24	1.80
Relative Bill Width	0.18	0.08	0.39	0.82

Discussion

Environmental gradients can induce gradual spatial differentiation in phenotypes among populations of common origin due to local adaptation along the gradient, resulting in correlations between spatial location and phenotype (Kirkpatrick and Barton, 1997; Keller et al. 2013). However, such pattern of gradual change can also appear in the case of more complex evolutionary scenarios such as secondary contact between previously isolated and differentiated populations because local adaptation along the gradient also induces gradual spatial differentiation in phenotypes within each set of populations in such circumstances and may drive such populations towards similar phenotypic optima where they meet. Thus, investigating the underlying neutral genetic structure to understand population history appears to be an important prerequisite to the study of the evolutionary dynamics of populations distributed over steep environmental gradients, such as elevational gradients (see *e.g.* Fuchs et al. 2011; Caro et al. 2013; Cheviron and Brumfield 2009). In this study, our analyses reveal that birds from the western slope of the island belong actually to two distinct genetic units that come into contact at a narrow hybrid zone mid-way at about 1400 m asl along a steep elevational and environmental gradient. It highlights that brown birds are in fact composed of two distinct genetic entities with highland birds (brown and grey ones) being in a different genetic cluster than lowland brown ones. Our results also show that size-related traits like body mass, wing, tail and tarsus length increase gradually with elevation across *Z. borbonicus* populations. Taken together, these results suggest that these phenotypic changes are likely to

be the result of a complex evolutionary scenario of secondary contact between two genetic entities combined with the action of selection along the gradient.

Elevational gradients in the tropics present particularly marked thermal stratification compared to temperate mountains (Janzen 1967). This is apparently well explained by the low seasonal variation in ambient temperature (Janzen 1967). As a consequence, it has been highlighted that mountain species that experience less variation in annual temperature in the tropics present narrower thermal tolerances than temperate species (McCain 2009; Cadena *et al.* 2011). These specificities are thought to give more opportunities for allopatric differentiation (Cadena et al. 2011). Thus, the contact zone between lowland and highland form of *Z. borbonicus* seems to match with the hypothesis that most of the diversification events along altitudinal gradients likely involved an allopatric phase (Cadena 2007; Cadena et al. 2011; Fjeldsa, Bowie and Rhabek, 2012). In-situ divergence of the two genetic units of *Z. borbonicus* in parapatry (or primary intergradation) could in principle produce a similar pattern of genetic variation along the gradient (Endler 1977; Barton and Hewitt 1985), but we find it to be a less parsimonious alternative to secondary contact between forms differentiated in allopatry. First, past allopatry is highly likely on Réunion Island due to the active eruptive history of the Reunion's volcanoes which may have restricted populations in isolated pockets of suitable habitat (discussed in Milá et al. 2010). Second, as shown by the PCA on environmental variables, the ecological change seems to be very gradual along the gradient and not associated with sharp changes in conditions, and lacks any major discontinuity that could support the parapatric divergence hypothesis. Even if primary intergradation is theoretically feasible on environmental gradients (Doebeli and Dieckmann 2003; Mizera and Meszéna 2003; Goldberg and Lande, 2006), empirical data tend to show that physical discontinuities or abrupt changes in the environment are necessary to impose the strong selection needed to give rise to the observed pattern (*e.g.* Schneider *et al.* 1999; Niemiller *et al.* 2008 but see Ohlberger et al. 2013). The position of the Reunion white-eye contact zone appears to correspond to an ecotone between native habitat (> 1400 m a.s.l.) and anthropogenic landscapes (< 1400 m a.s.l.), but the recent historical origin of this environmental break (less than 300 years) suggests that it may play a role in the position of the contact zone, but not its origin. Third, Milá *et al.* (2010) showed that more than 30% of their AFLP markers presented high F_{ST} values (> 0.1). Based on this distribution, they suggested that the differentiation between highland and lowland birds was the result of allopatric divergence. We also note that the marked pattern of genetic differentiation observed

in neutral microsatellite loci seems more consistent with differentiation in allopatry than primary intergradation, in the latter case we would have expected neutral loci to be less differentiated (but see Bierne, Gagnaire and David 2013).

In contrast to results from neutral genetic markers, morphological traits reveal the role of selection in driving local adaptation along the gradient. Traits related to body size (body mass, tarsus, tail and wing length) increased significantly with elevation and consistently with environmental variation. P_{ST}/F_{ST} comparisons at tarsus length and body mass showed that the degree of phenotypic differentiation (P_{ST}) largely exceeds the degree of neutral genetic differentiation (F_{ST}) even for small value of the c/h^2 ratio. Despite the short distances involved, the ~15°C differences in annual mean temperature between the two extreme localities along the gradient could explain why birds from the highlands are larger. Upland birds have often been found to be larger and heavier than lowland populations (e.g. Blackburn and Ruggiero 2001; Laiolo and Rolando 2001 Soobramoney et al. 2005; Milá et al. 2009). Hypotheses proposed to explain this pattern are either related to thermoregulation, feeding behaviour or to resistance to resource limitation (see Cornuault *et al. in press* - Chapitre 1 for more precisions about the causal agents of selection). Our results thus show that selection along the gradient seems strong enough to have driven convergence for an optimal size at mid elevations in genetic populations of different origins. Thus, ecological selection seems to shape the variation of these morphological traits regardless of the genetic background. This leads the contact zone to be cryptic to the human eye based on morphological criteria. It is thus possible that this kind of situation may be much more frequent than previously thought in nature but it has not yet been noticed.

Our study documents a morphologically cryptic contact zone among two genetic populations along an elevational gradient of less than 15 km even in a species with high dispersal potential. This situation is likely due to secondary contact after allopatric divergence. As the two genetic units colonized the elevational gradient, ecology has played an important role in driving phenotypic variation in fitness-related traits, driven by local adaptation and facilitated by low dispersal. Using model-based strategy with new markers would be very useful to estimate the time since secondary contact and to definitely rule out primary intergradation. The presence of a narrow hybrid zone likely indicates the existence of a barrier to gene flow between these two genetic units despite the high dispersal capacity of white-eyes. Further studies on this hybrid zone are required to gain a better understanding of the reproductive isolation mechanisms which are at play. This system might anyway provide

a good opportunity to investigate small spatial scale phenotypic and genetic differentiation and the evolutionary consequences of selection-constrained gene flow in highly mobile taxa (*e.g.* Garroway, et al. 2013; García-Navas et al. 2014) and in particular in spatially controlled island dwelling systems (*e.g.* De Léon et al. 2010; Arnoux et al. 2014).

Acknowledgements

Fieldwork was facilitated by the outstanding efforts of Guillaume Gélinaud, Dominique Strasberg, Ben Warren, Juli Broggi, Magali Thierry, René-Claude Billot, Jean-Michel Probst, Isabelle Henry and Vincent Leconte. We gratefully thank the Réunion National Park for permission to conduct fieldwork. Marc Salamolard and Benoît Lequette provided valuable help with fieldwork and logistics. J.B, Y.B, B.D and J.C were supported by MESR (Ministère de l'Enseignement Supérieur et de la Recherche) PhD scholarships. The research was supported by Agence Française pour le Développement grants to C.T, the Fondation pour la Recherche sur la Biodiversité (FRB) through its Centre for Synthesis and Analysis of Biodiversity (CESAB), the 'Laboratoire d'Excellence' TULIP (ANR-10-LABX-41), and the SYNTHESYS Project (http://www.synthesys.info/) which is financed by European Community Research Infrastructure Action under the FP7 "Capacities" Programme at the Museo Nacional de Ciencias Naturales (CSIC) of Madrid, Spain.

References

Åkesson, M., Bensch, S., Hasselquist, D. 2007. Genetic and phenotypic associations in morphological traits: a long term study of great reed warblers *Acrocephalus arundinaceus*. J Avian Biol 38: 58–72.

Angert, A. L. and Schemske, D. W. 2005. The evolution of species' distributions: reciprocal transplants across the elevation ranges of *Mimulus* cardinalis and *M. lewisii*. Evolution 59:1671-1684.

Arnoux, E., Eraud, C., Navarro, N., Tougard, C., Thomas, A., Cavallo, F., Vetter, N. Faivre, B., Garnier, S. 2014. Morphology and genetics reveal an intriguing pattern of differentiation at a

very small geographic scale in a bird species, the forest thrush *Turdus lherminieri*. Heredity 113: 514-525.

Badyaev, A.V., Young, R.L., Oh, K.P. and Addison, C. 2008. Evolution on a local scale: developmental, functional, and genetic bases of divergence in bill form and associated changes in song structure between adjacent habitats. Evolution 62:1951-1964.

Balbontín, J., Møller, A. P., Hermosell, I. G., Marzal, A., Reviriego, M. and De Lope, F. 2012. Lifetime individual plasticity in body condition of a migratory bird. Biol. J. Linn. Soc. 105: 420-434.

Barton, N.H. and Hewitt, G.M. 1985. Analysis of hybrid zones. Annu. Rev. Ecol. Evol. Syst. 16:113-148.

Barton, N.H. and Hewitt, G.M. 1989. Adaptation, speciation and hybrid zones. Nature 341: 497-503.

Bates D, Maechler M, Bolker B and Walker S (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, http://CRAN.R-project.org/package=lme4

Bears, H., C. Drever, M. and Martin, K. 2008. Comparative morphology of dark-eyed juncos *Junco hyemalis* breeding at two elevations: a common aviary experiment. J. Avian Biol. 39: 152-162.

Bertrand, J.A.M., García-Jiménez, R., Bourgeois, Y., Duval, T., Heeb P., Thébaud, C. and Milá, B. 2012. Isolation and characterization of twelve polymorphic microsatellite loci for investigating an extreme case of microgeographical variation in an island bird (*Zosterops borbonicus*). *Conserv. Genet. Resour*. 4:323-326.

Bertrand J. 2013. Causes de la différenciation génétique à une très petite échelle spatiale chez un oiseau insulaire (*Zosterops borbonicus*). PhD thesis. University of Toulouse, Toulouse.

Bertrand, J.A.M., Bourgeois, Y.X.C., Delahaie, B., Duval, T., García-Jiménez, R., Cornuault, J., Heeb, P., Milá, B., Pujol, B. & Thébaud, C. 2014. Extremely reduced dispersal and gene flow in an island bird. Heredity 112:190-196.

Bierne, N. Gagnaire, P.-A. and David, P. (2013) The geography of introgression in a patchy environment and the thorn in the side of ecological speciation. Curr. zool. 59-72-86.

Blackburn, T. M. and Ruggiero, A. 2001. Latitude, elevation and body mass variation in Andean passerine birds. Global Ecol. and Biogeogr. 10:245–259.

Brommer, J.E. 2011. Whither P_{ST}? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. J. Evol. Biol. 24:1160-1168.

Brommer, J. E., Hanski, I. K., Kekkonen, J. and Väisänen, R. A. 2014. Size differentiation in Finnish house sparrows follows Bergmann's rule with evidence of local adaptation. J. Evol. Biol., 27:737-747.

Brooks, S. and Gelman, A. 1998. General methods for monitoring convergence of iteratives simulations. J. Comp. Graph. Stat.7: 434-455

Cadena, C.D. 2007. Testing the role of interspecific competition in the evolutionary origin of elevational zonation: an example with Buarremon Brush-Finches (*Aves, Emberizidae*) in the neotropical mountains. Evolution 61: 1120-1136.

Cadena, C.D., Kozak, K.H., Gómez, J.P., Parra, J.L., McCain, C.M., Bowie, R.C.K. et al. 2011. Latitude, elevational climatic zonation and speciation in New World vertebrates. Proc. R. Soc. Lond. B 279: 194–201.

Caro, L.M., Caycedo-Rosales, P.C., Bowie, R.C.K., Slabbekoorn, H. and Cadena, C.D. 2013. Ecological speciation along an elevational gradient in a tropical passerine bird? J. Evol. Biol. 26:357–374.

Charmantier, A., Kruuk, L.E.B., Blondel, J., Lambrechts, M.M. 2004. Testing for microevolution in body size in three blue tit populations. J Evol Biol 17: 732-743.

Chessel, D., Dufour A.-B, and Thioulouse J. 2004. The ade4 package-I- One-table methods. R News 4:5-10.

Cheviron, Z.A. and Brumfield, R.T. 2009. Migration-selection balance and local adaptation of mitochondrial haplotypes in rufous-collared sparrows (*Zonotrichia capensis*) along an elevational gradient. Evolution 63: 1593-1605.

Cornuault, J., Khimoun, A., Harrigan, R. J., Bourgeois, Y. X. C., Milá, B., Thébaud, C. and Heeb, P. 2013. The role of ecology in the geographical separation of blood parasites infecting an insular bird. J. Biogeogr. 40:1313-1323.

Cornuault, J., Delahaie, B. Bertrand, J.A.M., Bourgeois, Y.X.C, Milá, B., Heeb, P. and Thébaud C. 2014. Morphological and plumage colour variation in the Réunion Grey White-eye (Aves: *Zosterops borbonicus*): assessing the role of selection. Biol. J. Linnean Soc. In Press.

De Léon, F., Bermingham, E., Podos, J and 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.

Doebeli M and Dieckmann U.2003. Speciation along environmental gradients. Nature 421:259-264.

DuBay, S. G. and Witt, C. C. 2014. Differential high-altitude adaptation and restricted gene flow across a mid-elevation hybrid zone in Andean tit-tyrant flycatchers. Mol. Ecol. 23:3551-3565.

Edelaar, P., Alonso, D., Lagerveld, S., Senar, J. C. and Björklund, M. 2012. Population differentiation and restricted gene flow in Spanish crossbills: not isolation-by-distance but isolation-by-ecology. J. Evol. Biol. 25:417–430.

Earl, D.A. and vonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Cons. Genet. Resour. 4: 359-361.

Endler, J.A. 1977. Geographic Variation, Speciation, and Clines. Princeton University Press, Princeton, New Jersey.

Endler, J.A. 1986. Natural selection in the wild. Princeton University Press, Princeton, New Jersey.

Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620.

Excoffier, L., Smouse, P.E. and Quattro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. Genetics 131: 479-491.

Falush, D., Stephens, M. and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Fjeldså, J., Bowie, R.C.K and Rahbek, C. 2012. The role of mountain ranges in the diversification of birds. Annu. Rev. Ecol. Evol. Syst. 43:249-265.

Frantz, A. C., Cellina, S., Krier, A., Schley, L. and Burke, T. 2009. Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? J. Appl. Ecol., 46: 493–505.

Fuchs, J., Fjeldsä, J. and Bowie, R. 2011. Diversification across an altitudinal gradient in the Tiny Greenbul (*Phyllastrephus debilis*) from the Eastern Arc Mountains of Africa. BMC Evol. Biol. 11: 1-17.

García-Navas, V., Ferrer, E. S., Sanz, J. J. and Ortego, J. 2014, The role of immigration and local adaptation on fine-scale genotypic and phenotypic population divergence in a less mobile passerine. J. Evol. Biol. 27:1590-1603.

Garnier, S., Alibert, P., Audiot, P., Prieur, B. and Rasplus, J.-Y. 2004, Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, *Carabus solieri*. Mol. Ecol. 13: 1883–1897.

Garroway, C. J., Radersma, R., Sepil, I., Santure, A. W., De Cauwer, I., Slate, J. and Sheldon, B. C. 2013. Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. Evolution 67: 3488-3500.

Gill, F.B. 1973. Intra-island variation in the Mascarene White-eye *Zosterops borbonica*. *Ornithological Monographs* 12.

Goldberg, E.E. and Lande, R. 2006. Ecological and reproductive character displacement on an environmental gradient. Evolution 60: 1344–1357.

Gonzalo-Turpin, H. and Hazard, L. 2009. Local adaptation occurs along altitudinal gradient despite the existence of gene flow in the alpine plant species *Festuca eskia*. J. Ecol., 97: 742–751.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices. Version 2: 3. J. Hered. 86: 485-486.

Grant, P.R. and Grant, B.R. 2007. How and Why Species Multiply: the Radiation of Darwin's Finches. Princeton University Press, Princeton, NJ.

Greenberg, R. Danner, R., Olsen, B., and Luther, D. 2012. High summer temperature explains bill size variation in salt marsh sparrows. Ecography 35: 146-152.

Hadfield, J.D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. J. Stat. Softw. 33: 1-22.

Hille, S. M. and Cooper, C. B. 2014. Elevational trends in life histories: revising the pace-of-life framework. Biol. Rev.: In press.

Hubisz, M.J., Falush, D., Stephens, M. and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9: 1322-1332.

Jakobsson, M. and Rosenberg, N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801-1806.

James, F.C. 1983. Environmental component of morphological differentiation in birds. Science 221:184-186.

Janzen, D.H. 1967. Why mountain passes are higher in the tropics. Am. Nat. 101: 233-249.

Kawecki, T.J. and Ebert, D. 2004. Conceptual issues in local adaptation. Ecol. Lett. 7: 1225-1241.

Kekkonen, J., Jensen, H. and Brommer, J.E. 2012. Morphometric differentiation across House Sparrow *Passer domesticus* populations in Finland in comparison with the neutral expectation for divergence. Ibis **154**: 846-857.

Keller, I., Alexander, J.M., Holderegger, R. and Edwards, P.J. 2013. Widespread phenotypic and genetic divergence along altitudinal gradients in animals. J. Evol. Biol.12:2527-2543.

Kirkpatrick, M. and Barton, N.H. 1997 Evolution of a species' range. American Naturalist 150: 1-23.

Körner, C. 2007. The use of 'altitude' in ecological research. *Trends Ecol. Evol.* 22: 569-574.

Laiolo, P. and Rolando, A. 2001. The evolution of vocalisations in the genus *Corvus*: effects of phylogeny, morphology and habitat. Evol. Ecol. 17 :111-123.

Landmann, A. and Winding, N. 1993. Niche segregation in high-altitude Himalayan chats (Aves, Turdidae): does morphology match ecology? Oecologia 95:506-519.

Landmann, A. and Winding, N. 1995. Guild organisation and morphology of high-altitude granivorous and insectivorous birds: convergent evolution in an extreme environment. Oikos 73:237-250.

Leinonen, T., Cano, J.M., Makinen, H. and Merila, J. 2006. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. J. Evol. Biol. 19:1803-1812.

Leinonen, T., O'Hara, R.B., Cano, J.M. and Merilä, J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. J. of Evol. Biol. 21:1-17.

Luquet, E., Léna, J.-P.; Miaud, C. and Plénet, S. 2015 Phenotypic divergence of the common toad (*Bufo bufo*) along an altitudinal gradient: evidence for local adaptation. Heredity 114: 69-79.

McCain, C.M. 2009. Global analysis of bird elevational diversity. Global Ecology and Biogeography 18: 346-360.

McCormack, J.E. and Smith, T.B. 2008. Niche expansion leads to small-scale adaptive divergence along an elevation gradient in a medium-sized passerine bird. Proc. R. Soc. B 275:2155-2164.

Meirmans, P.G. and Van Tienderen, P.H. 2004. Genotype and Genodive: Two programs for the analysis of genetic diversity of asexual organisms. Mol. Ecol. Notes 4:792-794.

Meirmans, P. G. 2012. The trouble with isolation by distance. Mol. Ecol., 21: 2839–2846.

Merila¨ J, Kruuk LEB, Sheldon BC (2001) Natural selection on the genetical component of variance in body condition in a wild bird population. J Evol Biol 14: 918–929.

Michalakis, Y. and Excoffier, L. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142: 1061-1064.

Milá, B., Wayne, R.K., Fitze, P. and Smith, T.B. 2009. Divergence with gene flow and fine-scale phylogeographical structure in the wedge-billed woodcreeper, *Glyphorynchus spirurus*, a Neotropical rainforest bird. Mol. Ecol. 18: 2979–2995.

Milá, B., Warren, B.H., Heeb, P. and Thébaud, C. 2010. The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene Grey White-eye (Aves: *Zosterops borbonicus*). BMC Evol. Biol.10:158.

Mizera, F. and Meszéna, G. 2003. Spatial niche packing, character displacement and adaptive speciation along an environmental gradient. Evol. Ecol. Res. 5:1-20.

Morjan, C. L. and Rieseberg, L. H. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol. Ecol. 13: 1341-1356.

Muir, A.P., Biek, R., Thomas, R. and Mable, B.K. 2014. Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria*). Mol. Ecol. 23:561-574.

Niemiller, M. L., Fitzpatrick, B. M. and Miller, B. T. 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. Mol. Ecol. 17: 2258-2275.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. and Shipley, P. 2004. Micro-Checker: Software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4:535-538.

O'Hara, R.B and Merilä, J. 2005. Bias and precision in Q_{ST} estimates: problems and some solutions. *Genetics* 171:1331-1339.

Ohlberger, J., Brännström, A. and Dieckmann, U. 2013. Adaptative phenotypic diversification along a temperature-depth gradient. Am. Nat. 182: 359-373.

Piersma, T. and Davidson, N.C. (1991) Confusion of mass and size. Auk 108: 441-444.

Pitchers, W. Pool, J.E. and Dworkin, I. 2013. Altitudinal clinal variation in wing size and shape in African *Drosophila melanogaster*: one cline or many? Evolution 67: 438-452.

Price T. 1991. Morphology and ecology of breeding warblers along an altitudinal gradient in Kashmir, India. J. Anim. Ecol. 60:643-664.

Price T. 2008. Speciation in birds. Roberts & company Publishers. Greenwood Village, CO.

Pritchard, J.K., Stephens, M. and Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

Porlier, M., Garant, D., Perret P. and Charmantier, A. 2012. Habitat-linked population genetic differentiation in the Blue Tit *Cyanistes caeruleus*. J. Hered. 103: 781-791

Pujol, B., Wilson, A.J., Ross, R.I.C. and Pannel, J.R. 2008. Are Q_{ST}–F_{ST} comparisons for natural populations meaningful? Mol. Ecol. 17:4782-4785.

Senar, J.C. and Pascual, J. 1997. Keel and tarsus length may provide a good predictor of avian body size. Ardea 85:269–274.

Schneider, C.J., Smith, T.B., Larison B., Moritz C. 1999. A test of alternative models of diversification in tropical rainforests: ecological gradients vs. rainforest refugia. Proc. Natl. Acad. Sci U. S. A. 96: 13869-13873.

Soobramoney, S., Downs, C.T. and Adams, N.J. 2005. Morphological variation in the Common Fiscal *Lanius collaris* along an altitudinal gradient in southern Africa. *Ostrich* 76:130-141.

Spitze, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozyme variation. Genetics 135: 367-374.

Strasberg, D., Rouget, M., Richardson, D.M., Baret, S., Dupont, J. and Cowling R.M. 2005. An assessment of habitat diversity and transformation on La Réunion Island (Mascarene Islands,

Indian Ocean) as a basis for identifying broad-scale conservation priorities. Biodiversity Conserv. 14:3015-3032.

Symonds, M.R.E and Tattersall, G.J. 2010. Geographical variation in bill size accross bird species provides evidence for Allen's rule. Am. Nat. 176: 188-197.

Teplitsky, C., Mills, J.A., Yarrall, J.W., Merilä, J.2009. Heritability of fitness components in a wild bird population Evolution. 63: 716–726.

Thébaud, C., Strasberg, D., Warren B.H. and Cheke, A. 2009. Mascarene islands, biology. In: Gillespie R.G. & Clague D.A. (eds) Encyclopedia of Islands. University of California Press, Berkeley, CA. pp 612-619.

Weir, B.S. and Cockerham, C.C. 1984. Estimating F-Statistics for the analysis of population structure. Evolution 38 :1358-1370.

Whitlock, M. C. 2008. Evolutionary inference from Q_{ST}. Mol. Ecol. 17:1885-1896.

Wright, S. 1951. The genetical structure of populations. A. Eug. 15: 323-354.

Supplementary information

Table S1: Microsatellite DNA specifications and diversity for the 12 loci accross 18 populations. PCR conditions and range of alleles sizes, sample size (n), average number of alleles per locus (A), Observed and expected heterozygosity (HO and HE), and inbreeding coefficient (FIS). Differentiation indices such as GST (Nei, 1987), G"ST (Meirmans & Hedrick, 2011) and Dest (Jost, 2008).

Locus	Pre-PCR Multiplex	Post-PCR Multiple	Fluoro-Label	T_a (°C)	Product size (bp)	*A*	A_R	H_O	H_E	F_{IS}
Z1	1	1	PET	56	148-172	7	4.137	0.64	0.64	-0.04
Z2	2	2	6-FAM	56	221-245	7	4.40	0.68	0.70	-0.02
Z3	No	3	VIC	56	170-220	16	6.972	0.87	0.85	-0.07
Z4	No	3	PET	56	146-228	29	8.081	0.84	0.84	-0.04
Z5	No	1	NED	56	184-224	11	5.225	0.76	0.76	-0.03
Z7	No	2	PET	54	95-189	31	9.734	0.89	0.91	0.01
Z15	1	1	VIC	56	155-282	39	10.49	0.91	0.92	-0.01
Z16	No	1	6-FAM	56	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Z22	2	1	PET	56	245-277	8	4.934	0.74	0.73	-0.05
Z24	2	2	NED	56	167-211	19	6.253	0.77	0.82	-0.01
Z28	No	2	PET	56	223-251	8	4.812	0.73	0.74	-0.01
Z31	2	2	VIC	56	147-183	10	NA	0.746	0.739	-0.010
Overall						16.82	6.504	0.78	0.79	-0.03

Table S2: Bioclimatic variables and elevation were obtained at high resolution from the French Meteorological Office (Météo-France, Toulouse). NDVI data were monitored by the MODIS device (NASA's Terra mission; http://e4eil01.cr.usgs.gov:22000/WebAccess/drill?attrib=esdt& esdt=MOD13A3.5&group=MOLT).

Variables	Description	Resolution
BIO1	Annual Mean Temperature	132 m x 132 m
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	132 m x 132 m
BIO3	Isothermality (BIO2/BIO7) (* 100)	132 m x 132 m
BIO4	Temperature Seasonality (standard deviation *100)	132 m x 132 m
BIO5	Max Temperature of Warmest Month	132 m x 132 m
BIO6	Min Temperature of Coldest Month	132 m x 132 m
BIO7	Temperature Annual Range (BIO5-BIO6)	132 m x 132 m
BIO8	Mean Temperature of Wettest Quarter	132 m x 132 m
BIO9	Mean Temperature of Driest Quarter	132 m x 132 m
BIO10	Mean Temperature of Warmest Quarter	132 m x 132 m
BIO11	Mean Temperature of Coldest Quarter	132 m x 132 m
BIO12	Annual Precipitation	132 m x 132 m
BIO13	Precipitation of Wettest Month	132 m x 132 m
BIO14	Precipitation of Driest Month	132 m x 132 m
BIO15	Precipitation Seasonality (Coefficient of Variation)	132 m x 132 m
BIO16	Precipitation of Wettest Quarter	132 m x 132 m
BIO17	Precipitation of Driest Quarter	132 m x 132 m
BIO18	Precipitation of Warmest Quarter	132 m x 132 m
BIO19	Precipitation of Coldest Quarter	132 m x 132 m
NDVI	Normalized Difference Vegetation Index	1000 m x 1000 m

Table S3: Factor loadings of the different PCAs on environmental variables.

Variable	PC1
BIO13	0.99
BIO16	0.98
BIO18	0.98
BIO11	-0.94
BIO9	-0.94
BIO11	-0.94
BIO6	-0.93
BIO5	-0.93
BIO10	-0.93
BIO8	-0.93
BIO12	0.89
BIO3	-0.78
BIO2	-0.76
BIO19	0.62
BIO17	0.58
BIO4	0.56
BIO14	0.51
NDVI	-0.27
BIO7	-0.27
BIO15	0.07

Table S4: Pairwise geographic distances (below the diagonal) and estimates of genetic differentiation based on 11 microsatellite markers (above the diagonal) surveyed in four sample localities. Differentiation index corresponds to θST (Weir and Cockerham 1984). Figures in bold face indicate significance at $P < 0.05$.

AMOVA F_{ST}	Tamarins 1	Tamarins 2	Alcide 1	Alcide 2	Feoga 1	Feoga 2	Petit Bernica	Ermitage	Tévelave	Sentier Tamarins	PK7	Makes 1400	Makes 1050	Bon Accueil	Canot	Etang Salé	St-Leu
Maïdo	**0.024**	**0.028**	**0.035**	**0.028**	**0.055**	**0.033**	**0.076**	**0.049**	**0.021**	**0.018**	**0.025**	**0.032**	**0.026**	**0.048**	**0.040**	**0.055**	**0.027**
Tamarins 1	0	**0.015**	0.007	**0.024**	**0.027**	**0.019**	**0.070**	**0.033**	**0.030**	**0.019**	**0.018**	**0.028**	**0.020**	**0.055**	**0.026**	**0.021**	**0.022**
Tamarins 2	3999	0	**0.026**	**0.033**	**0.046**	**0.038**	**0.104**	**0.052**	**0.022**	**0.023**	**0.035**	**0.030**	0.014	**0.055**	**0.042**	**0.042**	**0.040**
Alcide 1	7272	3374	0	**0.018**	**0.025**	0.001	**0.058**	**0.017**	**0.023**	**0.021**	0.007	**0.036**	0.018	**0.037**	**0.038**	**0.027**	0.009
Alcide 2	9588	5591	2763	0	**0.027**	0.008	**0.078**	**0.039**	**0.040**	**0.026**	**0.026**	**0.037**	0.013	**0.027**	**0.028**	**0.025**	**0.026**
Feoga 1	9984	6168	2806	2623	0	**0.020**	**0.077**	**0.038**	**0.051**	**0.040**	**0.030**	**0.043**	**0.014**	**0.053**	**0.026**	**0.026**	**0.034**
Feoga 2	10716	6770	3449	1964	1328	0	**0.046**	0.013	**0.029**	**0.029**	0.009	**0.026**	**0.021**	**0.035**	0.014	**0.022**	**0.019**
Petit Bernica	12497	9921	7613	8723	6114	7190	0	**0.054**	**0.077**	**0.075**	**0.057**	**0.070**	**0.086**	**0.102**	**0.075**	**0.093**	**0.058**
Ermitage	13968	13404	12727	14803	12476	13756	7253	0	**0.042**	**0.033**	**0.015**	**0.029**	**0.037**	**0.043**	**0.046**	**0.034**	**0.025**
Tévelave	6802	10635	13996	16084	16753	17404	18987	18876	0	**0.009**	**0.019**	**0.028**	0.016	**0.054**	**0.031**	**0.061**	**0.021**
Sentier Tamari	6938	10836	14180	16330	16917	17605	18970	18644	572	0	**0.010**	**0.023**	**0.018**	**0.037**	**0.026**	**0.040**	**0.012**
PK7	8498	12494	15750	18072	18401	19199	19798	18582	2925	2376	0	0.019	**0.014**	**0.044**	**0.022**	**0.024**	0.008
Makes 1400	10824	14285	17640	19379	20445	20919	23272	23419	4544	4807	5805	0	**0.025**	**0.041**	**0.024**	**0.045**	**0.031**
Makes 1050	12731	16376	19747	21601	22543	23081	25022	24595	6035	6095	6177	2413	0	**0.038**	0.016	**0.021**	**0.024**
Bon Accueil	10884	14703	18069	20104	20833	21467	22942	22237	4083	3991	3755	2807	2423	0	**0.037**	**0.039**	**0.043**
Canot	13860	17625	20998	22953	23776	24372	25953	25085	7063	7001	6503	4067	1696	3013	0	**0.026**	**0.027**
Etang Salé	16802	20763	23839	26343	26312	27257	26485	23225	11395	10914	8646	11698	10037	8979	8694	0	**0.028**
St-Leu	7621	9975	11892	14656	13675	14875	12490	9619	9872	9523	9061	14314	15191	12787	15544	14119	0

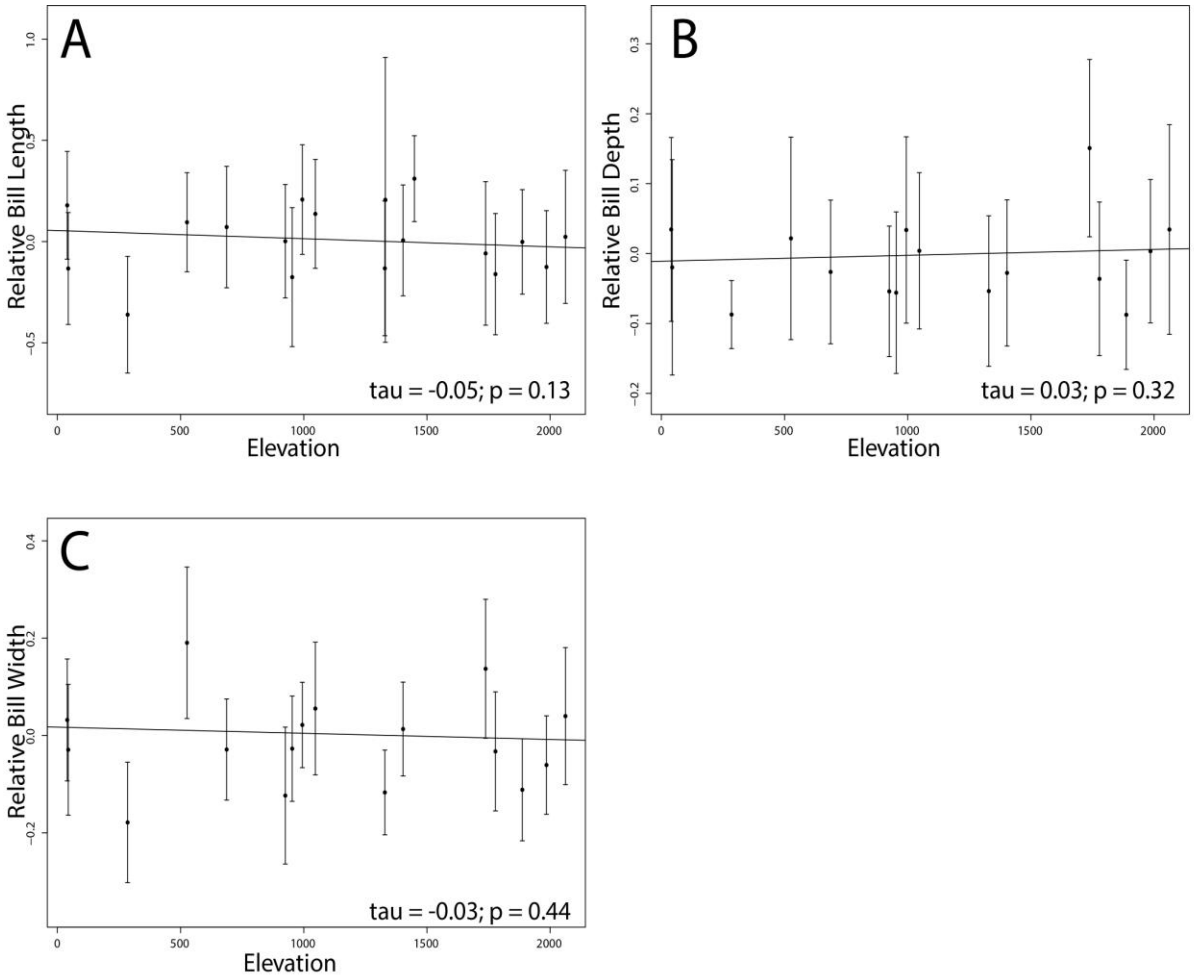

A
Relative Bill Length
tau = -0.05; p = 0.13
Elevation
B
Relative Bill Depth
tau = 0.03; p = 0.32
Elevation
C
Relative Bill Width
tau = -0.03; p = 0.44
Elevation

Figure S1:

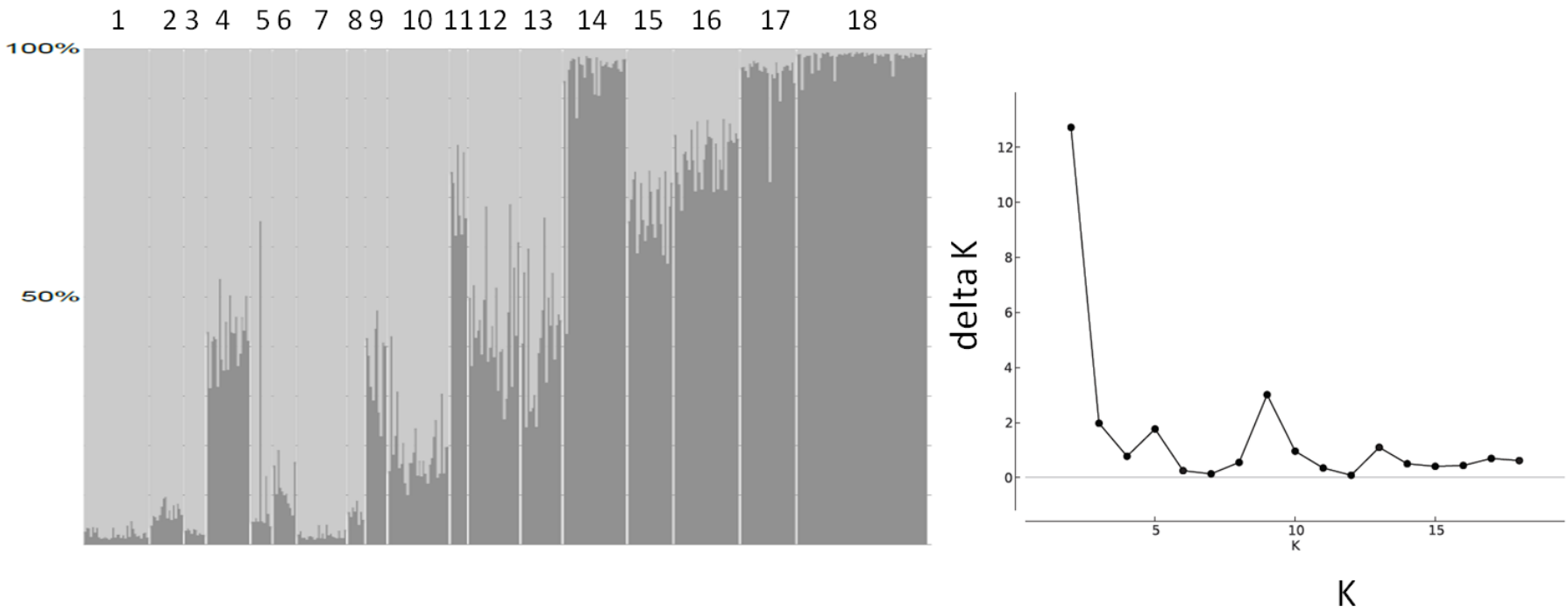


Figure S2: Left) Admixture proportions as inferred from genetic clustering analyses (STRUCTURE). Each bar represents an individual. Each colour reflects the likelihood of belonging to one of the inferred genetic clusters (at $K = 2$). **Right)** delta K statistic plotted against the number of cluster.

Chapitre 3

Zone hybride et gradient altitudinal abrupt

Présentation du chapitre

Ce chapitre traite de la zone hybride mise en évidence dans le chapitre précédent entre les oiseaux bruns de basse altitude (LBHB) et les oiseaux de haute altitude (bruns ou gris, HBHB ou G). Deux transects d'échantillonnage ont été réalisés au travers de cette zone hybride : l'un au nord-ouest de l'île et l'autre au sud-ouest (un certain nombre de localités d'échantillonnage sont partagées avec l'étude précédente). En utilisant des analyses de clines pour comparer les patrons de variation génétique, morphologique et de la couleur du plumage, nous nous sommes posés à trois questions : i) existe-t-il un isolement reproducteur entre la forme de basse altitude et la forme de haute altitude ? ii) observons-nous des patrons de variations phénotypiques compatibles avec l'action de la sélection exogène ? iii) quels sont les traits phénotypiques impliqués dans l'isolement reproducteur entre les formes ?

Contribution : Ce chapitre est constitué d'un article co-rédigé avec Joris Bertrand, nous nous partageons donc la position de premier auteur. Nous avons tous les deux participé à l'ensemble des étapes de la construction de cette article, depuis le terrain jusqu'à la rédaction. Joris a produit la majeure partie des données microsatellites. J'ai produit l'ensemble des données de coloration et effectué la majeure partie des analyses.

Hybrid zones along steep environmental gradients: a case study of two colour forms meeting at mid-elevation on a mountain slope in small-sized passerine bird

In preparation

Boris Delahaie*[1], Joris A. M. Bertrand*[1], Aurélie Khimoun[1], Josselin Cornuault[1], Yann X. C. Bourgeois[1], Thomas Duval[2], Philipp Heeb[1], Borja Milá[3] & Christophe Thébaud[1]

1. Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174, Centre National de la Recherche Scientifique (CNRS) - Université Paul Sabatier, Toulouse 3 - École Nationale de Formation Agronomique (ENFA), 118 Route de Narbonne, F-31062 Toulouse Cedex 9, France.

2. Société Calédonienne d'Ornithologie Nord, Nouvelle-Calédonie, France

3. National Museum of Natural Sciences, Spanish Research Council (CSIC), José Gutiérrez Abascal, 2, 28006 Madrid, Spain.

* These authors have contributed equally to this work.

Correspondance:

Boris Delahaie

Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174, Université Paul Sabatier, Toulouse 3, 118 Route de Narbonne, F-31062 Toulouse Cedex 9, France.

+33 (0)5.61.55.67.23

borisdelahaie@gmail.com

Running title: Narrow hybrid zone on a steep environmental gradient

Abstract

Hybrid zones provide unique opportunities to study the mechanisms underlying evolutionary divergence. Those placed on steep environmental gradient have however received little attention while they potentially allow to distinguish between traits under selection for local adaptation and those involved in reproductive isolation. Here, we used cline analysis to infer the role of selection acting on neutral genetic markers, morphology and plumage colour traits at a narrow hybrid zone between western highland and lowland forms of the Réunion Grey White-eye (*Zosterops borbonicus*), a passerine endemic to Réunion Island found along a steep elevational gradient ranging from sea level to more than 2000 m. Our results provide evidence for narrow cline widths for genetic markers localized around 1400 m above sea level, indicating that partial reproductive barriers exist between the two bird forms. Belly colouration showed a congruent steep cline, suggesting a role for plumage colouration in reproductive isolation among highland and lowland forms. Strikingly, body size did not match the neutral genetic cline and was almost linear over the entire gradient, supporting a role of exogenous selection pressures varying gradually along the gradient leading morphological traits to 'converge' at the contact zone. Our results demonstrate the usefulness of contact zones on environmental gradients for understanding how selection shapes trait variation in hybrid zones.

Keywords: clinal variation, hybrid zone, ecological gradient, phenotypic convergence, *Zosterops borbonicus*

Introduction

Secondary contact following allopatric divergence often leads to the formation of hybrid zones or geographic regions where two genetically distinct populations meet, interbreed and produce hybrids (Barton and Hewitt 1985). Remarkably, many hybrid zones appear to be maintained during tens of thousands of generations despite substantial gene flow (Moore 1977; Barton and Hewitt 1985; Hewitt 1988). The maintenance of the genetic integrity of parental populations is often explained by selection against hybrids (Barton and Hewitt 1985). While selection against hybrids may rely upon intrinsic attributes of species independently of local environmental conditions (tension zone model; Key, 1968; Slatkin, 1973; Barton and Hewitt, 1985), there are numerous cases in which counterselection of hybrids is environment-dependent and may rely upon a balance between the homogenizing effect of gene flow and the diversifying effect of natural or sexual selection between the two sides of the contact zone (environmental gradient selection model; May *et al.*, 1975; Endler, 1977). In this case, exogenous selection entails the maladaptation of hybrids to local ecological conditions, including abiotic or biotic factors such as temperature regimes or the interaction with other organisms at ecological transitions or ecotones (Schluter, 2001). However, it often remains unclear how selection works across hybrids zones and an outstanding issue is to determine to what extant hybrid zones are maintained by exogenous and/or endogenous selection.

Studies combining phenotypic, genetic and ecological data have led to a better understanding of hybrid zone dynamics. In particular, analyzing the environmental configuration of a hybrid zone may help to understand how selection shapes variation and maintains these zones. Hybrid zones located in homogeneous environment are basically expected to be maintained by endogenous barriers to reproduction as environment itself is not supposed to act as a divergent pressure. On the contrary, when hybrid zones are placed on ecotones, it is practically impossible to tease apart the relative contribution of endogenous and exogenous selections (Kruuk et al. 1999). However, a special case of interest is that of hybrid zones that occur along environmental gradients (e.g. elevational gradient; Brennan *et al.*, 2009; Abbott and Brennan, 2014; Dubay et al., 2014). In this case, patterns of variation at neutral genes and phenotypic traits involved in reproductive isolation and especially in inter-specific recognition may be similar to those observed in a tension zone by displaying an abrupt transition because divergence reduces the probability of heterospecific mating at such

contact zones. Alternatively, phenotypic optima at some other traits may vary gradually along the environmental gradient under the effect of exogenous natural selection, for example if selection is locally stabilizing with a moving optimum along the gradient (Endler 1977; Kirkpatrick & Barton 1997). In this case, phenotypes in both taxa would show 'convergence' at the hybrid zone centre with little or no step in the phenotype distribution at these traits, thus leading to cryptic hybrid zones that may be difficult to detect at the phenotypic level alone. Therefore hybrid zones along gradients provide interesting opportunity to distinguish traits that are under exogenous selection from those which are involved in reproductive isolation. A serious limitation is that documented evidences of hybrid zones on environmental gradients but in a spatially continuous area might not be common in nature. For example, only twelve studies on hybrid zones along altitudinal gradients have been referenced by Abbott and Brennan (2014) in a review on altitudinal hybrid zones in plants. In animals, some promising examples exist but are not numerous (Cheviron and Brumfield 2009; Dubay *et al.* 2014; Bertrand et al. *in prep.* - Chapitre 2). Studies that will account for hybrid zones placed on steep environmental gradients will contribute significantly to our understanding of how different modes of selection may shape traits variation and reproductive isolation between closely related taxa.

In this study, we focus on one of these hybrid zones found at a very small spatial scale on a steep environmental gradient between two geographic forms of the Mascarene Grey White-eye (Aves: *Zosterops borbonicus*) species complex. We used the methodological framework provided by cline theory (Endler, 1977; Barton and Hewitt, 1985) to investigate the nature of selection pressures involved by comparing patterns of variation at traits that have been shown to undergo exogenous selection, traits potentially involved in reproductive isolation, and patterns of neutral genetic differentiation. Specifically, we asked (i) Do we find patterns of variation which are consistent with the action of locally varying stabilizing selection? (ii) Do we detect differences between traits potentially involved in reproductive isolation and other traits?

To address these questions, we fitted clines on several characteristics related to genetics, coloration and morphological variation in *Z. borbonicus* along two replicate elevational transects spanning 2000 m in elevation across a 15 km linear distance. We compared cline positions (i.e. concordance of centers or not) and widths (i.e. concordance of slopes or not) to infer the role of selection acting on these traits. For genetics, we characterized variation at eleven microsatellites loci specifically developed for the species

(Bertrand *et al.* 2012). For coloration, an avian visual model was used to project plumage color spectral measurements in an avian-appropriate, tetrachromatic, color space (Goldsmith, 1990; Endler & Mielke, 2005). By using a tetrahedral avian color space instead of a direct analysis of reflectance spectra, we were able to examine how plumage color as perceived by the birds themselves varied along the two transects (Endler & Mielke, 2005; Stoddard & Prum, 2008). For morphology, clines were fitted at four morphological traits in order to characterize variation at body-size related traits and bill characteristics. According to clear evidences for a break at neutral genetic variation we previously documented at mid-elevations (Bertrand et al. *in prep* - Chapitre 2), we predicted that clines for traits involved in reproductive isolation between forms would show narrow cline width. By contrast, clinal variation at morphological traits under gradual exogenous elevational selection (i.e. here for which fitness optima covary gradually with environmental conditions) would be expected to be wider than the neutral cline with possibly offset in centers positions depending on the trait considered.

Material and methods

Study system

Zosterops borbonicus, an endemic passerine to Réunion (Mascarene Islands, Indian Ocean), represents an extraordinary case of within island diversification (Gill 1973; Milá et al. 2010). It is composed of four different forms whose distribution is geographically structured and temporally stable. The different forms are separated by narrow contact zones such as river beds and lava flows in the lowlands and less defined elevational transition zones between lowlands and uplands (Gill 1973; Mila et al. 2010). The highland form is polymorphic and comprises two plumage colour morphs: a grey one and a brown one, which do not exhibit any other morphological or neutral genetic difference distinguishing one from the other (Bourgeois 2013; Bertrand et al. *in prep.* - Chapitre 2). The lowland brown variant meets this polymorphic form on a steep elevation gradient on the western side of Réunion and form a hybrid zone at mid-elevations along the gradient (Bertrand et al. *in prep.* - Chapitre 2). This situation has been stable for at least 50 years, as the current distribution pattern matches exactly the pattern described by Frank Gill in the 1960s (Milá et al. 2010; Bertrand 2013).

Field sampling

Birds were sampled along two elevational transects crossing the hybrid zone between the lowland brown-headed brown and the highland grey/brown form of *Z. borbonicus* on the western slopes of the island of Réunion. The two transects were separated by an average of 20 km. The first (northern) transect (hereafter referred as the T1 transect) was located between the sampling localities Cambaie (T1-1) and Maïdo (T1-11), whereas the second (southern) transect (hereafter referred as the T2 transect) was located between the localities Étang du Gol (T2-1) and Tévelave (T2-9) (Fig. 1).

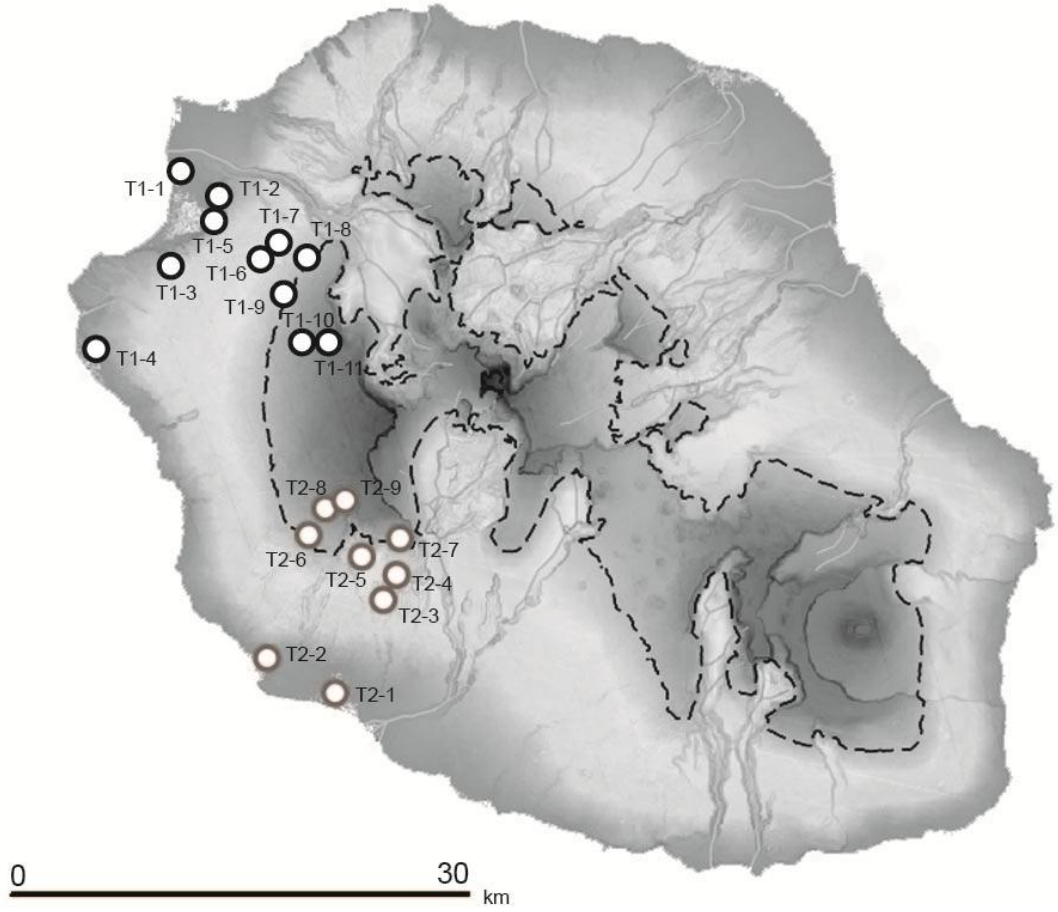


Figure 1: Map of Réunion Island with the geographical position of sampling sites. Dashed line indicates the 1400 m a.s.l. elevation.

Both transects run from near sea-level up to 2000 m elevation over distances of about 15 km (Table 1): *ca.* 150 m elevational gain for every kilometre of linear distance. These transects crossed five distinct ecological zones (as defined by Strasberg et al. 2005) from lowlands to highlands: lowland dry forest, semi-dry sclerophyllous forest, lowland rainforest, cloud forest and subalpine shrubland (see Table 1). Twenty localities were sampled along the two transects (Fig. 1; see Table 1 for details on localities and sample sizes). A total of 483 birds were sampled across these localities, using mist nets and following the procedures described in Milá et al. (2010).

Table 1. Summary of sampling station characteristics. Distance corresponds to the linear distance to the first locality of each transect (T1-1 and T2-1).

Sampling locality	Latitude	Longitude	Elevation	Distance (m)	Ecological zone	n_{tot}	n_{gen}	n_{tarsus}	n_{length}	n_{width}	n_{depth}	n_{back} (PC1)	n_{back} (PC2)	n_{belly} (PC1)	n_{belly} (PC2)	n_{head} (PC1)	n_{head} (PC2)	n_{flank} (PC1)	n_{flank} (PC2)	Prop. of grey birds (%)
T1-1	-20.97	55.28	8	0	Dry Lowland Forest	45	15	29	13	21	13	22	15	14	21	12	19	9	15	0
T1-2	-20.98	55.31	26	3287	Dry Lowland Forest	15	13	10	0	0	0	13	9	0	0	9	12	8	12	0
T1-3	-21.03	55.28	287	4129	Semi-Dry Lowland Forest	29	10	20	17	22	17	24	15	1	6	1	6	10	19	0
T1-4	-21.07	55.23	52	4153	Dry Lowland Forest	18	16	15	15	16	15	17	16	14	15	14	16	16	17	0
T1-5	-21.00	55.31	25	4195	Dry Lowland Forest	5	5	0	0	0	0	5	2	0	0	0	0	2	4	0
T1-6	-21.02	55.34	925	8373	Lowland Rainforest	25	24	24	24	25	24	24	23	20	21	23	24	22	23	4
T1-7	-21.01	55.35	956	8436	Lowland Rainforest	10	10	9	9	10	9	8	7	7	8	7	8	7	8	20
T1-8	-21.02	55.36	1328	10432	Cloud Forest	30	30	30	30	30	30	27	27	23	23	27	27	25	25	10
T1-9	-21.04	55.35	1324	11102	Cloud Forest	8	8	7	0	0	0	4	4	0	0	4	4	4	4	50
T1-10	-21.07	55.36	1731	14204	Cloud Forest	32	31	28	28	31	27	7	7	8	8	8	8	8	8	75
T1-11	-21.07	55.38	2055	15481	Subalpine Shrubland	64	64	63	63	64	63	10	10	10	10	10	10	10	10	70
T2-1	-21.28	55.38	4	0	Dry Lowland Forest	34	34	9	1	9	1	19	8	7	18	5	16	5	15	0
T2-2	-21.26	55.34	43	1677	Dry Lowland Forest	33	32	28	28	32	28	31	27	0	0	0	0	26	30	0
T2-3	-21.22	55.41	688	6710	Semi-Dry Lowland Forest	10	10	4	4	10	4	9	4	3	6	3	7	3	8	10
T2-4	-21.21	55.42	1041	8366	Lowland Rainforest	10	10	9	9	9	9	8	7	6	7	7	8	7	7	20
T2-5	-21.20	55.40	980	9150	Lowland Rainforest	8	8	5	5	8	5	7	4	4	6	4	7	4	7	12.5
T2-6	-21.19	55.37	1447	10215	Cloud Forest	21	20	18	0	0	0	12	11	0	0	12	13	14	16	19
T2-7	-21.19	55.42	1408	10789	Cloud Forest	25	25	12	12	25	12	19	9	9	17	9	19	7	17	20
T2-8	-21.17	55.38	1889	12275	Cloud Forest	34	32	23	23	32	23	16	13	9	12	12	14	13	16	47
T2-9	-21.17	55.39	1975	12495	Subalpine Shrubland	27	27	16	15	27	16	20	13	9	15	13	19	10	17	22

Individuals caught were aged as 'juvenile' or 'adult', whenever possible, using plumage characteristics, eye and gape colour, molt pattern and the degree of skull ossification, following criteria in Pyle (1997). Morphological measurements were taken at four morphological traits (tarsus length, bill length, depth and width, see Table 1 for details about sample sizes). Dial callipers were used to measure tarsus length (from intertarsal joint to the most distal undivided scute on the tarsometatarsus), bill length (from the anterior end of the nares to the tip of the upper mandible), and bill width and depth (both measured at the anterior end of the nares). Tarsus length is a good proxy for overall body size (*e.g.* Senar and Pascual 1997). Bill characteristics are mainly related to foraging, song and heat regulation (*e.g.* Edelaar et al. 2012; Caro et al., 2013; Greenberg et al., 2012). To characterize plumage colour, we collected feather samples on each bird captured (see Table 1 for details about sample sizes). At least ten feathers were collected from each of the following body regions on each bird: the head, back, flank and belly of each bird.

Molecular markers and genetic variation

We genotyped 424 (see Table 1 for details about sample sizes) individuals at 11 microsatellite loci that were specifically developed for *Zosterops borbonicus* (Z1, Z2, Z3, Z4, Z5, Z7, Z15, Z22, Z24, Z28, Z31; Bertrand *et al*. 2012). Amplification protocols and genotyping details are described elsewhere (see Bertrand et al. 2012; 2014). For the highland form, all analyses were done disregarding the specific colour morph as they do not significantly differ at these microsatellites (Bourgeois 2013). We performed basic tests to validate the reliability of our microsatellite dataset and assessed within-population genetic variation to quantify genetic diversity. The presence of null alleles was tested with MICRO-CHECKER v.2.2.3 (van Oousterhout et al. 2004). We checked for linkage disequilibrium among loci and estimated the allelic richness corrected for sample size (A_R). We also quantified within-population genetic variation by calculating the mean number of alleles per locus (A) along with expected and observed heterozygosities (H_E and H_O). Inbreeding coefficients (F_{IS}) and deviation from Hardy-Weinberg equilibrium were also estimated. R-package {diveRsity} was used was used to make all these calculations.

Variation in morphology and colour traits

For the highland form, all morphological analyses were done without regarding the morph as browns and greys birds have been shown to be morphologically and genetically similar (Bourgeois 2013). For coloration analyses, each bird was assigned visually to the brown or grey morph since the two morphs are easily distinguishable based on their plumage differences (Gill 1973; Cornuault *et al*. *in press* - Chapitre 1). Grey birds were then removed from all subsequent analysis in order to only focus on the quantification of subtle and cryptic variation in coloration across the transects. We used reflectance spectrophotometry to characterize variation in plumage colour because it provides an objective quantification of colour (Cuthill 1999; Endler 1990). Spectral data were recorded in the laboratory with a USB 2000 spectrophotometer connected to a PX-2 light source via a Qt-200 bifurcate optical fibre probe (Ocean Optics, Dunedin, FL, USA). Before each measurement, about 10 feathers were placed on a black surface in a fashion that mimicked the way feathers naturally lay on the bird. We positioned the optical fibre at a standardized three-millimetres distance from the feather surface at a 90° angle. Reflectance was measured relative to a white standard and each measurement consisted of three replicates that were averaged before analysis. After each individual replicate measurement, the feathers were separated and piled up again in a random order. For each colour patch, we calculated the brilliance (*B*) by summing the reflectance values over all wavelengths. We used the tetrahedral colour space model for the analysis of reflectance spectra (Goldsmith 1990; Endler and Mielke 2005). This model has several advantages as it allows the calculation of different meaningful variables in terms of bird vision (Endler and Mielke 2005; Stoddard and Prum 2008). All the following calculations were done using the package {pavo} (Maia et al. 2013) from R software. Following Goldsmith (1990), we calculated the relative idealized stimulation of the four colour cones {*u s m l*} using spectral sensitivity functions of the Blue Tit (*Cyanistes caeruleus)* and an irradiance spectrum which is constant across all visible wavelengths (as suggested by Goldsmith (1990) and Stoddard and Prum 2008). Each plumage colour patch was then described by a set of four values {*u s m l*} corresponding respectively to the relative stimulation of the ultraviolet-sensitive, short-wavelength-sensitive, medium-wavelength-sensitive and long-wavelength-sensitive avian retinal cones. This way of describing colour

takes into account the perception of colours by birds, and thus allows analyzing colour phenotypes as birds actually see them. Following Endler and Mielke (2005), we converted the relative stimulation values of each colour patch into their spherical coordinates (θ, φ and r) that define the colour vector in tetrahedral space. The angles θ and φ define the hue of the colour. r is defined as the length of the colour vector and corresponds to the chroma of the colour. Following Stoddard and Prum (2008), we calculated the achieved chroma (r_a) which corresponds to the chroma of a colour relative to the maximum chroma given its hue, which is more informative than r. Each colour patch is then described by 4 variables (B, θ, φ, r_a). For each body patch, we summarized colour variation by conducting a principal components analysis (PCA) on the set of 4 variables.

Quantifying phenotypic differences between the lowland and highland forms

We first used ANOVAs to test for potential age and sex effects on phenotypic traits (tarsus length, bill variables as well as the two first PC scores of each PCA for plumage coloration). For traits that differed by sex and/or age, we used the residuals of these ANOVA models for all further analyses in order to remove these potentially confounding effects.

Phenotypic traits were then compared between the lowland and highland forms to quantify differentiation between forms using ANOVAs. To perform these comparisons, we defined reference sampling localities as the terminal sites of each transect: T1-1 and T2-1 (for the lowland's form); T1-11 and T2-9 (for the highland's form).

All analyses were performed with *R v2.13.0* (R Development Core Team 2011).

Cline analysis

Level of admixture

To quantify the level of admixture which occurs between the two forms, we conducted an analysis of the population structure using the Bayesian model-based analysis implemented in STRUCTURE *v.2.3.2* (Pritchard et al. 2000). We applied the admixture model (Falush et al.

2003) with correlated allele frequencies and prior information on sampling locality (*locprior* model), a procedure that increases the algorithm's ability to find population clusters when the amount of genetic differentiation is limited, yet having no effect on the optimal number of clusters inferred (Hubisz et al. 2009). We conducted this analysis with a number of clusters ranging from one to eleven to determine whether a model with two clusters was best at explaining the data. Ten runs were performed for each *k* value. For each run, the program STRUCTURE used Markov Chain Monte Carlo (MCMC) chains of 600,000 iterations of which the first 100,000 were discarded as burn-in. Support for the optimal number of clusters was obtained by plotting the *ad hoc delta K* statistic following the method of Evanno et al. (2005). All individuals were assigned to clusters according to the outputs of *clump v.1.1.2b* (Jakobsson and Rosenberg 2007) which accounts for the variability in individual assignment probabilities across the different runs. Posterior probabilities of lowland cluster assignment (extracted from the *Q-matrix* of *structure* result) were used to define a genetic hybrid index. We decided to use a multilocus hybrid index instead of individual locus information because the very high level of polymorphism at individual loci prevent the use of common methods for summarizing variation at each locus (*e.g.* Bermond *et al.* 2012; Delmore *et al.* 2013; Smith *et al.* 2013). This index provides a measure of the genetic makeup of individuals by quantifying the genetic contribution of hybridizing clusters to individuals of unknown ancestry (hereafter referred as genetic HI). Individuals were considered as 'pure' lowland form when their genetic HI was smaller than 0.2, as 'pure' highland form when their genetic HI was greater than 0.8 and as hybrids for values between 0.2 and 0.8.

Orientation of the transect

Because cline models only handle one spatial dimension, we had to find out the transect line that best fitted our dataset. To determine the orientation of each transect, we estimated the direction of the maximum gradient of allele frequency change across the hybrid zone by fitting the genetic HI to a binomial generalized linear model (Raufaste et al. 2005). In these models, the logit transform of the genetic HI is a linear function of the geographical coordinates and we assume that genetic HI variation is sigmoidal and that the centre of the hybrid zone is a straight line. These models were fitted by maximum likelihood using the 'nlminb' optimizing function (R Development Core Team 2011). We then calculated transformed coordinates of each locality by projecting the original geographic coordinates

onto this new transect. These new coordinates correspond to the distance to the first lowland locality projected on the transect (T1-1 and T2-1, for the northern and the southern transect respectively).

Spatial cline analysis

Cline models were fitted to genetic HI and to the different phenotypic traits independently on each transect. Models were adjusted by maximum likelihood. We used a simple sigmoid cline model for each variable:

$$\boldsymbol{p_x} = \frac{\mathbf{1}}{\mathbf{1}+e^{\frac{\mathbf{4}}{\boldsymbol{w}}(\boldsymbol{x}-\boldsymbol{c})}} \qquad \text{(equation 1a)}$$

$$\boldsymbol{\mu_{x}} = \boldsymbol{p_{min}} + (\boldsymbol{p_{max}} - \boldsymbol{p_{min}})\boldsymbol{p_x} \qquad \text{(equation 1b)}$$

where p_x (equation 1a) is a monotone sigmoid function, c and w are the centre and width of the cline, x corresponds to the distance along the transect, μ_x (equation 1b) is the mean value of the trait, and p_{min} and p_{max} correspond to the minimum and maximum values of the trait, respectively.

Likelihood functions were optimized using the 'nlminb' function (R Development Core Team 2011). We allowed cline centre (c) to vary between the minimum and the maximum distance of each transect ([0:15481] for T1; [0:12495] for T2), cline width (w) to vary between [0,+∞[, and p_{min} and p_{max} to range between]-∞,+∞[for quantitative trait and between [0,1] for genetic HI. To determine whether cline centre or width differed between genetic markers and phenotypic traits (test for coincidence and concordance), we built a set of constrained and unconstrained models for each transect that included the following: (i) a model with constrained common centre and slope between each phenotypic trait and genetic HI to determine the single best-fitting curve for each trait combination; (ii) a model with common centre between each phenotypic trait and genetic HI; (iii) a model in which genetic HI and each phenotypic trait vary independently; and (iv) a model in which genetic HI varies

in a sigmoidal manner and where phenotypic variation is independent from the location on the transect. From each model, we extracted the maximum likelihood estimates of cline parameters (c, w, p_{min} and p_{max}) and their 95% confidence intervals. Confidence intervals were determined by approximating the sampling distribution by simulation from the best model (parametric bootstrap with 1000 samples). The different models were compared using AIC weights (AICw) following the approach described by Burnham and Anderson (1998).

All these analyses were done using R software *v2.13.0* (R Development Core Team 2011).

Neutral diffusion expectations

To test whether the neutral genetic cline is maintained by selection, we used the diffusion equation proposed by Barton and Gale (1993) to simulate the expected cline width following a secondary contact and subsequent homogenisation under neutral processes alone. If a barrier to gene flow is absent, neutral genetic clines are expected to become wider with time since secondary contact. Thus, if we assume that the neutral genetic cline is the result of secondary contact between highland and lowland populations, we can calculate the expected neutral genetic cline width using $2.51\sigma\sqrt{t}$, where σ is the root mean square (RMS, a surrogate of dispersal distance) and t is the time in generations since secondary contact. RMS for *Z. borbonicus* was estimated at 0.207 km/generation, a figure that to the best of our knowledge, stands among the smallest reported for birds (Bertrand 2013). We examined the relationship between the expected cline width under neutral diffusion and the time since secondary contact. We then determined the time since secondary contact needed to obtain the observed cline width. To make these calculations, we assumed a reasonable generation time for a passerine of two years.

Results

None of the 11 microsatellite loci exhibited significant deviation from Hardy-Weinberg equilibrium, and no significant linkage disequilibrium was found between pairs of loci. Microsatellite loci presented differences in their relative level of polymorphism, with

number of alleles per locus ranging from 7 (for Z1 and Z2) to 38 (for Z15). This important allelic polymorphism was associated with high mean heterozygosities (H_O = 0.76 and H_E = 0.78). The mean inbreeding coefficient (F_{IS}) could not be statistically differentiated from zero in all but two localities (T1-3, F_{IS} = - 0.09 and T2-9, F_{IS} = 0.07) indicating no major significant deviation from panmixia.

Results of STRUCTURE analysis confirmed that the most likely number of populations is two. The *delta K* statistic is more than ten times greater for $K = 2$ than for all other values of *K* tested (*delta* $K_{(K = 2)}$ = 45.7; Fig. S1). Assignment probabilities to these clusters are consistent with previous results obtained by Milá et al. 2010 and Bertrand *et al.* (*in prep* - Chapitre 2), with high altitude populations belonging to one cluster (*K highland*, Fig. 2) and low altitude populations belonging to the other cluster (*K lowland*, Fig**.** 2). One-hundred and fourteen individuals out of the 424 genotyped were found to be hybrids (26.9 %), with 52 hybrids out of 226 individuals (23 %) on the northern transect (T1) and 62 hybrids out of 198 individuals (31.3 %) on the southern transect (T2). On both transects, the occurrence of individuals with admixed ancestry was maximal at intermediate elevations. Indeed, on T1, 73.1% of the hybrids were found between 1300 m and 1400 m (T1-8 and T1-9; Fig. 2). On T2, most of the hybrids were found between 1000 m and 1500 m (95.2 %; T2-4, T2-5, T2-6, T2-7; Fig. 2).

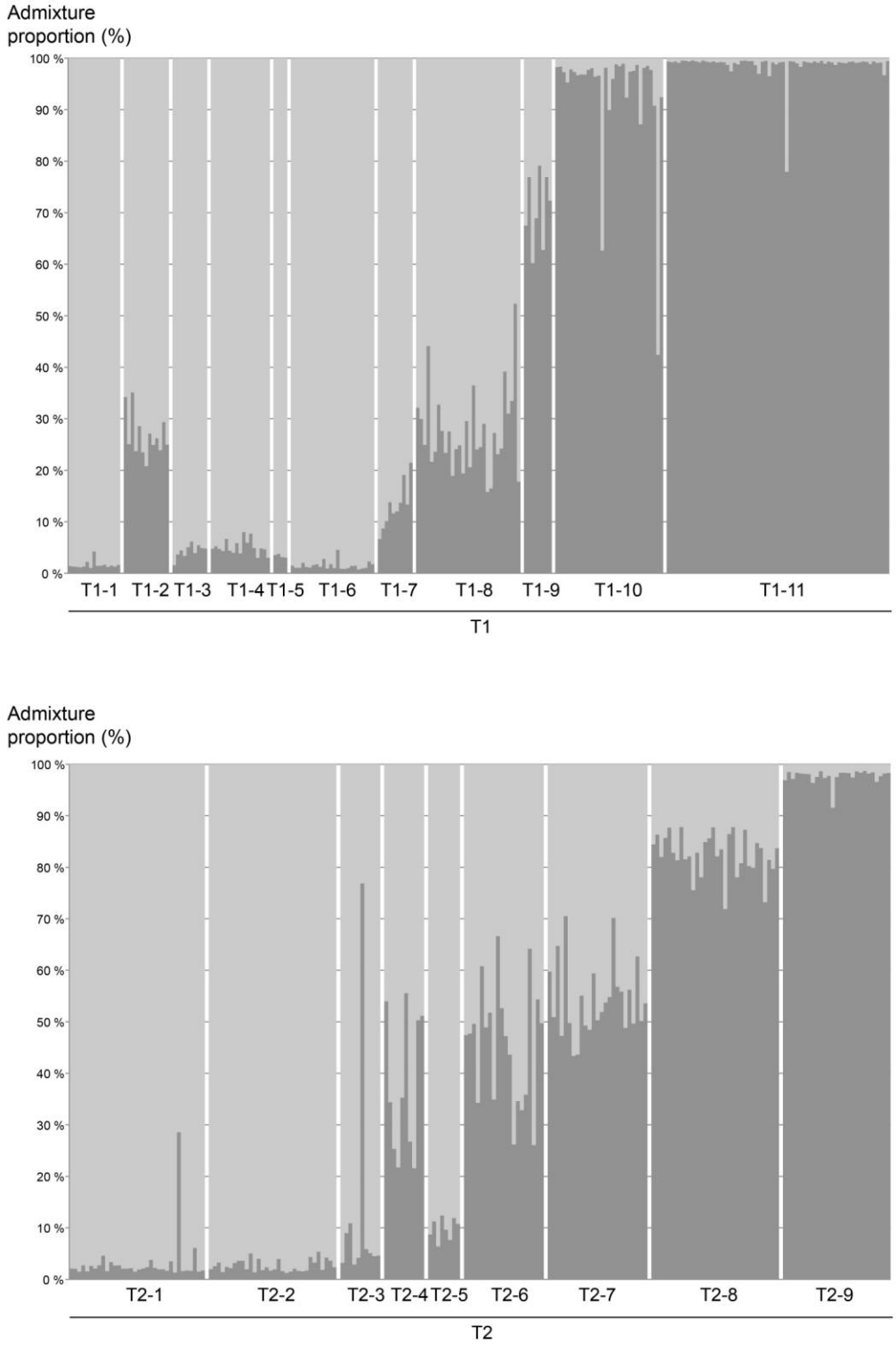


Figure 2: Admixture proportions inferred from STRUCTURE clustering. Graphic of the Q matrix after pooling all runs of *clump v.1.1.2b* for $k = 2$. Sampling stations are first organized by transect and by position on the transect lowland stations are on the left of each transect and high altitude populations on right of each transect.

Phenotypic differentiation

We found significant effects of age and/or sex on most of the phenotypic traits (Table 2). Lowland and highland birds (from reference sites) differed significantly in their phenotypes for back colouration (PC1), belly colouration (PC2), head colouration (PC2) and tarsus length (Table 2 and Fig. 3; see Table 3 for details about the factor loadings of the different PCAs).

Table 2: Age, sex and form effects on phenotypic traits (only brown birds for colour variables). Bold values indicate significant differences.

Variable	Age effect	Sex effect	Form effect
Back PC1	$F_{1,298} = 0.86$; p = 0.35	$F_{1,298} = 0.58$; p = 0.45	**$F_{1,69} = 14.81$; p = $2.6*10^{-4}$**
Back PC2	**$F_{1,298} = 10.99$; p = $1.0*10^{-3}$**	$F_{1,298} = 0.12$; p = 0.73	$F_{1,44} = 1.31$; p = 0.26
Belly PC1	**$F_{1,138} = 35.20$; p = $2.3*10^{-8}$**	$F_{1,138} = 0.43$; p = 0.52	$F_{1,38} = 0.15$; p = 0.70
Belly PC2	$F_{1,138} = 2.55$; p = 0.11	**$F_{1,138} = 4.42$; p = $3.7*10^{-2}$**	**$F_{1,62} = 56.49$; p = $2.7*10^{-10}$**
Flank PC1	**$F_{1,204} = 29.62$; p = $1.50*10^{-7}$**	$F_{1,204} = 0.37$; p = 0.54	$F_{1,32} = 0.83$; p = 0.37
Flank PC2	**$F_{1,204} = 59.25$; p = $5.88*10^{-13}$**	$F_{1,204} = 206$; p = 0.15	$F_{1,55} = 0.50$; p = 0.48
Head PC1	**$F_{1,175} = 19.11$; p = $2.11*10^{-5}$**	**$F_{1,175} = 5.62$; p = $1.9*10^{-2}$**	$F_{1,38} = 0.60$; p = 0.44
Head PC2	$F_{1,175} = 2.31$; p = 0.13	$F_{1,175} = 0.41$; p = 0.52	**$F_{1,62} = 9.01$; p = $3.9*10^{-3}$**
Tarsus	**$F_{1,356} = 9.00$; p = $2.9*10^{-3}$**	**$F_{1,356} = 7.68$; p = $5.9*10^{-3}$**	**$F_{1,115} = 289.47$; p < $2.2*10^{-16}$**
Width	$F_{1,292} = 0.51$; p = 0.48	**$F_{1,292} = 6.70$; p = $1.0*10^{-2}$**	**$F_{1,119} = 8.81$; p = $3.6*10^{-3}$**
Depth	**$F_{1,293} = 9.52$; p = $2.2*10^{-3}$**	**$F_{1,293} = 8.98$; p = $3.0*10^{-3}$**	$F_{1,91} = 0.74$; p = 0.39
Culmen	**$F_{1,293} = 33.10$; p = $2.2*10^{-8}$**	**$F_{1,293} = 10.26$; p = $1.5*10^{-3}$**	$F_{1,90} = 2.34$; p = 0.13

Table 3: Factor loadings of the different PCAs on plumage colouration.

	Back		**Belly**		**Flanks**		**Head**	
	PC1	**PC2**	**PC1**	**PC2**	**PC1**	**PC2**	**PC1**	**PC2**
B	-0.63	0.65	-0.4	0.83	-0.77	-0.24	0.1	0.94
ϑ	0.6	0.45	-0.82	0.27	-0.93	-0.02	0.75	0.15
φ	-0.14	-0.75	-0.84	-0.25	0.72	-0.68	-0.73	-0.31
r_a	0.92	0.04	0.66	0.51	0.84	0.33	0.68	-0.65
Explained Variance (%)	40	30	49	27	68	16	0.39	0.36

Highland brown birds had higher values for chroma in back colouration, and less brilliant bellies and heads than lowland birds (Fig. 3A, B and C). Belly colour (PC2) was the most

discriminating colour variable between highland brown birds and lowland ones (Table 2, Fig. 3A, B and C). Highland birds had significantly longer tarsi than lowland ones (Fig. 3D). Bill was significantly wider in highland birds, although there was a large overlap between forms (Fig. 3E).

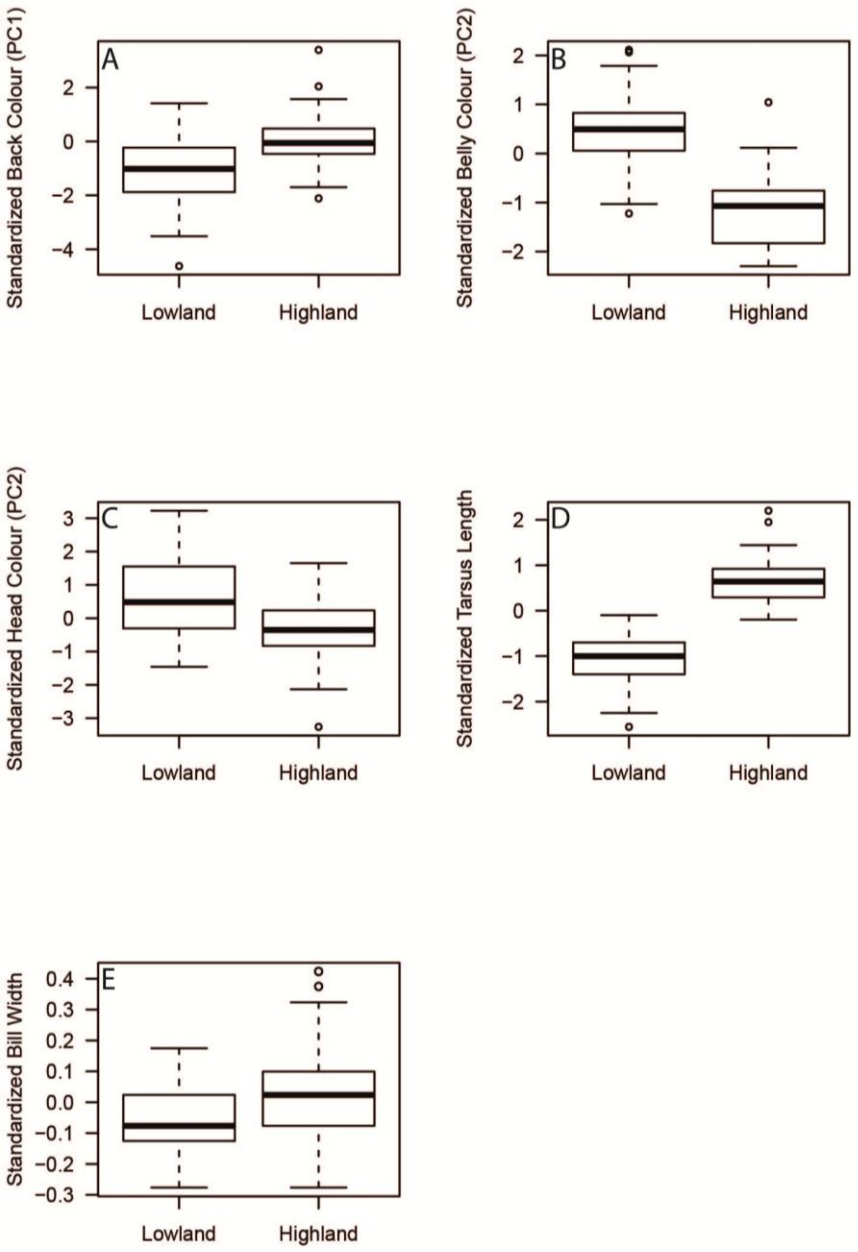


Figure 3: Differences between lowland and highland reference sites for phenotypic traits (only brown birds for colour variables).

Cline analysis

All cline models were fitted successfully. However, unsurprisingly, for traits that were not different between highland and lowland sites (*i.e.* back colour (PC2), belly colour (PC1), flank colour (PC1 and PC2), head colour (PC1) and bill depth and length), we obtained very wide confidence intervals for the different parameters and/or equivalent weights for the different models. Therefore, we did not interpret the results of cline comparisons at these traits.

Unconstrained maximum likelihood estimates for the centre of genetic cline on T1 was 10826 m (95% CI: 10769 – 10899) and 1262 m (95% CI: 1053 - 1512) for width (Fig. 4A). For the southern transect (T2), the estimated centre for the genetic HI was 10533 m (95% CI: 10291 - 10629) and 4035 m (95% CI: 3105 - 4458) for width (Fig. 4B). These centres are located between T1-8 and T1-9 on the northern transect (T1) and between T2-6 and T2-7 on the southern transect (T2). Considering neutral genetic markers, the hybrid zone centre is located in the cloud forest ecological zone (between 1400 and 1500 m). Cline width was twice as wide on the southern transect than on the northern one and their confidence intervals did not overlap.

Table 4: Model comparison for each variable that differ between lowland and highland forms. Maximum likelihood estimates and their confidence intervals are given for the different parameters of best models for each trait comparison.

			T1				T2			
Comparison	Model	Parameters	AIC	AICw	Centre	Width	AIC	AICw	Centre	Width
Tarsus	**Centre and slope constrained**	8	-34.41	0.00	-	-	**-194.96**	**0.32**	**10516 (10270 - 10622)**	**4107 (3110 - 4466)**
	Centre constrained	**9**	**-114.10**	**0.18**	**10826 (10764 - 10892)**	**49636 (18342 - 17500574)**	**-195.83**	**0.48**	**10534 (10298 - 10629)**	**8637 (3926 - 30352)**
	Unconstrained	**10**	**-117.07**	**0.82**	**7653 (5124 - 9585)**	**12519 (7056 - 35181)**	**-194.11**	**0.21**	**12495 (8083 - 12495)**	**12422 (1186 - 32208)**
	Null	7	118.23	0.00	-	-	-131.25	0.00	-	-
Belly PC2	**Centre and slope constrained**	**8**	**-226.58**	**0.66**	**10826 (10768 - 10888)**	**1261 (1084 - 1507)**	**-183**	**0.61**	**10530 (10296 - 10636)**	**4051 (3113 - 4441)**
	Centre constrained	**9**	**-224.59**	**0.24**	**10826 (10764 - 10903)**	**1217 (552 - 4550)**	**-181.44**	**0.28**	**10533 (10305 - 10636)**	**6249 (14 - 686017)**
	Unconstrained	10	-222.66	0.09	-	-	-179.47	0.10	-	-
	Null	7	-154.16	0.00	-	-	-169.54	0.00	-	-
Back PC1	Centre and slope constrained	8	69.22	0.00	-	-	41.96	0.00	-	-
	Centre constrained	9	52.90	0.19	-	-	43.92	0.00	-	-
	Unconstrained	10	50.01	0.81	54 (0 -8339)	14217 (120 - 39971)	29.79	0.99	0 (0 -1661)	230 (0 - 3854)
	Null	7	77.20	0.00	-	-	40.51	0.00	-	-
Head PC2	Centre and slope constrained	8	-27.12	0.03	-	-	-114.38	0.08	-	-
	Centre constrained	9	-30.80	0.20	-	-	-116.92	0.29	10533 (10297 - 10642)	769490 (39 - 10626128)
	Unconstrained	10	-33.53	0.76	8393 (0 - 12238)	94 (0 - 27720)	-117.98	0.49	782 (0 - 12495)	560 (17 - 20503)
	Null	7	-24.57	0.01	-	-	-115.4	0.14	-	-
Width	Centre and slope constrained	8	-712.40	0.06	-	-	-635.08	0.30	10541 (10296 -10639)	4012 (3039 - 4405)
	Centre constrained	9	-717.25	0.69	10826 (10771 - 10889)	16 (0 - 6758)	-631.6	0.05	-	-
	Unconstrained	10	-715.25	0.25	12809 (10488 - 15480)	140 (4 - 10457)	-636.61	0.65	11017 (10767 - 12495)	240 (16 - 8083)
	Null	7	-672.72	0.00	-	-	-619.02	0.00	-	-

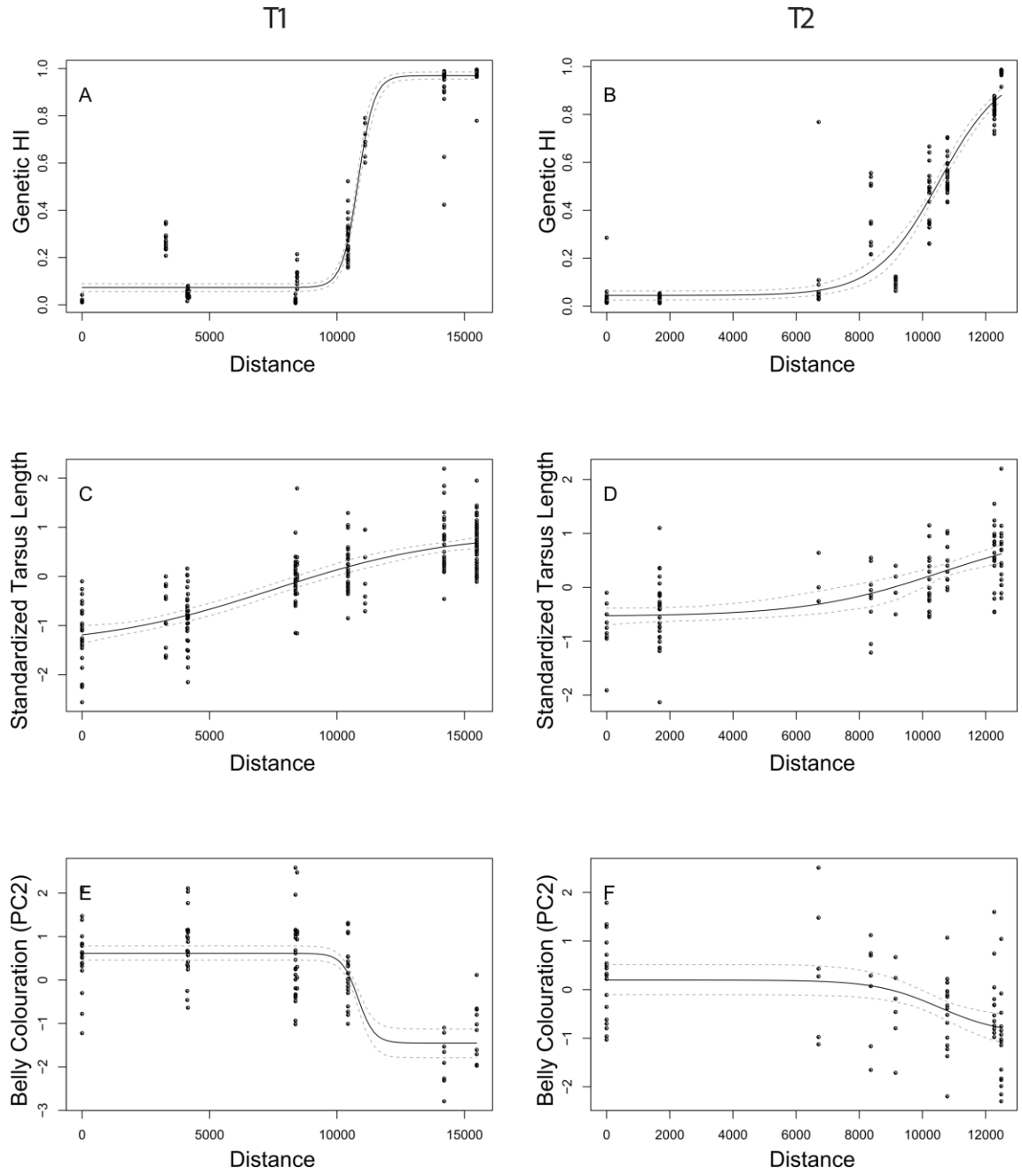


Figure 4. Spatial cline in genetic and phenotypic markers in the *Zosterops borbonicus* hybrid zone. Genetic markers on T1 (A) and T2 (B), tarsus length on T1 (C) and T2 (D), belly coloration on T1 (E) and flank coloration on T1 (F). Black points correspond to the observed values, black solid lines denote the best-fitting curve of each variable and dashed lines correspond to 95% confidence intervals. Abscissa corresponds to the distance to the first lowland locality projected on the transect (T1-1 and T2-1, for the northern and the southern transect respectively).

For the comparison between tarsus length and genetic HI on T1, the unconstrained model was found to be the best model. On this transect, cline width for tarsus length was much wider than neutral genetic cline width, with no overlap in their confidence intervals (Table 4, Fig. 4C). Centres for the genetic cline and the tarsus cline did not overlap. However, differences in centre location may not be meaningful as clines for tarsus length appeared to be very wide and approaching straight lines. On the southern transect the results were less clear, with the three clinal models having almost equal weight (Table 4). These results are probably due to the absence of data for a non-negligible part of the transect (between T2-2 and T2-3; Table 1). However, the most likely model gave estimates of tarsus cline width twice as large as the genetic one, which corresponds to a quasi-linear relationship with geographic distance (Fig. 4D).

On both transects, results for belly coloration clines (PC2) contrasted with morphological model results. These clines showed a similar pattern to the neutral genetic clines and the model with constrained centre and slope had the highest AICw (0.66 and 0.61 for T1 and T2, respectively). Belly colour (PC2) thus showed a steep step at 10826 m (95% CI: 10768 - 10888) on T1, and 10530 m (95% CI: 10296 -10636) on T2 (Table 4; Fig. 4E and 4F). Maximum likelihood width estimates for belly were estimated at 1261 m (95% CI: 1084 - 1507) and 4051 m (95% CI: 3113 - 4441) on T1 and T2, respectively. On both transects, the centre-constrained model received less support (AICw: 0.24 and 0.28, respectively). These two models were quantitatively similar to centre and slope constrained ones. However, they gave larger confidence intervals for width, especially for the southern transect (T2, Table 4).

Additionally, clines for back, head colour and bill width gave either very wide confidence intervals or senseless parameter estimates like centres at the very beginning of the transect, a pattern apparently due to the non-sigmoid form of the variation at these traits (Table 4, Fig. S2). This indicates that elevational variation exists in these traits too, but that their variation pattern cannot appropriately be described by sigmoid curves (Fig. S2).

Assessing the role of selection on the genetic cline

Given our estimates of cline width on both transects, the neutral diffusion model predicts that (in the absence of any selection pressure maintaining it), the contact between the two forms would have occurred very recently on both transects: 12 years ago (95% CI: 8 - 17)

based on width estimated on the northern transect (T1; Fig. 5) and 120 (95% CI: 71 - 147) years ago on T2 (Fig. 5). It strongly supports that these clines are likely to be maintained by selection as these dates are extremely recent.

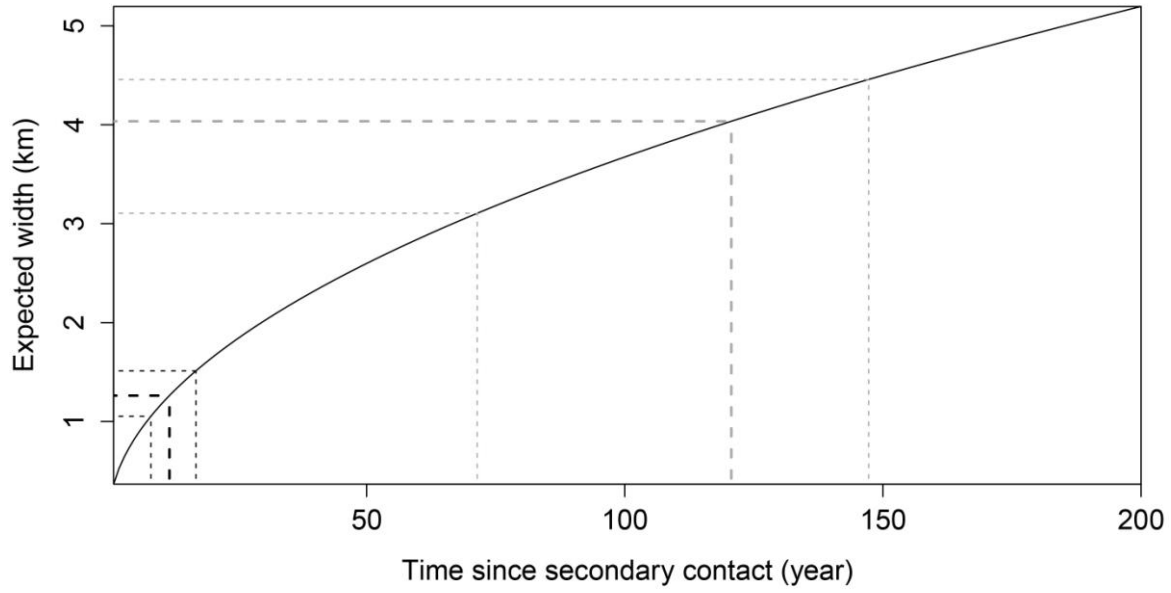


Figure 5: Expected width of the genetic cline according to time since secondary contact under neutral diffusion. Bold black dashed line represents expected time since secondary contact for T1, thin black dashed lines represent the 95% confidence intervals. Bold grey dashed line represents expected time since secondary contact for T2, thin grey ones represent the 95% confidence intervals.

Discussion

Hybrid zones are not rare among natural populations (e.g. more than 200 in birds reported by Price, 2008) and many of them have been proved to be powerful systems to investigate the relative contribution of selection and gene flow in maintaining population differentiation in contiguous areas (e.g. Szymura and Barton, 1986; Gay *et al.* 2008; Poelstra *et al.* 2014). Distinguishing between the effect of endogenous and exogenous selection is often difficult as they generally produce similar patterns (Kruuk *et al.* 1999). Moreover, these two kind of barriers often become coupled when environmental transitions such as ecotones exist (Bierne et al. 2011). Hybrid zones on regular environmental gradients have received

little attention whereas they could provide a good opportunity to investigate the role of exogenous and endogenous selection in the shaping of phenotypic variation and species barriers. As populations undergo spatially varying environmental selection pressures, traits under exogenous selection are expected to exhibit gradually varying phenotypes too whereas those involved in reproductive isolation are expected to present steeper changes in the hybrid zone vicinity. Here, we took advantage of a unique spatial configuration by studying a hybrid zone that occurs at a very small spatial scale but on a steep elevational gradient. We used cline analysis on phenotypic and genetic data on two transects crossing this hybrid zone to infer the role of selection on morphological, and plumage coloration clines the role of selection. Steep clines at neutral markers confirmed the existence of a barrier to gene flow and allowed to locate it accurately around 1400 m asl on both transects. Interestingly, some coloration components difficult to appreciate by the human eye also exhibited very narrow clines whose center position match the neutral genetic ones thus supporting the existence selection on plumage coloration and its possible implication in forms recognition and mate choice. On the contrary, morphological clines were wider than neutral ones and followed environmental variation. Thus, this study supports the idea that hybrid zones on environmental gradients may be very useful to distinguish traits under exogenous selection from those implied in reproductive isolation.

Spatial location and maintenance of the hybrid zone

Hybrid zones are likely to be trapped by local barriers (Barton and Hewitt 1985) and their centers are generally located in inhospitable areas where population density and habitat quality are low (Barton and Hewitt 1989). Here, we showed that the hybrid zone is located at the same elevation (ca. 1400-1500m) on both transects. This elevation matches anthropogenic boundaries. The elevation of 1400 m corresponds roughly to the lower limit of the native forest, below which natural habitats have been severely disturbed by anthropogenic activities (Strasberg *et al.* 2005). Despite the generalist habits of *Z. borbonicus*, this zone could constitute a slight ecological barrier, or at least a transition zone, triggering a decrease in bird density and consequently fixing the hybrid zone at this altitude. However, gathering new data to further test this idea will be necessary.

This hybrid zone occurs at an uncommonly small spatial scale for birds (*e.g.* McCormack & Smith 2008; see also references in Singhal & Moritz, 2012) as the observed

genetic cline width (on both transects) is very narrow (between 1 and 4.5 km). Two non-mutually exclusive hypotheses can be invoked to explain the abrupt clines observed on both transects. It could reflect the neutral expectation under a scenario of recent secondary contact between two divergent populations that still have not homogenized their genetic pool through hybridization (Haldane 1948, Endler 1977). However, using the diffusion equation, we showed that obtaining such genetic cline widths under neutral processes alone would have required an extremely recent secondary contact that appears to be unlikely in our system. As exemplified in another endemic passerine to Réunion (*Coracina newtoni*; Salmona *et al.* 2012), volcanic activities on Réunion could have played a role in demographic changes in birds and thus in the geographic isolation (*i.e.* allopatry) and subsequent contact of the different forms of *Z. borbonicus*. The last period of explosive activity of the Piton de la Fournaise volcano (between 3 000 and 10 000 years ago; Staudacher & Allègre, 1993; Mohamed-Abchir, 1996) is much older that the secondary contact date we obtained with the diffusion equation. It supports that barriers to gene flow exists between these forms and that selection (in conjunction with limited dispersal) is acting to maintain differentiation despite ongoing hybridization (Barton & Hewitt, 1989). This pattern suggests that differences between these forms are triggering pre-zygotic isolation and/or post-zygotic selection against hybrids such as genetic incompatibilities (Szymura & Barton, 1986; Barton & Hewitt, 1989). This also suggests that barriers to gene flow are probably not restricted to a small part of the genome. Moreover, it has already been shown with AFLP markers that divergence is probably genome wide.

Discrepancy between phenotypic traits: a role for endogenous and exogenous selection

The results of cline analysis provide evidence that belly colouration cline for brown birds shows coincidence and concordance with the molecular marker clines on both transects. Interestingly, the elevation at which the centers of these clines occur corresponds to the elevation at which the lowland brown forms meet the highland polymorphic form (*i.e.* from 1400 m upward grey birds appear to be common). The steep clines found in belly colouration suggest that its color might be involved in reproductive isolation by affecting assortative mating either as a by-product of previous isolation or as a consequence of secondary contact to avoid the production of unfit hybrids (i.e. endogenous constraints). If belly colouration is

involved in reproductive isolation with either direct or indirect selection acting on it, we expected that clines at its associated traits would have been narrower than neutral genetic ones. This may be due to heterogeneity among loci for cline width which is not captured by the use of a hybrid index (Gay *et al.* 2008). According to the physiological model of vision by Vorobyev and Osorio (1998), birds could use colour differences between lowland and highland populations for mate choice as these differences are visually detectable by birds (Cornuault et al., *in press* - Chapitre 1). The role of plumage colour in the maintenance of this hybrid zone would not be surprising as coloration signals are commonly used as criteria for mate choice in birds (Backer and Backer 1990; Roulin and Bize 2007, Price 2008; Moyle and Filardi 2009). Moreover, assortative mating has been recognized as a powerful agent of hybrid zone maintenance (Bridle *et al.*, 2006; Brodin & Hass, 2009; Taylor *et al.* 2012).

There was a striking discrepancy between clines for neutral molecular markers, belly colouration and body size; as none of the morphological traits measured exhibited the typical steep sigmoid shape found in genetic and colouration. Consistently with previous studies, tarsus length (a good proxy of structural body size; Senar and Pascual, 1997) exhibited a quasi-linear increase with elevation (Gill 1973; Milá et al. 2010; Bertrand *et al. in prep.* - Chapitre 2). This pattern of variation supports the role of exogenous selection through local adaption in shaping the response of this trait to gradually varying phenotypic fitness optimum (Cornuault et al. *in press* - Chapitre 1; Bertrand et al. - Chapitre 2). As expected from previous studies too, we did not find any clinal variation at bill related traits even after having controlled for allometric constraints. The increase in overall body size with altitude, is consistent with Bergmann's rule (Bergmann 1847, Zink and Remsen, 1986) which postulates that birds tend to be larger at higher altitudes. Potential selective agents on elevational gradient in this species may be either related to thermoregulation or resistance to starvation. They have already been discussed at length in elsewhere (see Cornuault *et al.*, *in press* - Chapitre 1; Bertrand, 2013). If we assume that the selective gradient is continuous and that body size is under exogenous selection associated to altitude, we may expect a linear increase in body size with elevation. At the contact zone, we also expect convergence of selective pressures linked to altitude and thus a convergence in phenotypic optima for both forms. Thus, the observed discrepancy between clines for molecular markers and body size, together with the positive relationship between body size and altitude, support this hypothesis of phenotypic 'convergence' at the hybrid zone centre due to strong environmental selection. Introgression may have facilitated the transfer of advantageous alleles for adaptation to

altitude between the different forms (*i.e.* adaptive introgression; Parsons *et al.* 1993; Martin *et al.* 2006). As recently highlighted by Abbott and Brennan (2014), hybrid zones on altitudinal gradients are thus promising systems to investigate the maintenance of divergence in the face of gene flow and to understand the role of adaptation and intrinsic barriers in hybrid zones.

Conclusions

Hybridization is common in birds (Grant and Grant 1992), yet the placement of this hybrid zone, the small spatial scale under consideration, and the discrepancy in the shapes of the different clines, suggests the role of various selective factors acting simultaneously to maintain this interesting hybrid zone. Strikingly, we demonstrated that variation for belly plumage coloration differed from body size clines and was concordant with variation at neutral markers. This suggests that differentiation in body size could be adaptive and that linkage disequilibrium might be sufficiently low to allow the different traits to vary independently. On the other side, coloration might be involved in reproductive isolation between the different forms. These findings also support the idea that some phenotypic traits in the hybrid zone may vary gradually leading to the formation of cryptic hybrid zones undetectable at the phenotypic level, which suggests that hybrid zones like the one reported here might have been previously undetected in other organisms and may be more common in nature than presently thought. Future studies on this system will allow us to better understand the interplay between gene flow, intrinsic and extrinsic selection forces in hybrid zones, and what they can reveal about the process of speciation. Genomic data will be particularly suitable, as they will allow to investigate to what extent introgression rate is variable within the genome, to identify specific genomic regions under selection and to specify the role of adaptation to the environmental gradient in such systems. Thus, hybrid zone on environmental gradients provide promising frameworks to study speciation and to understand patterns of differential introgression rates across genomes.

References

Abbott, R.J. and Brennan A.C.. (2014) Altitudinal gradients, plant hybrid zones and evolutionary novelty. *Philosphical Transactions of the Royal Society*, **369**: 20130346.

Balbontín J., Møller A.P., Hermosell I.G., Marzal A., Reviriego M., de Lope F. (2012) Lifetime individual plasticity in body condition of a migratory bird. *Biological Journal of the Linnean Society*, **105**: 420-434.

Barton, N.H. and Hewitt, G.M. (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**: 113-148.

Barton, N.H. and Hewitt, G.M. (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**: 497-503.

Barton, N.H. and Gale, K.S. Genetic analysis of hybrid zones. pp. 13-45 in Hybrid Zones and the Evolutionary Process, pp 13-45. Oxford: Oxford University Press; 1993.

Bears, H., Drever, M.C. and Martin, K. (2008) Comparative morphology of dark-eyed juncos *Junco hyemalis* breeding at two elevations: a common aviary experiment. *Journal Avian Biology*, **39:** 152-162.

Bergmann, C. (1847) Ueber die verhältnisse der wärmeökonomie der thiere zuihrer grösse. *Gottinger studien*, **3**: 595-708.

Bermond, G., Ciosi, M., Lombaert, E., Blin, A., Boriani, M., Furlan, L., Toepfer, S. and Guillemaud, T. (2012) Secondary contact and admixture between independently invading populations of the western corn rootworm, *Diabrotica virgifera virgifera* in Europe. *Plos One*, **7**: e50129.

Bertrand, J.A.M., García-Jiménez, R., Bourgeois, Y., Duval, T., Heeb, P., Thébaud, C and Milá, B. (*2012*) Isolation and characterization of twelve polymorphic microsatellite loci for investigating diversification of a "great speciator": the Mascarene grey white-eye (*Aves:Zosterops borbonicus*). *Conservation Genetic Resources*, **4**: 323-326.

Bertrand, J.A.M., Bourgeois, Y.X.C, Delahaie, B., Duval, T., García-Jiménez, R., Cornuault, J., Heeb, P., Milá, B., Pujol, B. and Thébaud, C. (2014) Extremely reduced dispersal and gene flow in an island bird. *Heredity*, **112**: 190-196.

Bertrand, J.A.M. (2013) Causes de la différenciation génétique à une très petite échelle spatiale chez un oiseau insulaire (*Zosterops borbonicus*). D. Phil. Thesis, University Paul Sabatier.

Bertrand, J.A.M., Delahaie, B., Bourgeois, Y.X.C., Duval, T., Garcia-Jimenez, R., Cornuault, J., Pujol, B., Mila, B., Thébaud, C. Population structure along an elevational gradient reveals cryptic variation between parapatric colour forms in an island bird. (*in prep.)*

Bierne, N., Welch, J., Loire, E., Bonhomme, F. and David, P. (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**: 2044-2072.

Blanchard, F. (2000) Guide des milieux naturels : La Réunion – Maurice – Rodrigues. Les éditions Eugen Ulmer, Paris, France.

Bourgeois, Y.X.C. (2013) Génétique évolutive d'un cas extrême de polymorphisme de la coloration du plumage chez un oiseau insulaire, *Zosterops borbonicus* (Zosteropideae). D. Phil. Thesis, University Paul Sabatier.

Brennan, A.C., Bridle, J.R., Wang, A.-L., Hiscock, S.J. and Abbott, R.J. (2009) Adaptation and selection in the *Senecio* (*Asteraceae*) hybrid zone on Mount Etna, Sicily. *New Phytologist*, **183**: 702-717.

Bridle, J.R., Saldamando, C.I., Koning, W. and Butlin, R.K. (2006) Assortative preferences and discrimination by females against hybrid male song in the grasshoppers *Chorthippus brunneus* and *Chorthippus Jacobsi* (*Orthoptera: Acrididae*). *Journal of Evolutionary Biology*, **19**: 1248-1256.

Brodin, A. and Haas, F. (2009) Hybrid zone maintenance by non-adaptive mate choice. *Evolutionary Ecology*, **23**: 17-29.

Brumfield, R.T., Jernigan, R.W, McDonald, D.B. and Braun, M.J. (2001) Evolutionary implications of divergent clines in an avian (Manacus: Aves) hybrid zone. *Evolution*, **55**: 2070–2087.

Burnham, K.P. and Anderson, D.R. Model selection and inference, a practical information-theoretic approach. Springer-Verlag; 2002.

Caro, L.M., Caycedo-Rosales, P.C., Bowie, R.C.K., Slabbekoorn, H. and Cadena, C.D. (2013) Ecological speciation along an elevational gradient in a tropical passerine bird? *Journal of Evolutionary Biology*, **26**: 357-374.

Cheviron, Z.A. and Brumfield, R.T. (2009) Migration-selection balance and local adaptation of mitochondrial haplotypes in rufous-collared sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution*: **63**: 1593-1605.

Cornuault, J., Delahaie, B., Bertrand, J.A.M., Bourgeois, Y.X.C., Milá, B., Thébaud, C. and Heeb, P. (*in press*) Morphological and plumage colour variation in the Réunion grey white-eye (Aves: *Zosterops* borbonicus): assessing the role of selection. *Biological Journal of the Linnean Society*.

Cuthill, I., Bennett, A., Partridge, J. and Maier, E. (1999) Plumage reflectance and the objective assessment of avian sexual dichromatism. *The American Naturalist*, **153**: 183-200.

Delmore, K.E., Brenneman, R.A., Lei, R., Bailey, C.A., Brelsford, A., Louis, E.E. Jr, Johnson, S.E. Clinal variation in a brown lemur (Eulemur spp.) hybrid zone: combining morphological, genetic and climatic data to examine stability. *Journal of Evolutionary Biology*, **26**: 1677-1690.

Dubay, S.G. and Witt C.C. (2014) Differential high-altitude adaptation and restricted gene flow across a mid-elevation hybrid zone in Andean tit-tyrant flycatchers. *Molecular Ecology*, **23**: 3551-3565.

Edelaar, P., Alonsi, D., Lagerveld, S., Senar, J.C. and Björklund, M. (2012) Population differentiation and restricted gene flow in Spanish crossbills: not isolation-by-distance but isolation-by-ecology. *Journal of Evolutionary Biology*, **25**: 417-429.

Endler, J.A. Geographic Variation, Speciation, and Clines. Princeton, New Jersey: Princeton University Press; 1977.

Endler, J.A. (1990) On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society*, **41**: 315-352.

Endler, J. A. and P. W. J. Mielke (2005) Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, **86**: 405–431.

Evanno G., Regnaut S. and Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**: 2611-2620.

Falush, D., Stephens, M. and Pritchard, J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**: 1567-1587

Gay, L., Crochet, P.-A., Bell, D.A. and Lenormand, T. (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution*, **62**: 2789-2806.

Gill, F.B. (1973) Intra-island variation in the mascarene white-eye *Zosterops borbonica*. *Ornithological Monographs*, **12**: 1-66.

Goldsmith, T. H. (1990) Optimization, constraint, and history in the evolution of eyes. *Quarterly Review of Biology*, **65**: 281–322.

Goudet, J. (2001) FSTAT, a program to estimate and test gene diversities and fixation indices, Version 2.9.3. http://www.unil.ch/izea/softwares/fstat.html

Grant, P.R. and Grant, B.R. (1992) Hybridization of bird species. Science, **256**: 193-197.

Greenberg, R., Danner, R., Olsen, B. and Luther, D. (2012) High summer temperature explains bill size variation in salt marsh sparrows. *Ecography*, **35**: 146-152.

Haldane, J.B.S. (1948) The theory of a cline. *Journal of Genetics*, **48**: 277-284.

Hewitt, G.M. (1988) Hybrid zones - Natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, **3**: 158-167.

Hubisz, M.J., Falush, D., Stephens, M. and Pritchard, J.K. (2009) Inferring weak population structure with the assistance of sample group information.*Molecular Ecology Resources*, **9**: 1322-1332.

Jakobsson, M. and Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**: 1801-1806.

Key, K. H. L. 1968. The concept of stasipatric speciation. Syst. Zool. 17: 14–22.

Kirkpatrick, M. and Barton, N.H. (1997) Evolution of a species' range. *American Naturalist*, **150:** 1-23.

Kruuk, L.E.B., Baird, S.J.E, Gale, K.S. and Barton, N.H. (1999) A Comparison of Multilocus Clines Maintained by Environmental Adaptation or by Selection Against Hybrids. *Genetics*, **153**: 1959-1971.

Maia, R; Eliason, CM; Bitton, P-P, Doucet, SM & Shawkey, MD. (2013) pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution*, **4**: 906-913.

Martin, N.H., Bouck, A.C. and Arnold, M.L. (2006) Detecting adaptive trait introgression between *Iris fulva* and *I. brevicaulis* in highly selective field conditions. Genetics, **172**: 2481-2489.

May, R.M., Endler, J.A. and McMurtie, R.E. (1975) Gene frequency clines in the presence of selection opposed by gene flow. *American Naturalist*, **109**: 659-676.

McCormack, J.E. and Smith, T.B. (2008) Niche expansion leads to small-scale adaptive divergence along an elevation gradient in a medium-sized passerine bird. *Proceedings of the Royal Society B: Biological Sciences*, **275**: 2155–2164.

Milá, B., R. K. Wayne, P. S. Fitze and T. B. Smith. (2009) Divergence with gene flow and fine-scale phylogeographic structure in the wedge-billed woodcreeper *Glyphorynchus spirurus*, a Neotropical rainforest bird. *Molecular Ecology*, **18**: 2979-2995.

Milá, B., Warren, B.H., Heeb, P. and Thébaud, C. (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**: 158.

Mohamed-Abchir, A. (1996) Les cendres de bellecombe (cb) : un evenement explosif majeur dans le passe récent du piton de la fournaise, île de la Réunion. D. Phil. Thesis, Paris.

Moore, W.S. (1977) An evaluation of narrow hybrid zones in vertebrates. *The quarterly review of biology*, **52**: 263-277.

Parsons, T. J., Olson, S. L. and Braun, M. J. (1993) Unidirectional spread of secondary sexual plumage traits across an avian hy- brid zone. *Science*, **260**: 1643-1646.

Poelstra, J.W., Vijay, N., Bossu, C.M., Lantz, H., Ryll, B., Müller, I, Baglione, V., Unneberg, P, Wikelski, M., Grabherr, G. and Wolf, J.B.W. (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, **344**: 1410-1414.

Price, T. 2008. Speciation in birds. Boulder, Colorado: Roberts and Co.

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945-959.

Pyle P: Identification Guide to North American Birds. Part I. Bolinas, California: Slate Creek Press; 1997.

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Raufaste N, Orth A, Belkhir K, Senet D, Smadja C, Baird SJE, Bonhomme F, Dod B, Boursot P. (2005) Inferences of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society* **84:** 593-616.

Roulin, A. and Bize, P. (2007) Sewual selection in genetic colour-polymorphic species: a review of experimental studies and perspectives. *Journal of Ethology*, **25**: 99-105.

Salmona J, Salamolard M, Fouillot D, Ghestemme T, Larose J, Centon, J.-F., Sousa, V., Dawson, D.A., Thébaud, C. and Chikhi, L. (2012) Signature of a Pre-Human Population Decline in the Critically Endangered Reunion Island Endemic Forest Bird *Coracina newtoni*. *PLoS ONE*, **7**: e43524.

Schluter, D. (2001) Ecology and the origin of species. *Trends in Ecology & Evolution*, **16**: 372-380.

Senar, J.E. and Pascual, J. (1997) Keel and tarsus length may provide a good predictor of avian body size. *Ardea*, **85**: 269-274.

Singhal S, Moritz C. (2012) Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. *Evolution* **66:** 1474-1489.

Slatkin, M. (1973) Gene flow and selection in a cline. *Genetics*, **75**: 733-756.

Smith, K.L., Hale, J.M., Kearney, M.R., Austin, J.J., Melville, J. (2013). Molecular patterns of introgression in a classic hybrid zone between the Australian tree frogs, *Litoria ewingii* and *L. paraewingi*: Evidence of a tension zone. *Molecular Ecology*, **22**: 1869-1883.

Staudacher, T. and Allegre, C.J. (1993) Ages of the second caldera of Piton de la Fournaise volcano (Reunion) determined by cosmic ray produced 3He and 21Ne. *Earth Planetary Science Letters*, **119**: 395-404.

Stoddard, M.C. and Prum, R.O. (2008) Evolution of avian plumage colour in a tetrahedral colour space: A phylogenetic analysis of new world buntings. *American Naturalist*, **171**: 755-776.

Strasberg, D., Rouget, M. Richardson, D.M., Baret, S., Dupont, J. and Cowling, R.M. (2005). An assessment of habitat diversity, transformation and threats to biodiversity on reunion island (Mascarene islands, indian ocean) as a basis for conservation planning. *Biodiversity & Conservation*, **14**: 3015-3032.

Szymura, J.M. and Barton, N.H. (1986) Genetic Analysis of a Hybrid Zone Between the Fire-Bellied Toads, Bombina bombina and B. variegata, Near Cracow in Southern Poland. *Evolution*, 40: 1141-1159.

Uy, J.A., Moyle, R.G., and Filardi, C.E. (2009) Plumage and song differences mediate species recognition between incipient flycatcher species of the Solomon Islands. *Evolution*, **63**: 153-164.

van Oosterhout C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**: 535-538.

Vorobyev, M. and Osorio, D. (1998) Receptor noise as a determinant of colour thresholds. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **265**: 351-358.

Zink, R.M. and Remsen, J.V. (1986) Evolutionary patterns and processes of geographic variation in birds. *Current Ornithology*, **4**: 1-69.

Supplementary information

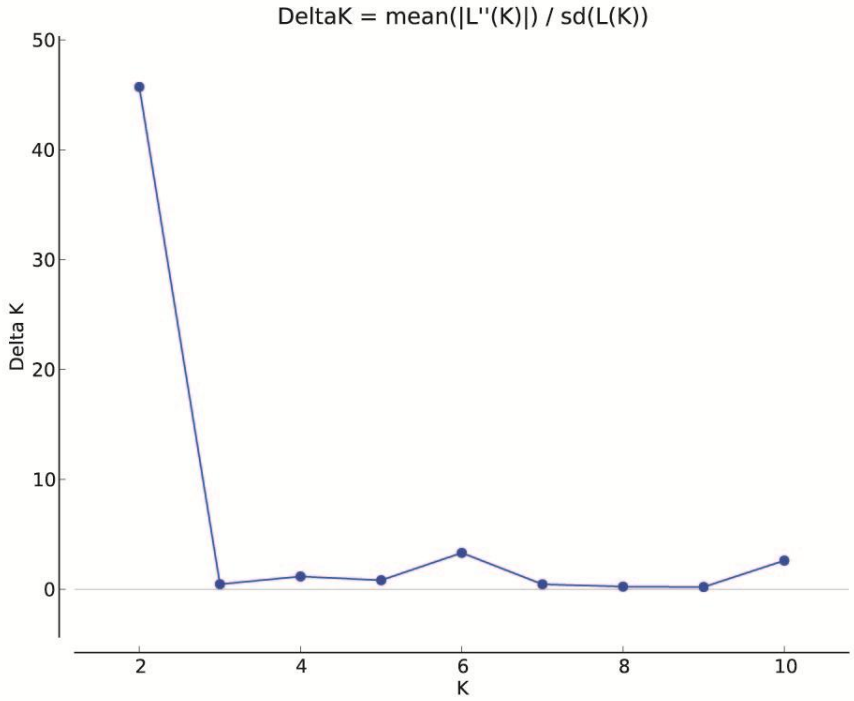


Figure S1: delta K plot from the STRUCTURE analysis.

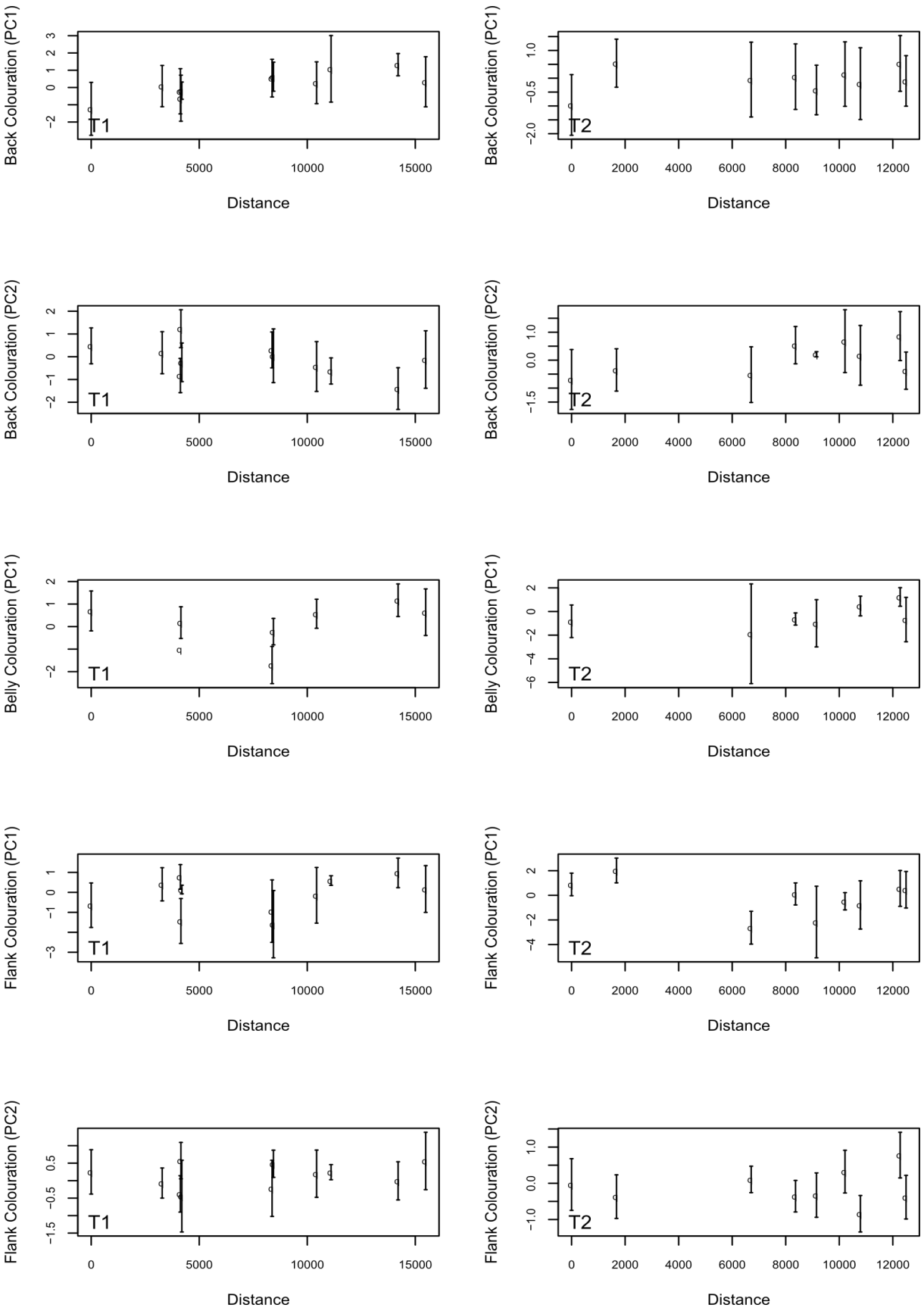


Figure S2: Morphological variation along the two transects (T1 on the left, T2 on the right).

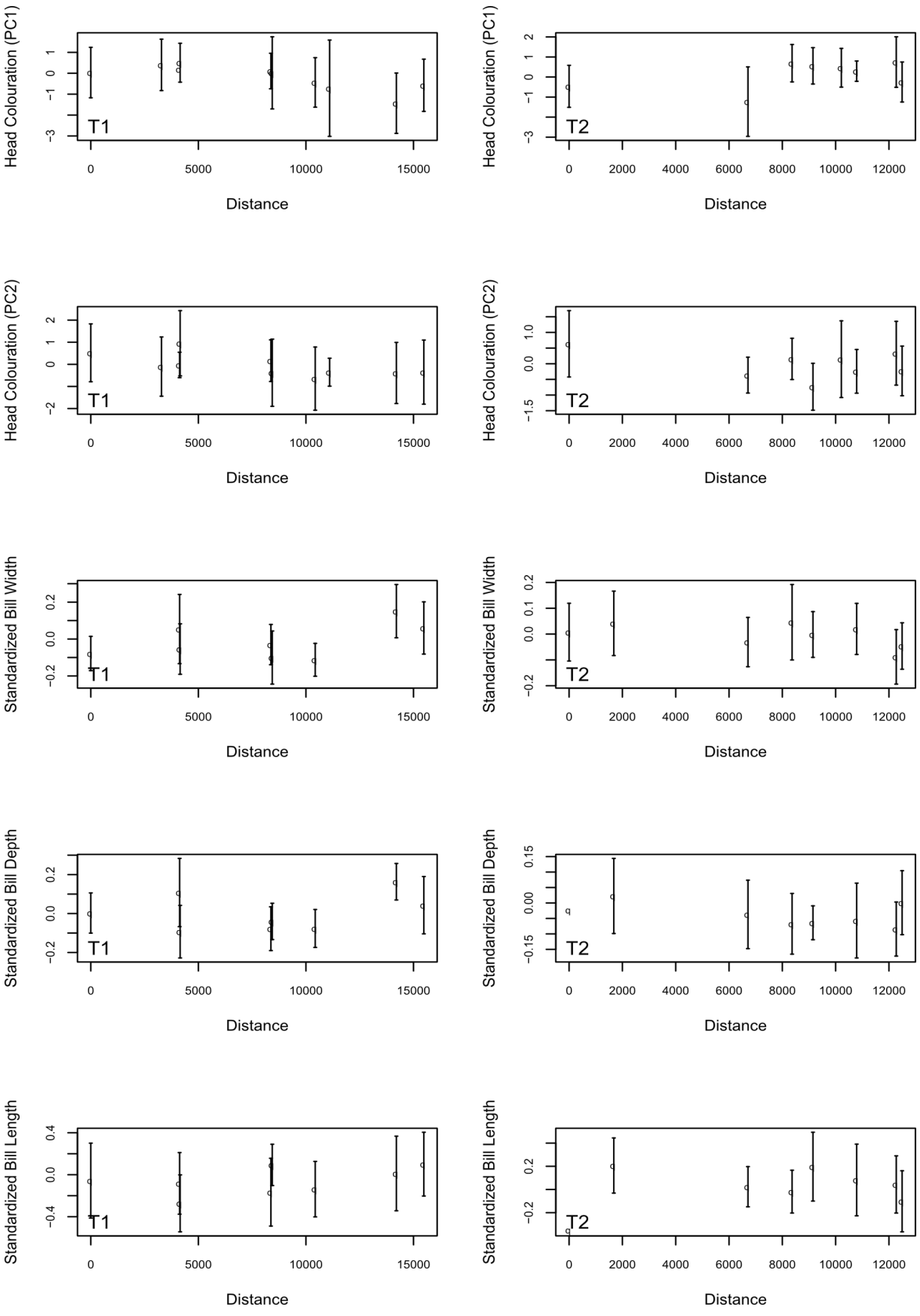


Figure S2 (continued)

Chapitre 4

Maintien des zones hybrides de basse altitude

Présentation du chapitre

Dans ce dernier chapitre, nous nous sommes intéressés aux trois zones hybrides séparant les formes de couleur de basse altitude. Ces différentes zones hybrides ne semblent pas être placées au niveau de transitions environnementales. Elles semblent, par contre, être localisées dans des habitats homogènes au niveau de discontinuités physiques : rivières et complexe de coulée de lave. Afin de comprendre comment ces différentes zones se maintiennent depuis plus de 50 ans, nous avons examiné les variations environnementales, génétiques et phénotypiques au travers de ces différentes zones hybrides.

Marked phenotypic differentiation across three hybrid zones in an island bird despite a lack of neutral genetic structure in microsatellite markers

In preparation

Boris Delahaie[1], Charline Masson[1], Joris AM Bertrand[1], Yann XC Bourgeois[1], Josselin Cornuault[1], Thomas Duval[2], Juli Broggi[3], Borja Milá[4], & Christophe Thébaud[1]

1. Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174 Centre National de la Recherche Scientifique (CNRS) - Université Paul Sabatier - Ecole Nationale de Formation Agronomique (ENFA), 118 Route de Narbonne, F-31062 Toulouse, France.

2. *BP 438, 98822 Poindimié, Nouvelle-Calédonie*

3. Estación Biológica Doñana, CSIC, 41092 Sevilla, Spain.

4 National Museum of Natural Sciences, Spanish Research Council (CSIC), Madrid 28006, Spain.

Correspondance:

Boris Delahaie

Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174, Université Paul Sabatier, Toulouse 3, 118 Route de Narbonne, F-31062 Toulouse Cedex 9, France.

+33 (0)5.61.55.67.23

borisdelahaie@gmail.com

Abstract

The Réunion grey white-eye (*Zosterops borbonicus*), a small passerine endemic to the island of Réunion (Mascarene archipelago), constitutes an extraordinary case of colour variation within a bird species, with four colour forms occupying distinct geographic ranges on this small island (2512km²). Its three lowland colour forms meet at narrow physical discontinuities where they form hybrid zones. To investigate how these hybrid zones are maintained, we characterized ecological, morphological and colour variation across them and examined genetic structure at 11 microsatellite loci. Strikingly, while these hybrid zones stand out among the narrowest in birds based on phenotypes, we were not able to detect any neutral genetic structure associated to colour forms based on a set of representative microsatellites loci. Moreover, hybrid zone centres did not correspond to any abrupt change in environmental conditions. Thus, we suggest that the differentiation between these forms is dominated by few genes located in a narrow part of the genome and that mate choice is probably involved in the maintenance of this zone.

Keywords: cline, *Zosterops borbonicus*, plumage colour, Réunion, hybrid zones

Introduction

Hybrid zones are geographic areas where parapatric taxa meet, reproduce and produce hybrids (Barton & Hewitt 1985). These areas provide ideal conditions to study the development of barriers to gene flow and the factors responsible for divergence. For this reason, they are often considered as 'windows' into evolutionary processes (Harrison 1990). Their maintenance is generally explained by the countervailing effects of gene flow and selection against hybrids (Barton & Hewitt 1985). Gene flow tends to homogenize populations through recombination whereas exogenous divergent selection or intrinsic hybrid incompatibilities (*i.e.* endogenous selection) reduces the introgression. In nature, hybrid zones are highly variable in their shape and in their position relative to ecological variation. At the phenotypic level, a continuum of situations exists from wide zones to sharp transition zones. Strikingly, it is very common that phenotypic characters and genetic markers present discordance in terms of width of the transition zone (Harrison & Larson 2014). Three main cases may be observed: i) overall, phenotypic traits and genetic markers show similar clinal pattern (Gay *et al.* 2008; Irwin *et al.* 2009); ii) some phenotypic traits show an abrupt transition at the hybrid zone whereas there are virtually no genetic breaks (Nadeau *et al.* 2013; Poelstra *et al.* 2014) iii) cryptic hybrid zone, where distinct genetic lineages do not show any differences in phenotypes (Singhal & Moritz 2012). This kind of differences is generally interpreted as a consequence of differential introgression rates across genomes.

It is now well recognized that species barriers are semi-permeable and thus gene flow does not affect the entire genome homogeneously (Rieseberg *et al.* 1999; Wu 2001; Harrison & Larson 2014). Introgression at neutral markers is influenced by several factors. First, the genetic architecture (*i.e.* the number, effect size and distribution across the genome of genomic regions) of the hybrid unfitness greatly influence the diffusion of neutral parts of the genome. When loci involved in reproductive isolation are numerous and widespread across the genome, a situation which may be common in nature (*e.g.* Larson *et al.* 2014; Baldassarre *et al.* 2014), nearby linked neutral loci are expected to diffuse very slowly across the hybrid zone (Barton & Hewitt 1985). In this case, we expect that phenotypes and neutral genetic markers will show similar steep transition patterns. In contrast, if only a few loci are involved in reproductive isolation or if they are located on a relatively small region of the genome, introgression in the rest of the genome is expected to be high relative to the small region under selection. In this case, it must lead to strong discrepancy between phenotypes and neutral genetic markers with no difference at most part of the neutral genome. Besides, the

proportion of assortative mating in a hybrid zone has also a strong impact on the diffusion of genetic material between the two taxa. In this respect, it has been highlighted that bimodal hybrid zones (*i.e.* 'zones where hybrids are rare and parental forms predominate') are all affected by quite strong pre-zygotic barriers (Jiggins & Mallet 2000). Many studies have documented genome-wide barriers to reproduction in hybrid zones leading gene flow to be globally restricted over the genome, but fewer have reported that the marked differences in phenotypes in hybrid zones may be maintained in the face of gene flow by divergence at a very small portion of the genome (Harrison & Larson 2014). These situations must be interesting as they probably represent early stages of speciation and may help to identify interesting speciation phenotypes or genes at the origin of reproductive isolation (Shaw & Mullen 2011; Harrison & Larson 2014). Thus, documenting variability in patterns of differentiation at different traits and markers in hybrid zones between closely related taxa provides much insight into how species limits evolve and are maintained. This work can be done either by using genome wide markers (*e.g.* Larson *et al.* 2014; Carneiro *et al.* 2014) or by comparing phenotypes to a representative set of neutral markers.

In this study, we focus on the heterogeneity of clinal variation patterns between neutral genetic markers and phenotypic traits in an avian hybrid zone. We chose *Zosterops borbonicus* because this species complex, endemic to the small island of Réunion (2512 km²) provides a unique set of hybrid zones between closely related but phenotypically distinct forms. Indeed, this species presents four distinct colour forms among which three are restricted to low elevations and occupy discrete geographic regions parapatrically distributed all around the island. Each colour form range is delimited by geographic barriers that correspond to rivers or dry lava flows. At these different contact zones, each pair of colour forms meet and birds with intermediate phenotypes are present at the core of each contact zone (Gill 1973; Milá *et al.* 2010). Frank Gill noticed that these zones occur at physical barriers with no conspicuous changes in vegetation types and proposed that natural selection was low and the physical barriers were restricting gene flow between the different forms (Gill 1973) although they are probably very narrow compared to the potential flying ability of the species. In this study, we sampled individuals at localities distributed around the island which correspond to three different transects cutting through the different hybrid zones. Using microsatellites, phenotypic and environmental data our objectives were three fold. First, we wanted to confirm with an objective dataset that the observed range limits between the different colour forms does not correspond to ecotones between habitat types or abrupt

transition in climate. In this respect, we expect the phenotypes differences not to be maintained by exogenous selection if the contact zones are not located on ecotones. Second, we evaluate the degree of genetic differentiation among each pair of colour forms and look for barrier to gene flow using eleven microsatellites loci specifically developed for the species. Third, we characterize the transition at different phenotypic traits using cline analysis in order to evaluate whether the hybrid phenotypes are restricted to the core of the hybrid zones.

Material and Methods

Study system and sampling

Lowland colour forms of *Z. borbonicus* mainly differ in head colouration. These three forms include a brown-headed brown form (also called 'lowland brown-headed brown - LBHB') in the West, a grey-headed brown form (GHB) in the North and the East of the island and a brown-naped brown form (BNB) in the South (Fig. 1 & 2). This situation has been stable for at least 50 years, as the current distribution pattern matches exactly the pattern described by Frank Gill in the 1960s (Milá *et al.* 2010). Sampling was done on three distinct transects cutting through the different contact zones between these three colour forms (Fig. 1). The first contact zone is located on the North-West side of the island (LBHB-GHB hybrid zone) and centered on the Galets River (Fig. 1). The second contact zone is located in the South-East of the island at the St-Etienne River (~ 500 m wide; LBHB-BNB hybrid zone; Fig. 1) and the third contact zone is located in the North of the 'Grand Brûlé' lava complex in the South-East of the island (BNB-GHB hybrid zone; Fig. 1). Both rivers are composed of a wide (~ 500m), dry and barren river bed, with a narrow river a few meters wide. There is very little vegetation and thus few favorable habitats in these river beds. In the Grand Brûlé (BNB-GHB hybrid zone), the vegetation is abundant except on the most recent lava flows. The end points of each transect represents parental populations corresponding to 'pure' phenotypes and are shared between the different transects (Fig. 1). The spatial configuration of this system precludes the establishment of straight sampling transects for each contact zone due to the circular shape of Réunion island, so we decided to use the 200-m elevation isocline as a transect for each contact zone.

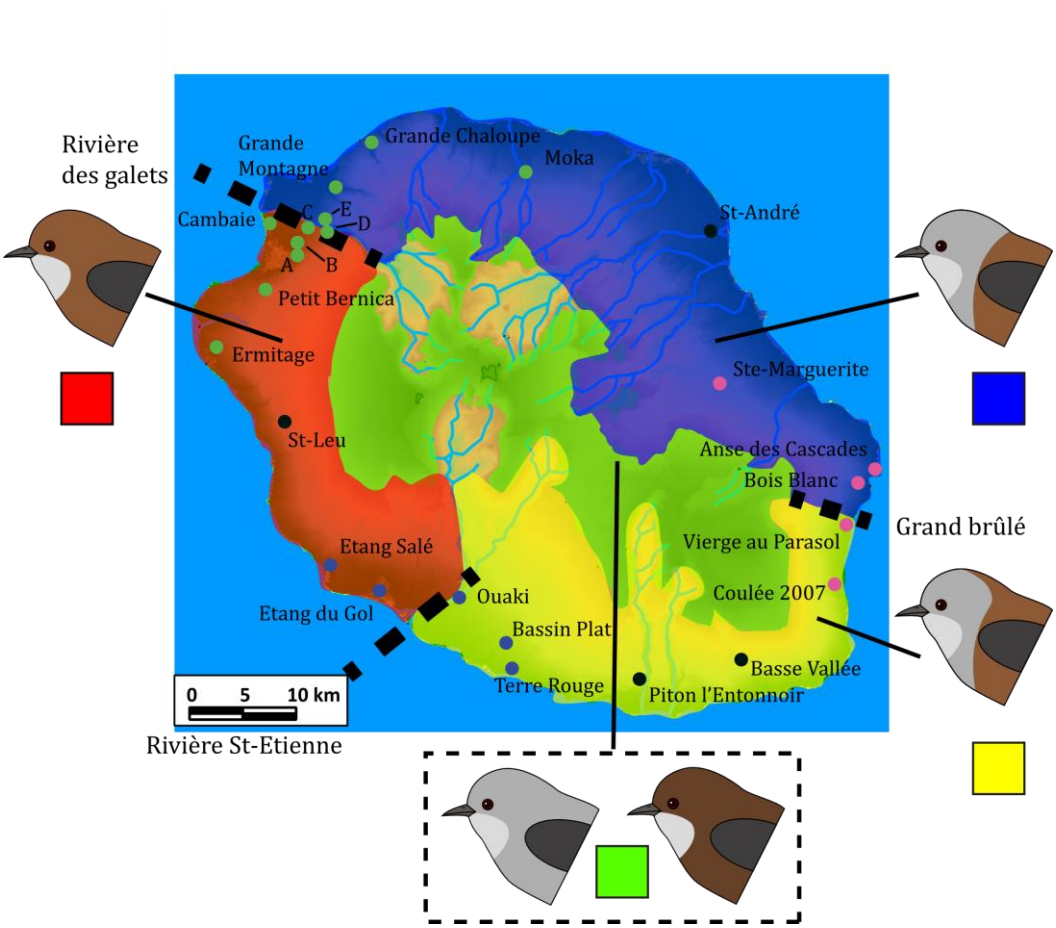


Figure 1: Map showing the geographical position of the 25 sampling localities. The different colour layers represent the range of the five Réunion grey white-eye plumage colour variants: LBHB in red, GHB in blue and BNB in yellow. The different colour used for the localities show the different transects: green dots correspond to the LBHB - GHB transect; pink ones to the GHB - BNB transect and blue ones to the BNB - LBHB transect. Black dots correspond to the end of each transect and belong to two adjacent transects. A: Moulin à eau; B: Déboulé; C: Lycée agricole; D: Sans Souci; E: Rivière des Galets.

Figure 2: Pictures of the different colour forms. From left to right: LBHB, GHB and BNB.

We chose this isocline as it corresponds roughly to the mean elevation of our sampling localities (Table 1). The numerous gullies characterizing the island were not taken into account as they are unlikely to act as barriers to dispersal for this species. We thus decided to cut them as the crow flies. Each transect thus corresponds to a line roughly parallel to the coastline. Sampling localities were then projected on each transect and the new coordinates were defined as the closest vertex of the isocline. Sampling stations were all comprised within the 4.5 km from the different transect lines. Distance along the transect was calculated from the following localities at each transect: 'St-Leu' for the LBHB - GHB one, 'St-André' for the GHB - BNB one and 'Gite Basse Vallée' for the BNB - LBHB one. Birds used in this study were sampled using mist nests. Three ringers performed all the bird manipulations: Borja Milá (BM), Thomas Duval (TD) and Juli Broggi (JB). Authorizations are described elsewhere (Milá *et al.* 2010). All birds were marked permanently with a uniquely numbered aluminum ring.

Morphological measurements and blood sampling

Individuals caught were aged as 'juvenile' or 'adult', whenever possible, using plumage characteristics, eye and gape colour, moult pattern and the degree of skull ossification, following criteria in (Pyle 1997). Morphological measurements were taken for each mist-netted bird at seven morphological traits : body mass, tarsus length, wing length, tail length, bill length, depth and width (Table 1). Dial calipers of 0.1 mm precision were used to measure tarsus length (from the proximal end of the tarsometatarsus to the first undivided scute), tail length (from the uropygial gland to the tip of the longest rectrix), bill length (from the anterior end of the nares to the tip of the upper mandible), and bill width and depth (both measured at the anterior end of the nares). Wing length was measured using a wing ruler of 0.5 mm precision. To characterize plumage colour, we collected feather samples on each bird captured (see Table 1 for details about sample sizes). At least ten feathers were collected from each of the following body regions on each bird: head, back, flank and belly (see Table 1 for details about sample sizes). Blood samples were taken and stored in Queen’s lysis buffer (Seutin *et al.* 1991). We extracted total genomic DNA from these blood samples using a DNeasy Blood & Tissue Kit (Qiagen Inc., Venlo, NL) in accordance with the manufacturer’s recommended protocol.

Table 1: Summary of sampling station characteristics and microsatellite diversity for the different loci. Habitat corresponds to the vegetation type, either lowland semi-dry forest ('Dry') or lowland rainforest ('Rain'). *A* corresponds to the mean number of allele per site; A_R to the mean allelic richness per site. Observed and expected heterozygosity (H_o and H_E), and inbreeding coefficient (F_{IS}). F_{IS} credible intervals which do not overlap with 0 are indicated in bold. Distance corresponds to the distance along the 200m isocline to the first locality of each transect (St Leu for the LBHB – GHB transect, St André for the GHB – BNB transect and Basse Vallée for the BNB – LBHB one).

Transect LBHB - GHB											
Sampling site	Latitude	Longitude	Distance	Colour Form	Habitat	n_{gen}	A	A_R	H_O	H_E	F_{IS}
St Leu	-21,14	55,30	0	LBHB	Dry	21	7,91	4,98	0,78	0,78	0
Ermitage	-21,07	55,23	9195	LBHB	Dry	16	7,18	4,86	0,77	0,78	0,01
Petit Bernica	-21,03	55,28	16555	LBHB	Dry	10	6,55	4,65	0,82	0,72	**-0,14**
Moulin à Eau	-21,00	55,31	21384	LBHB	Dry	5	-	-	-	-	-
Déboulé	-20,98	55,31	22743	LBHB	Dry	13	7,36	4,67	0,76	0,73	-0,04
Cambaie	-20,97	55,28	22743	LBHB	Dry	15	6,18	4,27	0,76	0,71	-0,07
Lycée agricole	-20,97	55,32	24159	LBHB	Dry	4	-	-	-	-	-
Sans Souci	-20,98	55,34	25968	LBHB	Dry	12	5,91	4,22	0,79	0,69	**-0,15**
Rivière des Galets	-20,97	55,33	26629	GHB	Dry	10	6,27	4,59	0,82	0,73	**-0,12**
la Grande Montagne	-20,94	55,34	29934	GHB	Dry	3	-	-	-	-	-
la Grande Chaloupe	-20,90	55,38	35434	GHB	Dry	7	4,73	3,93	0,8	0,68	**-0,17**
Moka	-20,93	55,52	53222	GHB	Rain	21	8,36	4,86	0,8	0,77	-0,04
St-André	-20,98	55,68	74676	GHB	Rain	9	7,18	4,97	0,74	0,79	0,07
Transect GHB - BNB											
Sampling site											
St André	-20,98	55,68	0	GHB	Rain	9	7,18	4,97	0,74	0,79	0,07
Sentier Ste-Marguerite	-21,11	55,69	10104	GHB	Rain	12	7,36	5,05	0,81	0,77	-0,05
Anse des Cascades	-21,18	55,83	26660	GHB	Rain	12	6,18	4,52	0,76	0,75	-0,01
Bois Blanc	-21,19	55,81	28089	GHB	Rain	12	7,36	5,05	0,83	0,78	-0,06
Vierge au Parasol	-21,22	55,81	30893	Hybrid	Rain	12	5,91	4,41	0,86	0,74	**-0,17**
Coulée 2007	-21,28	55,79	38107	BNB	Rain	22	6,91	4,42	0,79	0,74	**-0,07**
Basse Vallée	-21,34	55,71	52680	BNB	Rain	16	7,64	4,92	0,76	0,75	0
Piton de l'Entonnoir	-21,36	55,61	63904	BNB	Dry	7	6,18	4,78	0,74	0,76	0,03
Transect BNB - LBHB											
Sampling site											
Basse Vallée	-21,34	55,71	0	BNB	Rain	16	7,64	4,92	0,76	0,75	0
Piton de l'Entonnoir	-21,36	55,61	11224	BNB	Dry	7	6,18	4,78	0,74	0,76	0,03
Terre Rouge	-21,35	55,50	23559	BNB	Dry	4	-	-	-	-	-
Bassin Plat	-21,33	55,49	24953	BNB	Dry	6	4,18	3,4	-	0,68	0,18
Ouaki	-21,29	55,45	31321	Hybrid	Dry	40	8	4,61	0,78	0,76	-0,02
Etang du Gol	-21,28	55,38	40075	LBHB	Dry	34	6,64	4,39	0,75	0,72	-0,04
Etang Salé	-21,26	55,34	44744	LBHB	Dry	32	9,73	5,2	0,77	0,77	0
St Leu	-21,14	55,30	58824	LBHB	Dry	21	7,91	4,98	0,78	0,78	0

Molecular markers and within-population genetic variation

We genotyped 324 (see Table 1 for details about sample sizes) individuals at 11 microsatellite loci that were specifically developed for *Zosterops borbonicus* (Z1, Z2, Z3, Z4, Z5, Z7, Z15, Z22, Z24, Z28, Z31; (Bertrand *et al.* 2012)). Amplification protocols and genotyping details are described elsewhere (Bertrand *et al.* 2012, 2014). We performed basic tests to validate the reliability of our microsatellite dataset and assessed within-population genetic variation to quantify genetic diversity. The presence of null alleles was tested with MICRO-CHECKER v.2.2.3 (van Oosterhout *et al.* 2006). We checked for linkage disequilibrium among loci and estimated the allelic richness corrected for sample size (A_R). We also quantified within-population genetic variation by calculating the mean number of alleles per locus (A) along with expected and observed heterozygosities (H_E and H_O). Inbreeding coefficients (F_{IS}) and deviation from Hardy-Weinberg equilibrium were also estimated. R-package {diveRsity} (Keenan *et al.* 2013) was used to make all these calculations.

Plumage colour characterization

We used reflectance spectrophotometry to characterize variation in plumage colour because it provides an objective quantification of colour (Endler 1990; Cuthill *et al.* 1999). Spectral data were recorded in the laboratory with a USB 2000 spectrophotometer connected to a PX-2 light source via a Qt-200 bifurcate optical fibre probe (Ocean Optics, Dunedin, FL, USA). Before each measurement, about 10 feathers were placed on a black surface in a fashion that mimicked the way feathers naturally lay on the bird. We positioned the optical fibre at a standardized three-millimetre distance to the feather surface with a 90° angle. Each measurement consisted of three replicates that were averaged before analysis. After each individual replicate measurement, the feathers were separated and piled again in a random order. For each colour patch, we calculated the brilliance (B) by summing the reflectance values over all wavelengths. We used the tetrahedral colour space model for the analysis of reflectance spectra (Goldsmith 1990; Endler & Mielke 2005). This model has several advantages as it allows the calculation of different meaningful variables in terms of bird vision (Endler & Mielke 2005; Stoddard & Prum 2008). All the following calculations were done using the R-package {pavo} (Maia *et al.* 2013) from R software. We used the spectral sensitivity functions of the Blue Tit (*Cyanistes caeruleus*). Following Endler & Mielke

(2005), we estimated the spherical coordinates (θ, ϕ and r) that define the colour vector in the tetrahedral space. The angles θ and ϕ define the hue of the colour. r is defined as the length of the colour vector and corresponds to the chroma of the colour. Following Stoddard & Prum (2008), we calculated the achieved chroma (r_a) which corresponds to the chroma of a colour relative to the maximum chroma given its hue, which is more informative than r. Each colour patch was then described by four variables (B, θ, ϕ, r_a).

Quantifying environments around the island

Réunion presents considerable climatic heterogeneity. Variation exists not only along the steep altitudinal gradients but also between west and east side of the island both in terms of precipitation and temperature. The eastern side of the island is very rainy with more than 3000 mm of annual rainfall whereas the west side is relatively dry with less than 1000 $mm.yr^{-1}$. Climatic variation is well reflected by transition in habitat types around the island (Strasberg *et al.* 2005). Two main habitat types are found at low elevation around the island: dry lowland forest in the west of the island and lowland rainforest in the east (Strasberg *et al.* 2005; Thébaud *et al.* 2009).

In order to verify that the contact zone between the different lowland colour forms do not correspond to habitat or climate transition but only to physical barriers, we first compared the distribution of the original extent of the habitat types in the lowland to the position of the geographic barriers. To do so, we used ArcGIS10.2 (ESRI, Redlands, CA) and extracted the habitat type at 256 localities around the island on the 200 m elevation isocline. On this isocline, we decided not to take into account the numerous gullies and cut them as the crow flies. The 256 sites correspond to localities separated by 1 km along the isocline (the mean linear distance between two consecutive sites being 730 m). Besides, we obtained 20 environmental data layers covering the whole island of Réunion. 19 of these layers, obtained from the French Meteorological Office (Météo-France, Toulouse), summarize climatic data reflecting different aspects of temperature and precipitation over the last 30 years (see Table S1). The remaining layer was the mean normalized difference vegetation index (NDVI) for the year 2009. NDVI is a good proxy of standing biomass or vegetation cover (Myneni *et al.* 1995). We summarized variation at these different variables by doing a Principal Components Analysis (PCA) on these 20 variables. We then looked at the variation on the two first principal component (PC) (56.1 % and 21.1 % of variance explained, respectively) scores

over distance on the isocline in order to detect any abrupt change in these scores at the contact zone between the different colour forms.

Population structure

To detect restriction in gene flow (if any) between color forms, we used two approaches. First, we evaluated pairwise population differentiation with F_{ST} statistics (Weir & Cockerham 1984). We also computed another genetic index *D* (Jost 2008) because its assumptions differ from F_{ST} estimators that are under debate in the literature. Significance was evaluated with 95% confidence intervals calculated by bootstrapping over 1000 iterations. All these calculations were done using the R-package {diveRsity} (Keenan *et al.* 2013). Identifying Isolation-By-Distance (IBD) pattern may help to understand the different processes involved in the structuring among populations (Garnier *et al.* 2004). We assessed IBD over the entire lowland area by testing with a Mantel test the relationship between $F_{ST}/(1-F_{ST})$ and geographic distance between populations (Rousset 1997). This test was done with the R-package {vegan} (Dixon 2003) using 10 000 permutations. All populations with very low sample sizes ($n < 5$) were removed from these analyses. Moreover, we assumed that birds are more prone to disperse through the lowlands than across the mountain range present in the centre of the island. Thus, to calculate pairwise geographical distance between populations, we decided to use the shortest distance between the projected coordinates of the populations on the 200-m isocline. Second, we performed the Bayesian model-based analysis implemented in STRUCTURE v2.3.2 (Pritchard *et al.* 2000). The rationale for using both F_{ST} statistics and clustering analysis is that F_{ST} allows to assess among-population differentiation whereas STRUCTURE allows to identify clusters separated by barriers to gene flow (Garnier *et al.* 2004; Gayathri Samarasekera *et al.* 2012). We applied the admixture model with correlated allele frequencies and prior information on sampling locality (*locprior* model), a procedure that increases the algorithm's ability to find population clusters when the amount of genetic differentiation is limited, yet has no effect on the optimal number of clusters inferred (Falush *et al.* 2003; Hubisz *et al.* 2009). We conducted this analysis with a number of clusters ranging from one to twenty to determine which number of clusters was best explaining the data. Ten runs were performed for each *k* value. For each run, the program STRUCTURE used Markov Chain Monte Carlo (MCMC) chains of 600,000 iterations of which the first 100,000 were discarded as burn-in. Support for the optimal number of clusters was obtained

by plotting the *ad hoc delta K* statistic (Evanno *et al.* 2005). All individuals were assigned to clusters according to the outputs of *clump v.1.1.2b* (Jakobsson & Rosenberg 2007) which accounts for the variability in individual assignment probabilities across the different runs. The program *distruct 1.1* was used to graphically display the results (Rosenberg 2003). We also repeated the clustering analysis for each transect independently in order to examine potential substructure and test whether the different forms were different in terms of neutral markers.

Phenotypic transition

Quantifying phenotypic differences between each pair of forms

To identify which colour traits were different between forms and to quantify this differentiation for a given plumage patch, we evaluated disparity in the different components of colour (hue, chroma, brilliance). Hue disparity measures the difference in the hue of any two colours and was calculated using a corrected version of equation A3 in Stoddard & Prum (2008). Achieved chroma disparity and normalized brilliance disparity were measured as the absolute value of the difference between the achieved chroma and the normalized brilliance of two colours, respectively. We tested for differences among colour forms in each plumage colour patch using npANOVAs. npANOVA is the univariate equivalent of non-parametric multivariate analysis of variance (npMANOVA; Anderson 2001). These tests were carried out with the function *adonis* in the R-package {vegan}. For each plumage patch the dependent variable was a distance matrix whose elements consisted of colour disparity measures between each pair of individual birds (i.e. normalized brilliance disparity, achieved chroma disparity and hue disparity). In all cases, colour form was used as the independent variable. To test for potential effects of age and sex, these variables were added in these analyses. We also tested for morphological differences between colour forms by performing npANOVAs on each morphological trait with age, sex and ringer identity as factors to control.

For these analyses, we excluded populations placed at the core of hybrid zones (‘Vierge au Parasol’ in the Grand Brulé and ‘Ouaki’ in the St-Etienne river) because they were not attributable to a particular form.

Cline analysis

We used cline models to characterize the position and width of the phenotypic transitions between the different colour forms. For plumage colour, we fitted clines on the metric which best captured the variation between form for each patch, *i.e.* the variable with the highest R^2 for each patch. For morphological traits, we fitted clines only on traits for which colour form explained more than 5 % of variance.

We used a simple sigmoid cline model for each variable:

$$\boldsymbol{p_x} = \frac{\mathbf{1}}{\mathbf{1}+e^{\frac{\mathbf{4}}{\boldsymbol{w}}(\boldsymbol{x}-\boldsymbol{c})}} \qquad \text{(equation 1a)}$$

$$\boldsymbol{\mu_{x}} = \boldsymbol{p_{min}} + (\boldsymbol{p_{max}} - \boldsymbol{p_{min}})\boldsymbol{p_x} \qquad \text{(equation 1b)}$$

where p_x (equation 1a) is a monotone sigmoid function, c and w are the centre and width of the cline, x corresponds to the distance along the transect, μ_x (equation 1b) is the mean value of the trait, p_{min} and p_{max} correspond to the minimum and maximum values of the trait.

Models were adjusted by maximum likelihood. Likelihood functions were optimized using the 'nlminb' function (R Development Core Team). We allowed cline centre (c) to vary between the minimum and the maximum distance of each transect, cline width (w) to vary between $[0,+\infty[$, and p_{min} and p_{max} to range between $]-\infty,+\infty[$. From each model, we extracted the maximum likelihood estimates of cline parameters (c, w, p_{min} and p_{max}) and their 95% confidence intervals. Confidence intervals were determined by approximating the sampling distribution by simulation from the best model (parametric bootstrap with 1000 samples).

All these analyses were done using the R software *v2.15.2* (R Development Core Team).

Results

Physical barriers and environmental transition

The comparison between habitat type and colour form ranges confirmed that habitat transitions do not match the different contact zones (Fig. 3). The three barriers are all located within a homogenous habitat type. The nearest transition between habitat types to one of the contact zone is located at 18 km from the Grand Brulé lava flow complex (transition between GHB and BNB forms; Fig. 3). Besides, using independent data summarizing climatic variation and actual vegetation openness we showed that contact zones between the different colour forms do not correspond to abrupt bioclimatic transition zones. These barriers are all located either in smooth environmental transition or in homogeneous climatic zones (Fig. 3).

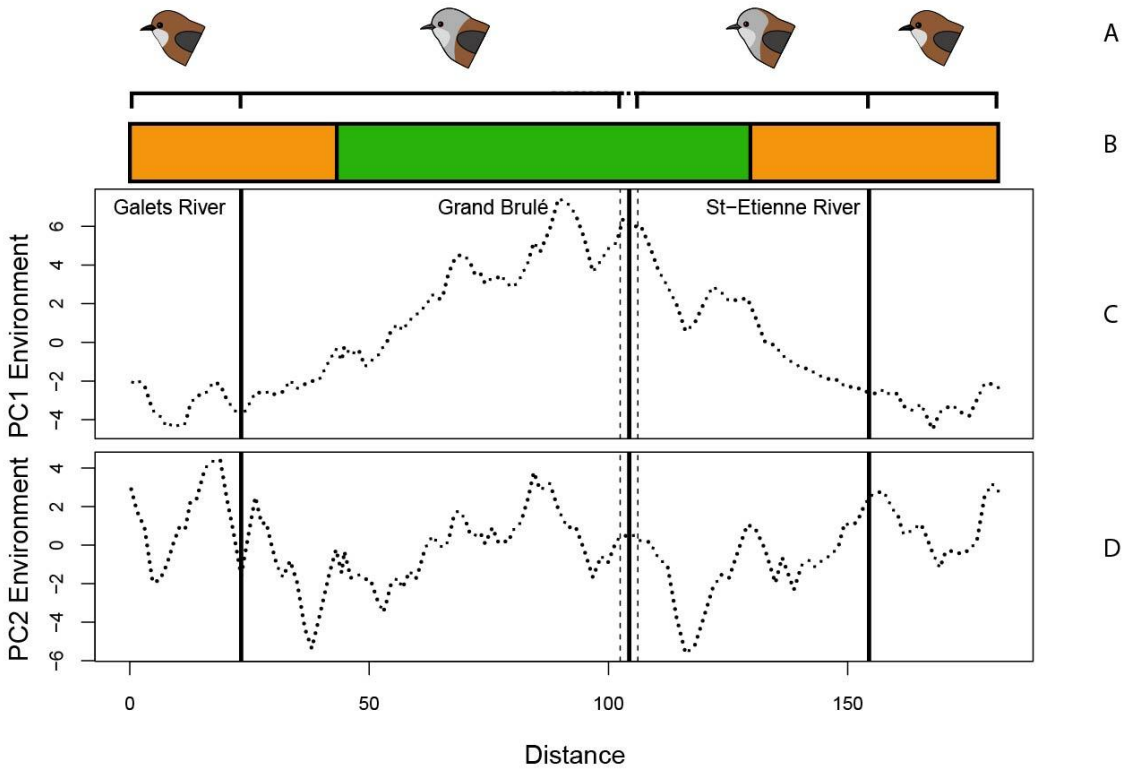


Figure 3: Representation of different characteristics over distance around the island (on the 200 m elevation isoclines; distance from the middle of the range of LBHB going clockwise around the island). **A)** Geographic range of the three colour forms. **B)** Original habitat extant: dry lowland forest in orange and lowland rainforest in green. **C)** Scores of the first principal component (56.1 %) of a PCA on environmental data **D)** Scores of the second principal component (21.1 %) of a PCA on environmental data. Black lines represent the different physical barriers, from left to right: Galets River, Grand Brulé lava complex and St-Etienne River.

Genetic structure and among population differentiation

None of the 11 microsatellite loci exhibited significant deviation from Hardy-Weinberg equilibrium. No significant linkage disequilibrium was found across all pairs of loci. All the localities sampled presented similar levels of within-population polymorphism (with *A* ranging from 4.18 to 9.73 and A_R ranging from 3.4 to 5.2; Table 1). Microsatellite loci presented differences in their relative level of polymorphism with number of alleles per locus ranging from 6 (for Z1) to 38 (for Z15) (Table 1). This important allelic polymorphism was associated with high mean heterozygosities (H_O = 0.78 and H_E = 0.75). The mean inbreeding coefficient (F_{IS}) could not be statistically differentiated from zero in all but seven localities indicating no major significant deviation from panmixia in most of the localities (Table 1). These deviations were all negative, denoting an excess of heterozygotes.

The overall F_{ST} value suggested a low but significant pattern of among-population neutral genetic differentiation (θ_{ST} = 0.0594; 95% CI: 0.0397 - 0.0720). Almost all pairwise F_{ST} comparisons between localities were significant (lower CI > 0) and are consistent with a pattern of global neutral genetic differentiation (Table S2). Pairwise *D* values were highly correlated with F_{ST} values (Mantel test: r = 0.97, p = < 0.001) and indicated the same pattern (Table S3). Genetic differentiation was slightly correlated with geographic distance indicating a significant pattern of IBD (Mantel test: r = 0.12, p = 0.02; Fig S1).

Evanno's criterion (ΔK) derived from the STRUCTURE analysis on the whole data set suggested an optimal number of genetic clusters equal to two (L(*K*) = -13 996.7; ΔK = 19.05 ; Fig. 4A and 4B) and indicates that differentiation exists among populations. These two clusters are consistent with geography but do not correspond to the geographic range of the different colour forms. The first genetic cluster includes the sampling localities found in the North-West of the island. This cluster is homogeneous from Ermitage to St-André and includes the contact zone between the LBHB and GHB form (Fig. 4). The second genetic cluster mainly includes the sampling localities found in the South of the island (Fig. 4). In other words, the different contact zones are found within homogeneous genetic clusters. Populations with mixed ancestry are not geographically structured as we found them at the core of the geographic range of each form (Fig. 4). STRUCTURE results for each separate transect did not identify any biologically meaningful clusters (Fig. S2, S3 and S4). Overall,

these results indicate that there is small-scale genetic differentiation among populations but genetic structure is unrelated to colour forms.

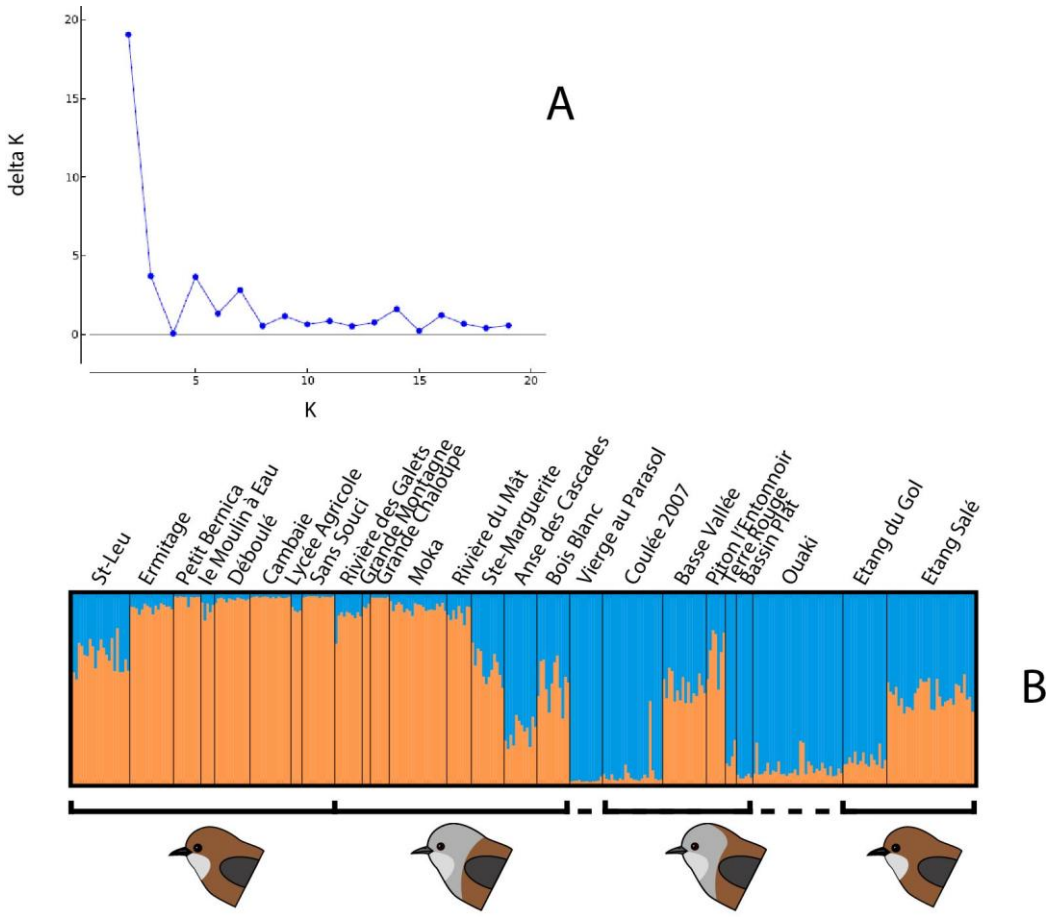


Figure 4: A) delta K statistic plotted against the number of clusters (K). **B)**Admixture proportions as inferred from genetic clustering analyses (STRUCTURE). Each bar represents an individual. Each colour reflects the likelihood of belonging to one of the inferred genetic clusters (at $K = 2$).

Table 2: Between-form differences in morphology and colour. Measures of morphological difference were calculated for each traits and each pair of individuals. Measures of colour disparity were calculated for each plumage patch and each pair of individuals. The within- and among-variant averages are given, showing for which patches and which variables the plumage of *Z. borbonicus* is more different among variants than within. Results of npANOVAs are given with *P*-values significant under the 5% error threshold in bold.

			npANOVAs							
			Ringer		Age		Sex		Form	
	Average Within-Variant	Average Between-Variant	R^2	*p*	R^2	*p*	R^2	*p*	R^2	*p*
Morphology:										
Tarsus length difference (mm)	6.68 x 10-1	7.01 x 10-1	0.02	**0.02**	0	0.47	0.01	0.09	0.02	**0.06**
Wing length difference (mm)	2.43	2.27	0.01	0.05	0.05	**< 0.001**	0.02	**0.003**	0.02	**0.05**
Tail length difference (mm)	1.94	2.33	0.01	0.05	0.03	**0.003**	0.01	0.22	0.07	**< 0.001**
Bill length difference (mm)	3.89 x 10-1	6.87 x 10-1	0.01	0.15	0.02	**0.003**	0.01	**0.04**	0.21	**< 0.001**
Bill width difference (mm)	2.40 x 10-1	2.58 x 10-1	0.01	0.17	0.03	**< 0.001**	0.01	0.24	0.08	**< 0.001**
Bill depth differnce (mm)	2.30 x 10-1	2.22 x 10-1	0.04	**< 0.001**	0.03	**< 0.001**	0.02	**0.002**	0.08	**< 0.001**
Back:										
Normalized brilliance disparity (%)	8.65 x 10-3	8.56 x 10-3			0.02	0.05	0	0.86	0.02	0.11
Achieved chroma disparity (%)	9.10 x 10-2	9.15 x 10-2			0.04	**0.006**	0.01	0.33	0.01	0.15
Hue disparity (radian)	9.02 x 10-2	9.60 x 10-2			0.11	**< 0.001**	0.01	0.39	0.02	0.26
Flank:										
Normalized brilliance disparity (%)	1.41 x 10-2	1.67 x 10-2			0.01	0.14	0.01	0.23	0.11	**< 0.001**
Achieved chroma disparity (%)	1.25 x 10-1	1.23 x 10-1			0.04	**0.005**	0	0.93	0.03	0.07
Hue disparity (radian)	9.92 x 10-1	9.76 x 10-2			0.02	**0.07**	0.02	0.06	0.01	0.69
Head:										
Normalized brilliance disparity (%)	1.04 x 10^{-2}	1.16x 10^{-2}			0.03	**0.01**	0	0.81	0.05	**0.004**
Achieved chroma disparity (%)	9.7 x 10-2	2.10 x 10-1			0.02	**0.017**	0.01	**0.08**	0.44	**< 0.001**
Hue disparity (radian)	1.46 x 10-1	2.07 x 10-1			0.11	**< 0.001**	0.01	0.25	0.24	**< 0.001**
Belly:										
Normalized brilliance disparity (%)	3.8 x 10-2	5.00 x 10-2			0.01	0.46	0.1	**< 0.001**	0.15	**< 0.001**
Achieved chroma disparity (%)	8.26 x 10-2	9.44 x 10-2			0.03	**0.06**	0	0.73	0.07	**0.02**
Hue disparity (radian)	3.07 x 10-1	3.26 x 10-1			0.21	**< 0.001**	0.02	0.07	0.14	**< 0.001**

Clinal variation

Colour forms differed for three of the four plumage patches, for tail length and bill characteristics (Table 2). We thus fitted clines only for these different traits. We showed that the cline for head coloration was very narrow (w = 681 m) on the LBHB - GHB transect (Fig. 5; Table 3). On the same contact zone, we found evidence for change in bill length with GHB birds having a longer bill, and showing a wider cline (~20km wide; Table 3). These clines were centred around the Galets river. The transect between LBHB and GHB form was the best sampled among the three contact zones with dense sampling around the centre of the hybrid zone, a prerequisite for accurate estimation of width of the hybrid zone (Raufaste *et al.* 2005). For the GHB - BNB transect, head colouration showed also a marked transition in the North of Grand Brulé (Fig. 5; Table 3) but the estimated width was substantially larger than on the LBHB - GHB, probably because of the limited sampling at the centre of the zone. On St-Etienne river, the contact zone between GHB and BNB forms, feather sampling was too limited and we were not able to fit clines. The results of clinal analysis for the other traits are not presented because model did not converge either because of poor sample sizes or because sigmoid curve did not describe well the data.

Based on visual observations and previous results from Gill (1973), the width of head colour transition is likely to be very similar in the three hybrid zones. Thus, we are confident that the width inferred for the LBHB - GHB transect (< 1km) is comparable to that of the two other hybrid zones.

Table 3: Maximum likelihood estimates and their confidence intervals for the different models.

	centre (c)	width (w)	p_{min}	p_{max}
LBHB - GHB (Galets river)				
Head Colour ($r_{achieved}$)	29853.5 (28782.1 - 29958. 1)	681.2 (98 - 7004)	0.13 (0.1 - 0.16)	0.42 (0.40 - 0.44)
Bill Length	25810.9 (22245.7 - 30655.3)	18551.0 (3715.3 - 52190.9)	-0.30 (-0.49 - -0.20)	0.39 (0.29 - 0.54)
GHB - BNB (Grand Brulé)				
Head Colour ($r_{achieved}$)	30877.4 (25543.0 - 34449.0)	23697.9 (6872.10 - 57749.2)	0.13 (0.05 - 0.18)	0.42 (0.07 - 0.10)

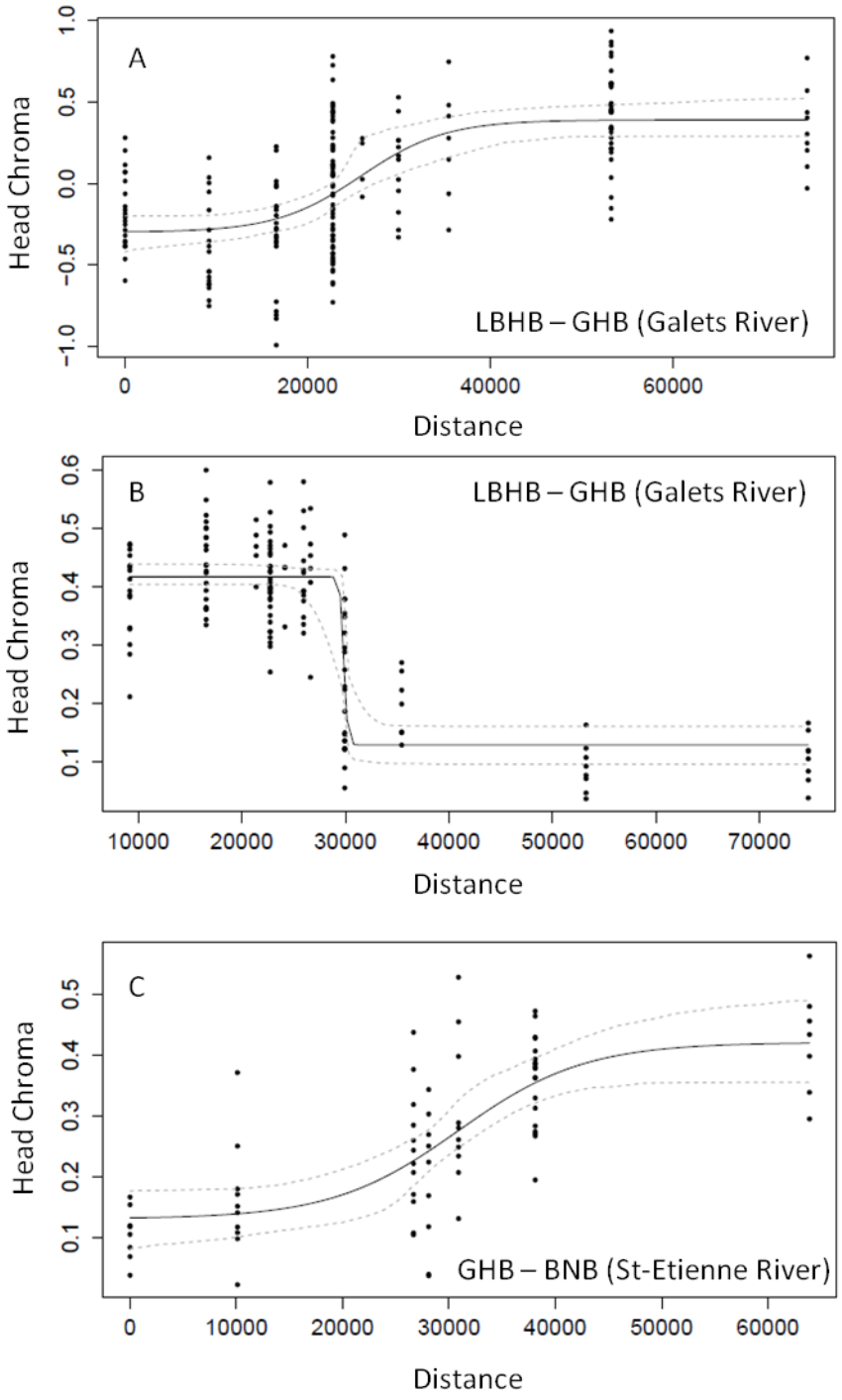


Figure 5: Spatial cline in phenotypic traits in the different hybrid zones. Culmen length on the LBHB-GHB transect (**A**). Head colour (chroma) on the LBHB-GHB transect (**B**). Head colour (chroma) on the GHB-BNB transect (**C**). Black points correspond to the observed values, black solid lines denote the best-fitting curve of each variable and dashed lines correspond to 95% confidence intervals.

Discussion

Studies investigating variation in hybrid zones have shown that introgression does not affect all parts of the genome homogeneously. This is well illustrated by cases in which genomic data are available but also by comparisons between phenotypic traits and neutral genetic markers. However, in most cases, even if variability in introgression rates exists, genome-wide patterns of divergence are found which means that divergence is widespread across the genome (Harrison & Larson 2014). This generally entails reduced introgression at most part of the genome and can be detected even with a modest number of microsatellite loci (Harrison & Larson 2014). Using population structure and clinal analyses, we made the first examination of neutral genetic differentiation and phenotypic variation patterns across each lowland hybrid zones between colour forms of *Z. borbonicus*. Strikingly, we showed that the abrupt spatial transition in plumage characteristics at the different contact zones is in sharp contrast with the absence of neutral genetic structure associated with these colour forms. Although neutral genetic differentiation exists, the patchy, small-scale population differentiation found is consistent with limited dispersal as previously demonstrated in the species; Bertrand *et al.* 2014; Annexe 1). Head colouration was not the only trait that differed between lowland forms as bill length presented as well sigmoidal clinal transition at the Galets river. We also confirmed with objective data on climate, vegetation openness and habitat types that the 'physical' barriers at the limit between the different colour forms did not correspond to steep environmental changes. The stability of these zones, the absence of neutral genetic structure associated to colour forms is thus remarkable considering the very distinct phenotypic differences and the extreme narrowness of the different hybrid zones. These results, together with a previous molecular study from Milá *et al.* (2010), suggest that these hybrid zones are likely associated with plumage colour associated genes with very little divergence at other markers.

There are several possible and non-mutually exclusive hypotheses to explain this pattern. First, the divergence between the different colour forms may have occurred without geographic isolation and corresponds to primary intergradation. It has been shown that divergence with gene flow is promoted by strong ecological selection (Pinho & Hey 2010) which is not likely to be the case among lowland forms of *Z. borbonicus.* In such a case, we would have expected colour genes to be related to such exogenous factors. Here, the different colour forms experience different habitat types and the contact zones are not associated with

particular environmental transitions. Therefore, environmental and climatic conditions do not seem to be involved in the maintenance of these different hybrid zones. Second, geographic isolation may have played an important role in the emergence of this pattern. Réunion is a very small island (2,512 km²) but its active volcanic history potentially created barriers between different isolated populations. Thus, past allopatry and subsequent secondary contact is very likely on the island (see Milá *et al.* 2010 for more details). Then, two different phenomena may have produced the actual pattern. On one hand, it is possible that divergence has been more important in the past but admixture following secondary contact has already homogenized neutral genetic pools. On the other hand, incomplete lineage sorting is known to leave similar footprint in the genome than recent gene flow (Hudson & Coyne 2002; Broughton & Harrison 2003). Shared ancestral polymorphism is favored by recent divergent times and large effective population sizes, two characteristics which are likely in the species. Indeed, Warren *et al.* (2006) have shown that divergence in *Z. borbonicus* took place during the last 420,000 years and it has been highlighted that melanin pigmentation can evolve rapidly in birds (Omland & Lanyon 2000; Ödeen & Björklund 2003; Milá *et al.* 2007). Besides, Bourgeois (2013) has already shown that effective population size estimates were high. Regardless of the scenario which produced this pattern, there are mechanisms which are maintaining these hybrid zones stable.

The maintenance of abrupt transitions in head colouration may be the product of pre-zygotic barriers. Plumage coloration is well known to play a role in the maintenance of avian hybrid zones (*e.g.* Sæther *et al.* 2007; Haas *et al.* 2010; Hughes *et al.* 2011) through assortative mating. Moreover, Cornuault *et al.* (*in press* - Chapitre 1) have demonstrated that colour differences between lowland forms are distinguishable by birds and thus usable as mate choice criteria. Two main mechanisms are recognized to explain assortative mating in hybrid zones. It can either be due to genetically determined mate preferences or to sexual imprinting. Genetically preferred mate choice may play a predominant role in the maintenance of hybrid zones (Coyne & Orr 2004; Sæther *et al.* 2007). However, gene flow is expected to break down association between preference and mate choice traits (Felsenstein 1981). A good brake to recombination seems to be sex-linkage of the genes involved in mate choice and phenotypic differences. This linkage is supposed to prevent hybrid zone disappearance (Hall & Kirkpatrick 2006). It has been shown in several avian hybrid zones that introgression was reduced on the Z chromosome and this potentially contributes to prezygotic isolation through sexual selection (*e.g.* Sæther *et al.* 2007; Carling *et al.* 2008; Storchová *et*

al. 2010). Sexual imprinting might also play a role in the maintenance of this hybrid zone. It has been defined as the process by which the nestlings will learn their taxa characteristics to choose their mates as adults. Sexual imprinting has been shown to trigger hybrid disadvantage and thus impede gene flow (Cooke & McNally 1975; Irwin & Price 1999; Servedio *et al.* 2009). Nevertheless, determining which mechanism of prezygotic isolation is at play in this system will require further analyses.

Our study sheds light on hybrid zones where neutral genetic markers show no differentiation associated to colour forms whereas plumage colour show abrupt transition at the different physical barriers. We can conclude that the maintenance of these different hybrid zones probably involves few genomic regions which are linked to colour differences and that the majority of the genome, which is unlinked to the plumage genes, shows no differentiation at the level of these hybrid zones. This pattern strongly contrasts with many hybrid zones where many genomic island of divergence are involved in reproductive isolation and most of molecular markers show reduced introgression and steep clines (Harrison & Larson 2014). These hybrid zones stand out as ones of the handful of cases which present similar patterns (*e.g.* Whibley *et al.* 2006; Nadeau *et al.* 2013; Poelstra *et al.* 2014). Strikingly, in most of these cases pre-mating barriers seem to play an important role in genetic and phenotypic differentiation. Thus, even if it remains to be tested, studying mate choice at the core of the different hybrid zones will allow us to determine if homogamy effectively maintains the forms separated. Comparisons of neutral markers and phenotypes in hybrid zones give only indirect information on the underlying genetic architecture of the isolation between the different colour forms. Thus, we will much benefit from the genomic analysis of these hybrid zones which is currently underway. Determining the number and the placement of the genomic regions which are indeed involved in genetic and phenotypic differentiation will be an important step towards the understanding of the maintenance of these hybrid zones. It will also be nice to examine whether the same genomic regions are responsible for the observed differences between each pairs of forms. Moreover, an interesting comparative analysis with another hybrid zone in the system in which environmental differences are marked and neutral genetic differentiation is steep (Chapitre 3) will provide new insights into the dynamics of these different kind of hybrid zones.

Acknowledgements

Fieldwork was facilitated by the outstanding efforts of Guillaume Gélinaud, Dominique Strasberg, Ben Warren, Magali Thierry, René-Claude Billot, Jean-Michel Probst, Isabelle Henry and Vincent Leconte. We gratefully thank the Réunion National Park for permission to conduct fieldwork. Marc Salamolard and Benoît Lequette provided valuable help with fieldwork and logistics. J.B, Y.B, B.D and J.C were supported by MESR (Ministère de l'Enseignement Supérieur et de la Recherche) PhD scholarships. The research was supported by Agence Française pour le Développement grants to CT, the Fondation pour la Recherche sur la Biodiversité (FRB) through its Centre for Synthesis and Analysis of Biodiversity (CESAB), the 'Laboratoire d'Excellence' TULIP (ANR-10-LABX-41), and the SYNTHESYS Project (http://www.synthesys.info/) which is financed by European Community Research Infrastructure Action under the FP7 "Capacities" Programme at the Museo Nacional de Ciencias Naturales (CSIC) of Madrid, Spain.

References

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, **26**, 32–46.

Baldassarre DT, White TA, Karubian J, Webster MS (2014) Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. *Evolution*, **68**, 2644–2657.

Barton NH, Hewitt GM (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Bertrand JAM, Bourgeois YXC, Delahaie B *et al.* (2014) Extremely reduced dispersal and gene flow in an island bird. *Heredity*, **112**, 190–196.

Bertrand JAM, García-Jiménez R, Bourgeois Y *et al.* (2012) Isolation and characterization of twelve polymorphic microsatellite loci for investigating an extreme case of microgeographical variation in an island bird (Zosterops borbonicus). *Conservation Genetics Resources*, **4**, 323–326.

Bourgeois Y (2013) Génétique évolutive d'un cas extrême de polymorphisme de la coloration du plumage chez un oiseau insulaire, Zosterops borbonicus (Zosteropideae). Université Toulouse 3 Paul Sabatier, Toulouse.

Broughton RE, Harrison RG (2003) Nuclear gene genealogies reveal historical, demographic and selective factors associated with speciation in field crickets. *Genetics*, **163**, 1389–1401.

Carling MD, Brumfield RT, Webster M (2008) Haldane's Rule in an Avian System: Using Cline Theory and Divergence Population Genetics to Test for Differential Introgression of Mitochondrial, Autosomal, and Sex-Linked Loci Across the Passerina Bunting Hybrid Zone. *Evolution*, **62**, 2600–2615.

Carneiro M, Albert FW, Afonso S *et al.* (2014) The Genomic Architecture of Population Divergence between Subspecies of the European Rabbit. *PLoS Genet*, **10**, e1003519.

Cooke F, McNally CM (1975) Mate selection and colour preferences in lesser snow geese. *Behaviour*, **53**, 151–170.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates Inc.,U.S., Sunderland, Mass.

Cuthill IC, Bennett ATD, Partridge J. C., Maier EJ (1999) Plumage Reflectance and the Objective Assessment of Avian Sexual Dichromatism. *The American Naturalist*, **153**, 183–200.

Dixon P (2003) VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, **14**, 927–930.

Endler JA (1990) On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society*, **41**, 315–352.

Endler JA, Mielke PW (2005) Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, **86**, 405–431.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, **164**, 1567–1587.

Felsenstein J (1981) Skepticism Towards Santa Rosalia, or Why are There so Few Kinds of Animals? *Evolution*, **35**, 124–138.

Garnier S, Alibert P, Audiot P, Prieur B, Rasplus J-Y (2004) Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, Carabus solieri. *Molecular Ecology*, **13**, 1883–1897.

Gayathri Samarasekera GDN, Bartell NV, Lindgren BS *et al.* (2012) Spatial genetic structure of the mountain pine beetle (Dendroctonus ponderosae) outbreak in western Canada: historical patterns and contemporary dispersal. *Molecular Ecology*, **21**, 2931–2948.

Gay L, Crochet P-A, Bell DA, Lenormand T (2008) Comparing Clines on Molecular and Phenotypic Traits in Hybrid Zones: A Window on Tension Zone Models. *Evolution*, **62**, 2789–2806.

Gill FB (1973) Intra-island variation in the Mascarene white-eye Zosterops borbonica. *Ornithological Monographs*, iii–66.

Goldsmith TH (1990) Optimization, constraint, and history in the evolution of eyes. *The Quarterly Review of Biology*, **65**, 281–322.

Haas F, Knape J, Brodin A (2010) Habitat preferences and positive assortative mating in an avian hybrid zone. *Journal of Avian Biology*, **41**, 237–247.

Hall DW, Kirkpatrick M (2006) Reinforcement and sex linkage. *Evolution; International Journal of Organic Evolution*, **60**, 908–921.

Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, **7**, 69–128.

Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, **105**, 795–809.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution; International Journal of Organic Evolution*, **56**, 1557–1565.

Hughes JM, Toon A, Mather PB, Lange CL (2011) Maintenance of a Hybrid Zone: The Role of Female Mate Choice. *The Auk*, **128**, 688–695.

Irwin DE, Brelsford A, Toews DPL, MacDonald C, Phinney M (2009) Extensive hybridization in a contact zone between MacGillivray's warblers Oporornistolmiei and mourning warblers O. philadelphia detected using molecular and morphological analyses. *Journal of Avian Biology*, **40**, 539–552.

Irwin DE, Price T (1999) Sexual imprinting, learning and speciation. *Heredity*, **82**, 347–354.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics (Oxford, England)*, **23**, 1801–1806.

Jiggins null, Mallet null (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, **15**, 250–255.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782–788.

Larson EL, White TA, Ross CL, Harrison RG (2014) Gene flow and the maintenance of species boundaries. *Molecular Ecology*, **23**, 1668–1678.

Maia R, Eliason CM, Bitton P-P, Doucet SM, Shawkey MD (2013) pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution*, **4**, 906–913.

Milá B, McCormack JE, Castañeda G, Wayne RK, Smith TB (2007) Recent postglacial range expansion drives the rapid diversification of a songbird lineage in the genus Junco. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 2653–2660.

Milá B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: Zosterops borbonicus). *BMC Evolutionary Biology*, **10**, 158.

Myneni RB, Hall FG, Sellers PJ, Marshak AL (1995) The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 481–486.

Nadeau NJ, Martin SH, Kozak KM *et al.* (2013) Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology*, **22**, 814–826.

Ödeen A, Björklund M (2003) Dynamics in the evolution of sexual traits: losses and gains, radiation and convergence in yellow wagtails (Motacilla flava). *Molecular Ecology*, **12**, 2113–2130.

Omland KE, Lanyon SM (2000) Reconstructing Plumage Evolution in Orioles (icterus): Repeated Convergence and Reversal in Patterns. *Evolution*, **54**, 2119–2133.

Van Oosterhout C, Joyce DA, Cummings SM *et al.* (2006) Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (poecilia reticulata). *Evolution*, **60**, 2562–2574.

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 215–230.

Poelstra JW, Vijay N, Bossu CM *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, **344**, 1410–1414.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945–959.

Pyle P (1997) *Identification Guide to North American Birds, Part I: Columbidae to Ploceidae*. State Creek Press, Bolinas, CA.

Raufaste N, Orth A, Belkhir K *et al.* (2005) Inferences of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society*, **84**, 593–616.

R Development Core Team R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713–727.

Rosenberg NA (2003) distruct: a program for the graphical display of population structure: PROGRAM NOTE. *Molecular Ecology Notes*, **4**, 137–138.

Rousset F (1997) Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance. *Genetics*, **145**, 1219–1228.

Sæther SA, Sætre G-P, Borge T *et al.* (2007) Sex Chromosome-Linked Species Recognition and Evolution of Reproductive Isolation in Flycatchers. *Science*, **318**, 95–97.

Servedio MR, Sæther SA, Sætre G-P (2009) Reinforcement and learning. *Evolutionary Ecology*, **23**, 109–123.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69**, 82–90.

Shaw KL, Mullen SP (2011) Genes versus phenotypes in the study of speciation. *Genetica*, **139**, 649–661.

Singhal S, Moritz C (2012) Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. *Evolution*, **66**, 1474–1489.

Stoddard MC, Prum RO (2008) Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. *The American Naturalist*, **171**, 755–776.

Storchová R, Reif J, Nachman MW (2010) Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. *Evolution; International Journal of Organic Evolution*, **64**, 456–471.

Strasberg D, Rouget M, Richardson DM *et al.* (2005) An Assessment of Habitat Diversity and Transformation on La Réunion Island (Mascarene Islands, Indian Ocean) as a Basis for Identifying Broad-scale Conservation Priorities. *Biodiversity & Conservation*, **14**, 3015–3032.

Thébaud C, Warren BH, Cheke AC (2009) Mascarene Islands, Biology. In: *Encyclopedi of Islands* , pp. 612 – 619. Berkeley, USA.

Warren BH, Bermingham E, Prys-Jones RP, Thébaud C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. *Molecular Ecology*, **15**, 3769–3786.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, **38**, 1358.

Whibley AC, Langlade NB, Andalo C *et al.* (2006) Evolutionary Paths Underlying Flower Color Variation in Antirrhinum. *Science*, **313**, 963–966.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Supporting information

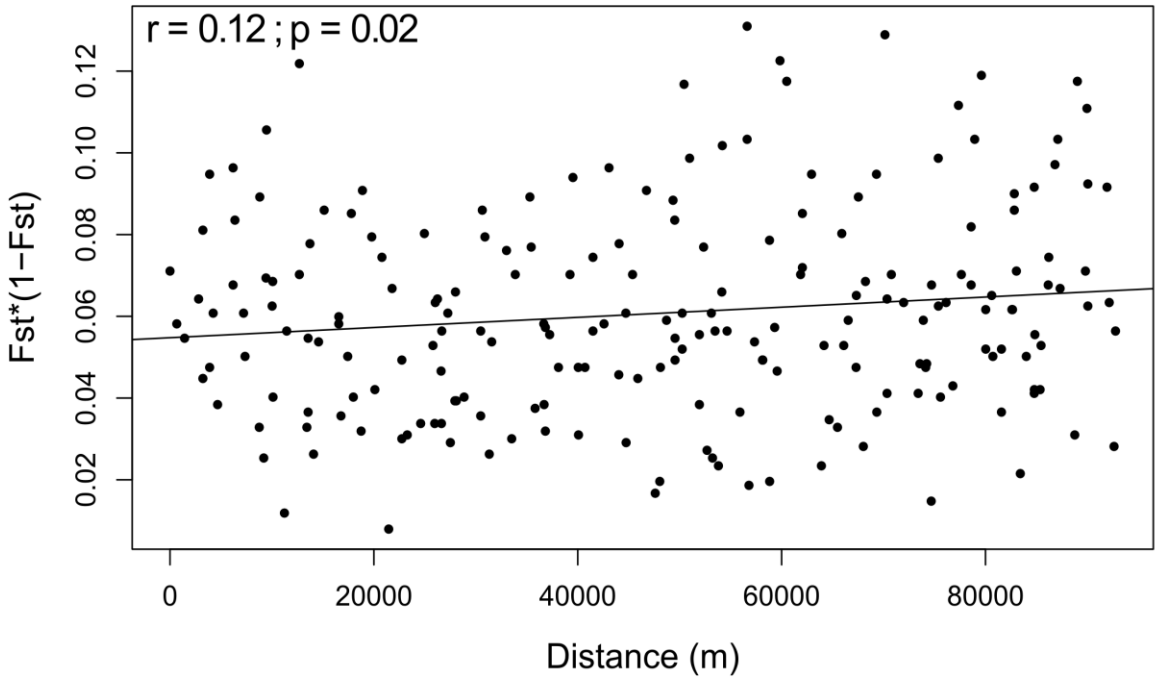


Figure S1: Relationship between FST and geographic distance: a slightly significant pattern of Isolation-by-Distance. Results of the mantel test are given.

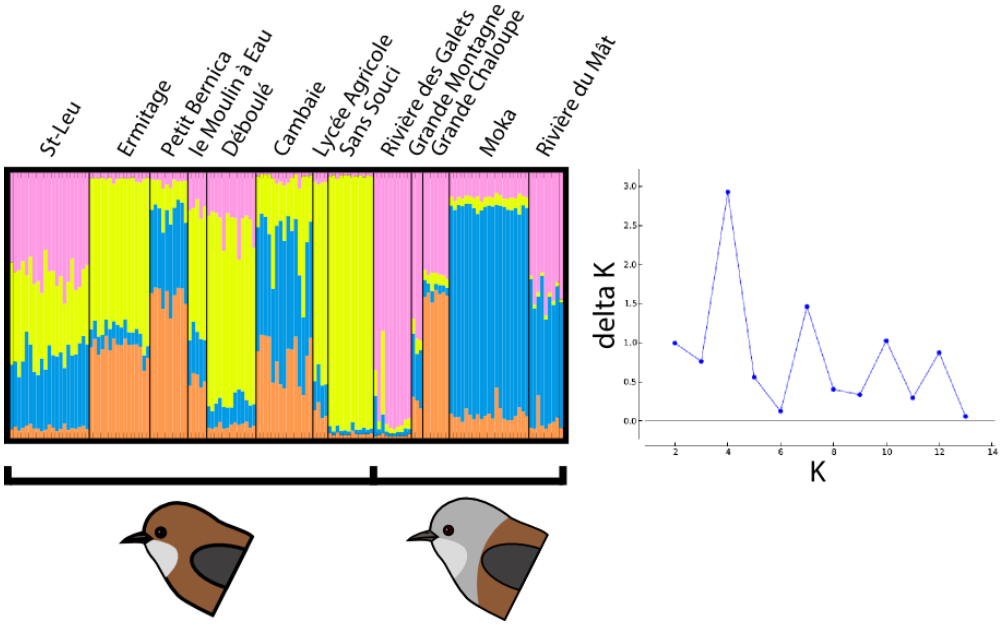


Figure S2: STRUCTURE analysis on the LBHB-GHB transect (Galets River). **Left)** Admixture proportions as inferred from genetic clustering analyses (STRUCTURE). Each bar represents an individual. Each colour reflects the likelihood of belonging to one of the inferred genetic clusters (at *K* = 2). **Right)** delta K statistic plotted against the number of clusters (K).

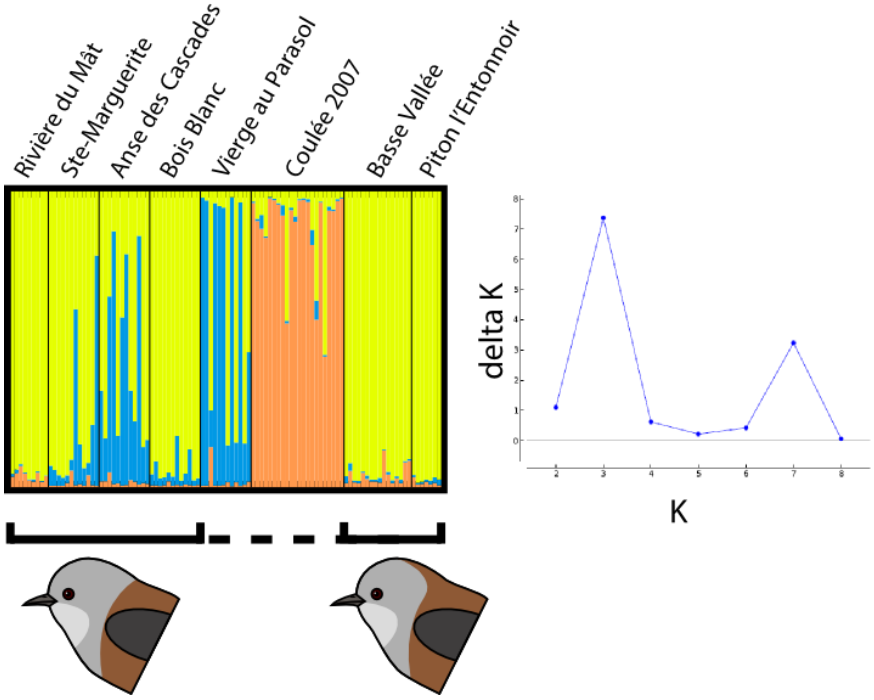


Figure S3: STRUCTURE analysis on the GHB-BNB transect (Grand Brûlé). **Left)** Admixture proportions as inferred from genetic clustering analyses (STRUCTURE). Each bar represents an individual. Each colour reflects the likelihood of belonging to one of the inferred genetic clusters (at *K* = 2). **Right)** delta K statistic plotted against the number of clusters (K).

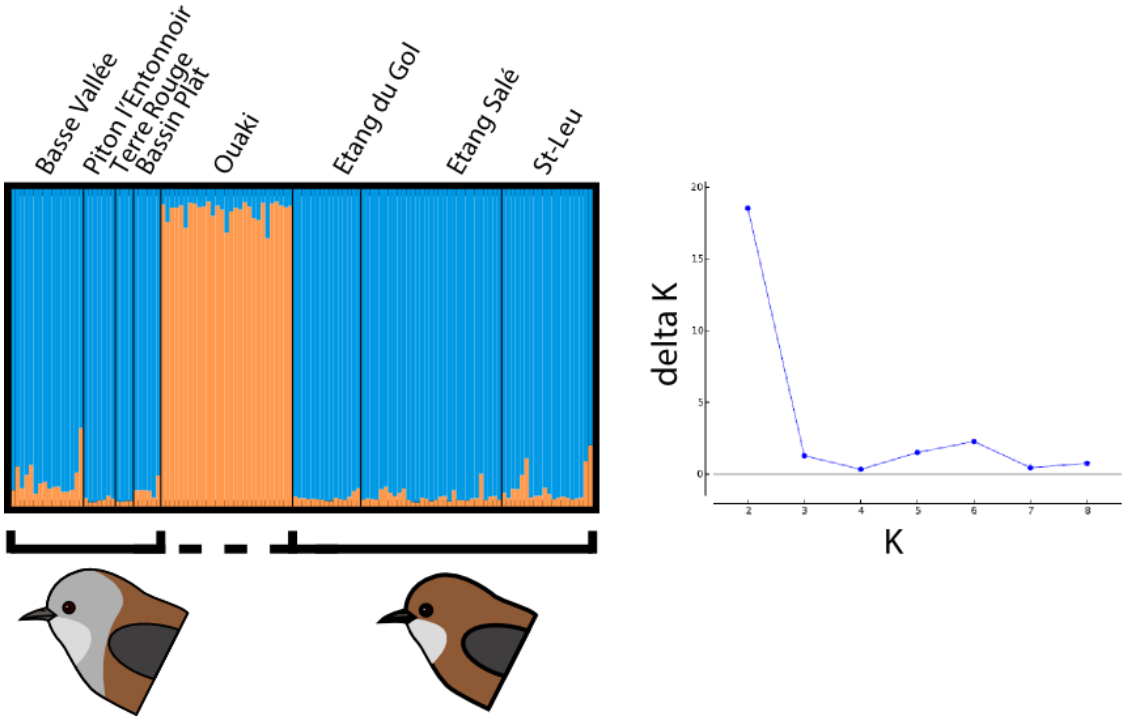


Figure S4: STRUCTURE analysis on the BNB-LBHB transect (St-Etienne River). **Left)** Admixture proportions as inferred from genetic clustering analyses (STRUCTURE). Each bar represents an individual. Each colour reflects the likelihood of belonging to one of the inferred genetic clusters (at $K = 2$). **Right)** delta K statistic plotted against the number of clusters (K).

Table S1: Bioclimatic variables and elevation were obtained at high resolution from the French Meteorological Office (Météo-France, Toulouse). NDVI data were monitored by the MODIS device (NASA's Terra mission; http://e4eil01.cr.usgs.gov:22000/WebAccess/drill?attrib=esdt& esdt=MOD13A3.5&group=MOLT).

Variables	Description	Resolution
BIO1	Annual Mean Temperature	132 m x 132 m
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	132 m x 132 m
BIO3	Isothermality (BIO2/BIO7) (* 100)	132 m x 132 m
BIO4	Temperature Seasonality (standard deviation *100)	132 m x 132 m
BIO5	Max Temperature of Warmest Month	132 m x 132 m
BIO6	Min Temperature of Coldest Month	132 m x 132 m
BIO7	Temperature Annual Range (BIO5-BIO6)	132 m x 132 m
BIO8	Mean Temperature of Wettest Quarter	132 m x 132 m
BIO9	Mean Temperature of Driest Quarter	132 m x 132 m
BIO10	Mean Temperature of Warmest Quarter	132 m x 132 m
BIO11	Mean Temperature of Coldest Quarter	132 m x 132 m
BIO12	Annual Precipitation	132 m x 132 m
BIO13	Precipitation of Wettest Month	132 m x 132 m
BIO14	Precipitation of Driest Month	132 m x 132 m
BIO15	Precipitation Seasonality (Coefficient of Variation)	132 m x 132 m
BIO16	Precipitation of Wettest Quarter	132 m x 132 m
BIO17	Precipitation of Driest Quarter	132 m x 132 m
BIO18	Precipitation of Warmest Quarter	132 m x 132 m
BIO19	Precipitation of Coldest Quarter	132 m x 132 m
NDVI	Normalized Difference Vegetation Index	1000 m x 1000 m

Table S2: Pairwise geographic distances (below the diagonal) and estimates of F_{ST} based on 11 microsatellite markers (above the diagonal).

	St Leu	Ermitage	Petit Bernica	Moulin à Eau	Cambaie	Sans Souci	Rivière des Galets	Grande Chaloupe	Moka	St André	Ste Marguerite	Anse des Cascades	Bois Blanc	Vierge au Parasol	Coulée 2007	Basse Vallée	Piton l'Entonnoir	Bassin Plat	Ouaki	Etang du Gol	Etang Salé
St Leu	-	0.026	0.062	0.031	0.052	0.035	0.035	0.084	0.026	ns	0.043	0.059	ns	0.070	0.043	ns	ns	0.076	0.030	0.033	0.027
Ermitage	9.20	-	0.053	ns	0.058	0.037	0.053	0.069	0.048	0.034	0.042	0.068	0.029	0.077	0.066	0.029	0.019	0.108	0.040	0.041	0.032
Petit Bernica	16.56	7.36	-	0.073	0.108	0.075	0.074	0.101	0.062	0.052	0.074	0.102	0.081	0.136	0.127	0.067	0.056	0.135	0.085	0.099	0.095
Moulin à Eau	22.74	13.55	6.19	-	0.077	0.047	0.050	0.076	0.037	0.040	0.078	0.073	0.055	0.100	0.067	0.038	ns	0.117	0.055	0.060	0.033
Cambaie	22.74	13.55	6.19	0.00	-	0.089	0.106	0.142	0.060	0.059	0.094	0.090	0.066	0.095	0.103	0.055	0.069	0.155	0.065	0.081	0.061
Sans Souci	25.97	16.77	9.41	3.23	3.23	-	0.062	0.120	0.065	0.063	0.086	0.111	0.045	0.138	0.109	0.044	0.051	0.143	0.060	0.065	0.050
Rivière des Galets	26.63	17.43	10.07	3.89	3.89	0.66	-	0.099	0.049	ns	0.052	0.073	0.068	0.117	0.073	0.056	0.051	0.136	0.071	0.076	0.050
Grande Chaloupe	35.43	26.24	18.88	12.69	12.69	9.47	8.81	-	0.094	0.076	0.098	0.088	0.070	0.152	0.128	0.102	0.077	0.106	0.106	0.115	0.092
Moka	53.22	44.03	36.67	30.48	30.48	27.25	26.59	17.79	-	ns	0.057	0.050	0.052	0.084	0.049	0.050	0.044	0.117	0.053	0.068	0.050
St André	74.68	65.48	58.12	51.93	51.93	48.71	48.05	39.24	21.45	-	ns	0.060	0.041	0.087	0.050	ns	ns	0.076	0.053	0.060	0.032
Ste Marguerite	84.78	75.59	68.23	62.04	62.04	58.81	58.15	49.35	31.56	10.10	-	0.064	0.042	0.081	0.071	0.062	ns	0.099	0.063	0.066	0.072
Anse des Cascades	84.84	92.14	84.78	78.59	78.59	75.37	74.71	65.90	48.11	26.66	16.56	-	0.058	0.065	0.060	0.068	0.059	0.111	0.057	0.056	0.076
Bois Blanc	83.42	92.61	86.21	80.02	80.02	76.80	76.14	67.33	49.54	28.09	17.99	1.43	-	0.069	0.067	0.035	ns	ns	0.038	0.036	0.038
Vierge au Parasol	80.61	89.81	89.01	82.83	82.83	79.60	78.94	70.14	52.35	30.89	20.79	4.23	2.80	-	0.065	0.072	0.083	0.101	0.065	0.076	0.063
Coulée 2007	73.40	82.59	89.95	90.04	90.04	86.82	86.15	77.35	59.56	38.11	28.00	11.45	10.02	7.21	-	0.057	0.056	0.105	0.047	0.060	0.061
Basse Vallée	58.82	68.02	75.38	81.57	81.57	84.79	85.45	91.92	74.13	52.68	42.58	26.02	24.59	21.79	14.57	-	ns	0.088	0.027	0.032	0.030
Piton l'Entonnoir	47.60	56.80	64.16	70.34	70.34	73.57	74.23	83.03	85.36	63.90	53.80	37.24	35.82	33.01	25.80	11.22	-	ns	0.044	0.042	ns
Bassin Plat	33.87	43.07	50.43	56.61	56.61	59.84	60.50	69.31	87.09	77.63	67.53	50.97	49.54	46.74	39.53	24.95	13.73	-	0.092	0.095	0.087
Ouaki	27.50	36.70	44.06	50.25	50.25	53.47	54.13	62.94	80.73	84.00	73.90	57.34	55.91	53.11	45.89	31.32	20.10	6.37	-	0.034	0.034
Etang du Gol	18.75	27.94	35.30	41.49	41.49	44.72	45.38	54.18	71.97	92.76	82.65	66.10	64.67	61.86	54.65	40.08	28.85	15.12	8.75	-	0.040
Etang Salé	14.08	23.28	30.64	36.82	36.82	40.05	40.71	49.51	67.30	88.76	87.32	70.76	69.34	66.53	59.32	44.74	33.52	19.79	13.42	4.67	-

Table S3: Pairwise geographic distances (below the diagonal) and estimates of *D* (Jost 2008)based on 11 microsatellite markers (above the diagonal).

	St Leu	Ermitage	Petit Bernica	Moulin à Eau	Cambaie	Sans Souci	Rivière des Galets	Grande Chaloupe	Moka	St André	Ste Marguerite	Anse des Cascades	Bois Blanc	Vierge au Parasol	Coulée 2007	Basse Vallée	Piton l'Entonnoir	Bassin Plat	Ouaki	Etang du Gol	Etang Salé
St Leu	-	0.11	0.24	0.12	0.19	0.13	0.14	0.33	0.11	0.07	0.19	0.25	0.10	0.28	0.17	0.08	0.07	0.28	0.12	0.12	0.10
Ermitage	9.20	-	0.21	0.15	0.21	0.13	0.21	0.26	0.20	0.16	0.18	0.29	0.13	0.32	0.26	0.11	0.08	0.41	0.16	0.15	0.12
Petit Bernica	16.56	7.36	-	0.25	0.36	0.24	0.26	0.33	0.23	0.22	0.29	0.39	0.33	0.50	0.46	0.24	0.22	0.45	0.31	0.34	0.36
Moulin à Eau	22.74	13.55	6.19	-	0.25	0.15	0.17	0.25	0.14	0.17	0.31	0.28	0.22	0.36	0.23	0.14	0.17	0.39	0.20	0.20	0.12
Cambaie	22.74	13.55	6.19	0.00	-	0.27	0.37	0.47	0.21	0.23	0.36	0.32	0.25	0.32	0.35	0.18	0.25	0.50	0.22	0.25	0.21
Sans Souci	25.97	16.77	9.41	3.23	3.23	-	0.20	0.37	0.23	0.24	0.31	0.39	0.16	0.48	0.36	0.15	0.18	0.44	0.20	0.20	0.17
Rivière des Galets	26.63	17.43	10.07	3.89	3.89	0.66	-	0.33	0.19	0.08	0.21	0.28	0.29	0.44	0.26	0.21	0.21	0.47	0.27	0.26	0.19
Grande Chaloupe	35.43	26.24	18.88	12.69	12.69	9.47	8.81	-	0.35	0.32	0.37	0.31	0.27	0.54	0.44	0.36	0.29	0.32	0.38	0.39	0.33
Moka	53.22	44.03	36.67	30.48	30.48	27.25	26.59	17.79	-	0.03	0.24	0.20	0.23	0.34	0.18	0.19	0.19	0.43	0.21	0.25	0.20
St André	74.68	65.48	58.12	51.93	51.93	48.71	48.05	39.24	21.45	-	0.20	0.28	0.21	0.39	0.20	0.12	0.13	0.32	0.23	0.25	0.13
Ste Marguerite	84.78	75.59	68.23	62.04	62.04	58.81	58.15	49.35	31.56	10.10	-	0.27	0.19	0.33	0.27	0.25	0.11	0.37	0.25	0.25	0.29
Anse des Cascades	84.84	92.14	84.78	78.59	78.59	75.37	74.71	65.90	48.11	26.66	16.56	-	0.25	0.25	0.22	0.26	0.26	0.40	0.22	0.20	0.30
Bois Blanc	83.42	92.61	86.21	80.02	80.02	76.80	76.14	67.33	49.54	28.09	17.99	1.43	-	0.29	0.27	0.14	0.18	0.22	0.15	0.13	0.15
Vierge au Parasol	80.61	89.81	89.01	82.83	82.83	79.60	78.94	70.14	52.35	30.89	20.79	4.23	2.80	-	0.23	0.27	0.34	0.33	0.24	0.26	0.24
Coulée 2007	73.40	82.59	89.95	90.04	90.04	86.82	86.15	77.35	59.56	38.11	28.00	11.45	10.02	7.21	-	0.20	0.21	0.34	0.17	0.20	0.22
Basse Vallée	58.82	68.02	75.38	81.57	81.57	84.79	85.45	91.92	74.13	52.68	42.58	26.02	24.59	21.79	14.57	-	0.04	0.30	0.10	0.11	0.11
Piton l'Entonnoir	47.60	56.80	64.16	70.34	70.34	73.57	74.23	83.03	85.36	63.90	53.80	37.24	35.82	33.01	25.80	11.22	-	0.33	0.17	0.15	0.12
Bassin Plat	33.87	43.07	50.43	56.61	56.61	59.84	60.50	69.31	87.09	77.63	67.53	50.97	49.54	46.74	39.53	24.95	13.73	-	0.31	0.30	0.30
Ouaki	27.50	36.70	44.06	50.25	50.25	53.47	54.13	62.94	80.73	84.00	73.90	57.34	55.91	53.11	45.89	31.32	20.10	6.37	-	0.11	0.12
Etang du Gol	18.75	27.94	35.30	41.49	41.49	44.72	45.38	54.18	71.97	92.76	82.65	66.10	64.67	61.86	54.65	40.08	28.85	15.12	8.75	-	0.14
Etang Salé	14.08	23.28	30.64	36.82	36.82	40.05	40.71	49.51	67.30	88.76	87.32	70.76	69.34	66.53	59.32	44.74	33.52	19.79	13.42	4.67	-

Discussion

L'objectif principal de cette thèse était de comprendre les processus permettant la mise en place et le maintien de la diversité génétique et phénotypique dans un contexte spatial restreint. Nous avons focalisé nos recherches sur le Zosterops des Mascareignes qui constitue un cas remarquable de diversification à faible échelle spatiale chez les oiseaux (Gill 1973; Milá *et al.* 2010). Le contexte de divergence de l'espèce est relativement bien maîtrisé : des études phylogéographiques et démographiques ont permis de mettre en évidence que la diversification de l'espèce avait très vraisemblablement eu lieu à l'intérieur de l'île au cours des 420 000 dernières années (Warren *et al.* 2006; Bourgeois 2013). En analysant des données microsatellites et phénotypiques (morphologie et couleur du plumage) à plusieurs échelles spatiales, nous avons pu mettre en évidence plusieurs aspects des causes de la diversification de l'espèce et du maintien des zones hybrides entre les différentes formes de couleur de l'espèce. Ces éléments de réponses seront synthétisés dans la première section de cette discussion tandis que la deuxième section sera consacrée à la présentation de quelques perspectives de recherche

1 - Conclusions générales

La diversification des organismes à des échelles spatiales restreintes et notamment au sein d'îles océaniques de petites tailles est controversée (Coyne & Price 2000; Kisel & Barraclough 2010). La caractérisation des processus promouvant la diversification *in situ* est donc primordiale afin de mieux comprendre le rôle de l'écologie, de la géographie et des flux de gènes dans l'apparition de ces phénomènes rares. La dérive génétique et les effets fondateurs ont longtemps été considérés comme les processus principaux d'évolution sur les îles (Barton 1996). Cependant, le rôle de la sélection divergente comme agent de diversification sur les îles a récemment été mis en avant. Dans le chapitre 1, nous avons montré, en analysant les données collectées par Frank Gill avec de nouvelles méthodes d'analyses indisponibles dans les années 1970, que l'évolution des différences phénotypiques au sein de *Z. borbonicus* était vraisemblablement due à l'action de la sélection naturelle. Cette étude apporte donc une nouvelle preuve du rôle de la sélection naturelle comme moteur de la diversification sur les îles à des échelles spatiales restreintes. Néanmoins, l'action conjointe de la sélection et de processus neutres dans l'évolution des différences phénotypiques n'est pas à exclure. En effet, il a été montré que *Z. borbonicus* présentait une dispersion très limitée (Bertrand 2013; Bertrand *et al.* 2014 - Annexe 1) ce qui pourrait favoriser l'action de la

sélection naturelle mais aussi de la dérive génétique. L'action combinée de la sélection naturelle et de la dispersion très faible des individus pourrait être la raison principale pour laquelle la diversification de cette espèce a pris place à une échelle spatiale si réduite.

Dans le chapitre 2, nous nous sommes intéressés aux populations présentes sur un gradient altitudinal court (~15 km de distance linéaire) mais abrupt (~ 2500 m d'altitude). Les gradients altitudinaux sont généralement caractérisés par des changements forts de nombreuses caractéristiques abiotiques (*e.g.* température, pression atmosphérique, radiation UV, etc.) et biotiques (quantité de compétiteurs, prédateurs, etc.) de l'environnement (Körner 2007). Ces zones sont donc souvent considérées comme des laboratoires naturels pour étudier les effets des changements climatiques sur les populations (*e.g.* Colwell *et al.* 2008; Laurance *et al.* 2011). Dans cette étude, nous avons pu montrer qu'un ensemble de traits morphologiques liés à la taille des oiseaux variaient de manière graduelle avec l'altitude. Néanmoins, l'examen de la différentiation génétique neutre sous-jacente nous a permis d'identifier une zone de contact entre deux groupes génétiquement distincts vraisemblablement issue d'un contact secondaire. Dans une revue de littérature sur la diversification sur les gradients altitudinaux tropicaux, Cadena *et al.* (2011) ont montré que les événements de diversification par isolement géographique étaient les plus probables sur les gradients altitudinaux. Chez les oiseaux, la grande majorité des événements de diversification le long des gradients semblent d'ailleurs être le fait de diversifications allopatriques (*e.g.* Fuchs *et al.* 2011; Caro *et al.* 2013). Notre étude ne fait pas office d'exception quant au scénario de divergence mais souligne la nécessité et l'intérêt d'examiner la structure génétique des populations chez les espèces qui présentent de vastes répartitions altitudinales avant de conclure sur les processus en jeux et l'histoire des populations. Sans un examen attentif de la structure génétique des populations, l'existence de deux groupes génétiquement distincts et de cette zone hybride aurait pu passer inaperçue en raison de la convergence phénotypique observée au milieu du gradient altitudinal. Cette étude montre que l'isolement géographique au sein même de l'île de la Réunion a vraisemblablement joué un rôle dans la diversification génétique et phénotypique de *Z.* borbonicus.

Dans les chapitres 3 et 4, nous nous sommes intéressés plus spécifiquement aux zones hybrides présentes entre les différentes formes de couleur. Même si le corpus d'études empiriques et théoriques sur les zones hybrides est bien fourni, certains verrous de connaissances restent en place. Les zones hybrides sur gradients altitudinaux ont globalement été peu étudiées car elles semblent assez rares et/ou peu détectées, notamment chez les

animaux. Toutefois, l'étude de ce type de zones hybrides, où l'environnement change généralement de manière forte sur de courtes distances géographiques, pourrait permettre de distinguer les traits soumis à la sélection exogène de ceux impliqués dans l'isolement reproducteur et nous renseigner sur les bases de l'adaptation à l'altitude ou encore sur le maintien des différences interspécifiques face aux flux de gènes (Abbott & Brennan 2014). Grâce à l'utilisation de modèles de clines sur les données génétiques et phénotypiques collectées sur deux transects altitudinaux traversant la zone hybride décrite dans le chapitre 2, nous avons pu mettre en évidence dans le chapitre 3 que la zone hybride était étroite. Ceci suggère la présence d'un isolement reproducteur partiel entre les deux formes malgré la faible échelle spatiale de ce gradient. De manière frappante, nous avons aussi montré que la coloration du plumage présentait une transition abrupte concordante avec les variations génétiques neutres tandis que les traits morphologiques variaient de façon graduelle en suivant la variation environnementale. Dans le chapitre 4, nous avons étudié les trois zones hybrides de basse altitude. Dans celles-ci, bien que les transitions phénotypiques soient remarquables, car plus abruptes que sur les gradients altitudinaux de l'Ouest, nous n'avons pas trouvé de structuration génétique neutre associée aux différentes formes de couleur.

Les zones hybrides sont de parfaits exemples du caractère semi-perméable des barrières au flux de gènes. Cela se traduit par des variations des taux d'introgression au travers du génome. Dans les deux derniers chapitres de cette thèse nous avons pu montrer, en comparant des marqueurs génétiques neutres et des traits phénotypiques, qu'il existait des patrons de variation très hétérogènes en fonction des traits considérés et de la position de la zone hybride. Sur les gradients altitudinaux, la transition abrupte des marqueurs neutres et de la coloration du plumage suggère qu'il existe des barrières au flux génique qui pourraient être réparties sur l'ensemble du génome. La comparaison des deux catégories de traits phénotypiques, coloration du plumage et morphologie, avec les marqueurs neutres indiquent que la coloration du plumage est probablement impliquée dans l'isolement reproducteur alors que les variables morphologiques seraient soumises à l'action de la sélection exogène. Cette étude confirme l'intérêt des zones hybrides placées sur les gradients altitudinaux pour étudier les patterns de variation de traits soumis à la sélection exogène et de ceux impliqués dans l'isolement reproducteur. La situation est très différente dans les zones hybrides de basse altitude puisque les patrons observés sont vraisemblablement le résultat d'une divergence limitée à un nombre restreint de gènes, laissant les flux de gènes homogénéiser le reste du génome. Le choix de partenaire, basé sur la coloration du plumage, pourrait jouer un rôle fort

dans le maintien de ces zones hybrides. Dans la majorité des zones hybrides, même si les taux d'introgression sont hétérogènes au sein du génome, il apparait que la divergence est répartie sur l'ensemble du génome (Harrison & Larson 2014). Il existe finalement assez peu de cas où la divergence ne concerne qu'une partie du génome (*e.g.* Nadeau *et al.* 2013; Poelstra *et al.* 2014). De plus, ces cas semblent très intéressants car le faible nombre de région divergente du génome limite le nombre de candidat possible lors de la recherche des gènes responsables de l'isolement reproducteur.

L'ensemble de ces résultats ne nous permet pas de statuer sur le scénario exact ayant conduit à la divergence des populations de *Z. borbonicus*. Néanmoins, cette thèse apporte plusieurs éléments de réponses quant aux différents processus impliqués dans la divergence de l'espèce. Nous avons pu démontrer que la sélection naturelle (Chapitre 1, 2 et 3) et l'isolement géographique (Chapitre 2) ont probablement joué un rôle fort dans l'émergence des différentes formes de couleur. De plus, nos résultats suggèrent que le choix de partenaires, probablement basé sur la couleur du plumage, serait impliqué dans l'isolement reproducteur entre les différentes formes (Chapitre 3 et 4). Enfin l'analyse des différentes zones hybrides indique que la divergence est vraisemblablement plus marquée entre les formes de couleur sur les gradients d'altitude qu'entre les formes de basse altitude (Chapitre 3 et 4). La divergence entre les populations de haute et de basse altitude semble aussi concerner plus de régions génomiques que celle entre les formes de basse altitude. Les taux d'introgression sont supposés être fort sur la majeure partie du génome si le contact entre les populations intervient peu après la phase d'isolement géographique, tandis qu'il devrait être réduit s'il intervient longtemps après le début de l'isolement géographique (Wu 2001; Harrison & Larson 2014; Dufresnes *et al.* 2014; Fig. 5). Par conséquent, il pourrait être intéressant de réviser le statut taxonomique des différentes formes de couleur de l'espèce. En effet, les menaces pesant sur la biodiversité à la Réunion sont très hétérogènes spatialement et il conviendrait de tenir compte de la forte diversité génétique et phénotypique existant au sein de l'espèce pour établir les priorités de conservation. En dehors des réponses spécifiques à la différenciation du Zostérops des Mascareignes, ces résultats semblent importants à considérer dans le cadre des événements de divergence à très faibles échelles spatiales.

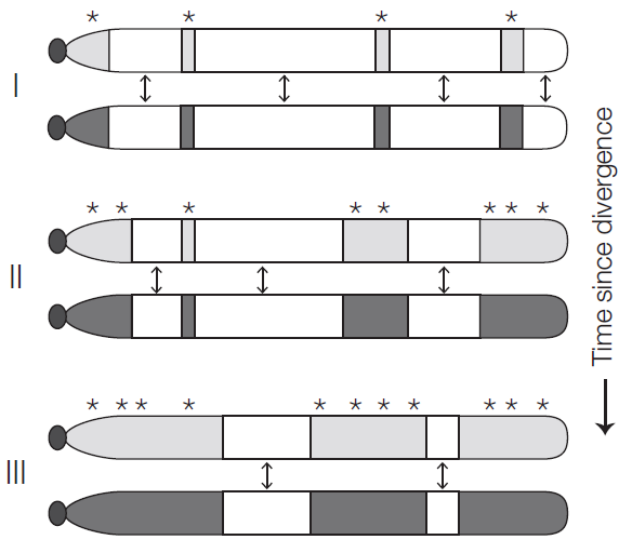


Figure 5 : Accumulation des barrières à la reproduction au cours du temps. Plus la divergence est récente, moins il y a de barrières à la reproduction et plus il y a de flux de gènes dans le reste du génome. Chaque paire de barres représente les chromosomes de taxons frères ayant divergé plus ou moins récemment (du plus récent au plus ancien : I à III). Les astérisques matérialisent les gènes contribuant à l'isolement reproducteur. Tiré de Harrison & Larson (2014) selon l'idée de Wu (2001).

2 - Perspectives

2.1 - Ecologie de l'espèce

Les études menées sur le Zostérops des Mascareignes nt pour le moment principalement utilisé des approches indirectes basées sur l'utilisation de marqueurs génétiques. Ces approches présentent un intérêt fort car elles permettent notamment de mesurer la différenciation génétique entre populations, les taux de migration ou encore les distances de dispersion en capturant seulement quelques individus par population. Néanmoins, elles ne peuvent pas totalement se substituer aux observations de terrain qui permettent de collecter de nombreuses informations sur l'histoire naturelle des espèces. Or, alors que certains éléments seraient vraisemblablement clés pour nous offrir une meilleure compréhension du système, nous connaissons assez mal l'écologie fine de l'espèce.

2.1.1 - Dispersion, socialité et *preferendum* d'habitat

Les analyses de structuration des populations à partir des marqueurs microsatellites ont permis de montrer qu'il existait une structure génétique à très faible échelle spatiale (< 10km; Bertrand *et al.* 2014 - Annexe 1). Cette structuration génétique remarquable peut notamment être expliquée par une faible dispersion des individus. Cette hypothèse semble d'ailleurs être confirmée par l'absence de recapture d'oiseaux marqués entre localités d'échantillonnage malgré leur proximité géographique et les taux de recaptures très élevés sur les sites d'échantillonnage suivis sur le long terme. Les estimations de distance de dispersion parent-enfant qui ont été réalisées avec des données de marqueurs neutres (~200m ; Bertrand 2013) se situent parmi les plus faibles rapportées chez les oiseaux. Cette espèce ne semble pourtant pas présenter de limitations anatomiques au vol et plusieurs observations de terrain montrent qu'elle peut soutenir le vol pendant plusieurs centaines de mètres. Il a été proposé que *Z. borbonicus* pourrait présenter des propensions réduites au vol pour des raisons « psychologiques » (« psychological flightlessness » ; Diamond 1981). Sous cette hypothèse, bien que les oiseaux présentent les capacités anatomiques pour soutenir le vol, ils présentent une réticence à utiliser le vol de longue distance (Komdeur *et al.* 2004). Une autre hypothèse liant les caractéristiques comportementales de l'espèce et sa faible propension à disperser concerne la socialité de l'espèce. Même s'il conviendrait d'étudier plus en détails les comportements de reproduction chez *Z. borbonicus*, il est évident que l'espèce présente toutes les caractéristiques d'une espèce sociale. Les oiseaux se déplacent la majeure partie du temps en groupe et semblent peu territoriaux. Au sein des groupes, des comportements d'« allopreening » (toilettage mutuel) sont régulièrement observés. De plus, il a été montré que des « helpers » (*i.e.* individus apparentés aux parents) participaient au nourrissage des jeunes (Gill 1973). Un dénominateur commun aux « cooperative breeders » comme *Z. borbonicus* semble être une différenciation génétique neutre à faible échelle spatiale (*e.g.* Painter *et al.* 2000; Double *et al.* 2005; Woxvold *et al.* 2006). Le lien entre socialité et structuration génétique des populations est assez mal connu (Painter *et al.* 2000), mais la fidélité au site de nidification, la philopatrie et des faibles distances de dispersion pourraient être à l'origine de la structure génétique à faible échelle observée chez ces espèces. Il serait par conséquent logique de caractériser la philopatrie des individus et le caractère sexe-spécifique de la dispersion afin de mieux comprendre la structuration génétique des populations. Pour ce faire, le suivi à long terme réalisé dans des populations de haute altitude pourrait s'avérer très utile. D'autre part, la structure sociale des populations pourrait impacter

les inférences issues de la génétique des populations (Balloux *et al.* 1998; Lawson Handley & Perrin 2007). Il est par exemple probable que l'échantillonnage d'individus, issus de plusieurs groupes sociaux ou issus d'un seul groupe social au sein d'une même localité, aient un impact sur nos inférences sur la structure des populations.

Une deuxième caractéristique qui influence fortement la structuration génétique des populations et notamment la position des zones hybrides entre les différentes formes de couleur sont les densités de populations (cf. Encadré 1 en introduction). Les zones hybrides de basse altitude sont d'ailleurs localisées sur des discontinuités physiques de l'environnement où la végétation est peu abondante. Bien que les densités n'aient pas été mesurées sur l'ensemble de l'île, il semble très probable qu'elles soient plus faibles au sein des deux rivières (rivière des Galets et St-Etienne) que dans les habitats adjacents du fait de la faible abondance de la végétation dans le lit de ces rivières. La végétation du Grand Brûlé est beaucoup plus abondante et les densités en *Zosterops* pourraient y être importantes. Néanmoins, la succession des coulées de lave éliminant toute végétation sur leurs passages dans cette zone géographique depuis environ 4500 ans a probablement contribué à y limiter l'expansion de l'espèce. Par exemple, entre 1900 et 2007, 12 coulées de laves se sont écoulées dans le Grand Brûlé. Il existe 13 rivières permanentes autour de l'île, il serait intéressant de savoir si elles sont systématiquement associées à des diminutions de densités de population ou si seuls les barrières qui délimitent les aires de répartition des différentes formes de couleur sont concernées. Si toutes les rivières constituent des zones de faible densité, il faudrait déterminer s'il existe une corrélation entre l'âge des rivières et la position des zones hybrides. Enfin, bien qu'il ne semble pas exister de baisse de densité au niveau des zones hybrides altitudinales, il serait logique d'examiner les variations fines de densités des populations le long des différents gradients altitudinaux afin de d'évaluer la possibilité d'une correspondance entre zones hybrides et chute de densité.

2.1.2 - Choix de partenaire : couleur du plumage et chants

Nous avons proposé (Chapitre 3 et 4) que des barrières pré-zygotiques à la reproduction pourraient contribuer au maintien des zones hybrides (et même probablement en être un des agents principaux dans les zones hybrides de basse altitude). Au sein des zones hybrides, les barrières pré-zygotiques résultent généralement de l'homogamie (*i.e.* les partenaires préfèrent un partenaire qui leur ressemble) ou du choix de partenaire en fonction

d'un critère espèce-spécifique (*e.g.* comportement de parade nuptiale). Chez les oiseaux, le choix de partenaire est généralement basé sur des critères visuels (*e.g.* coloration du plumage) ou acoustiques. Par exemple, dans une zone hybride bien décrite entre le Gobemouche noir (*Ficedula hypoleuca*) et le Gobemouche à collier (*Ficedula albicollis*), il existe une forte préférence des femelles pour les mâles conspécifiques qui sont reconnus en fonction de la couleur de leur plumage. Les méthodes pour tester l'homogamie reposent principalement sur des observations comportementales mais peuvent se révéler complexes à mettre en œuvre chez les espèces où l'expérimentation est impossible et les nids difficiles à trouver, comme chez *Z. borbonicus*.

La caractérisation des vocalises et de leur utilisation dans le choix de partenaire sera aussi nécessaire pour comprendre les barrières pré-zygotiques à la reproduction. Nous avons déjà enregistré quelques chants et cris au hasard des missions de terrain et il semble exister des variations entre les différentes localités d'échantillonnage. Cependant, il n'existe pour le moment pas assez d'enregistrements pour éclaircir les patrons de variations entre les différentes formes. Le rôle des chants dans la reconnaissance interspécifique a fait l'objet d'un grand nombre d'études dont les résultats sont très contrastés (Slabbekoorn & Smith 2002). D'une part, les chants ou les réponses aux chants sont parfois plus divergents dans les zones de sympatrie qu'en allopatrie (*e.g.* Kirschel *et al.* 2009). Ces différences marquées aident alors à choisir les partenaires conspécifiques et renforcent les barrières aux flux de gènes. Au contraire, les chants ou les réponses comportementales peuvent converger au sein des zones hybrides (*e.g.* Secondi *et al.* 2003). Dans cette situation, les barrières aux flux de gènes s'en trouvent diminuées car il y a plus de risques de se reproduire avec un partenaire hétérospécifique. Afin de tester ces différentes hypothèses, il conviendrait premièrement d'étudier l'hétérogénéité spatiale des chants en examinant les chants au travers des aires de répartition des différentes formes de couleur et des zones hybrides. Il serait ensuite primordial de caractériser les chants aux alentours des différentes zones hybrides. Enfin dans un deuxième temps, il serait intéressant de réaliser des expériences de repasse des chants afin d'examiner les réponses comportementales.

2.2 - De nouveaux marqueurs génétiques

Les marqueurs microsatellites se sont révélés très utiles afin d'aborder la structuration génétique des populations à différentes échelles spatiales chez *Z. borbonicus* (Bertrand *et al.*

2014 - Annexe 1; Chapitre 2 à 4). Bien que peu nombreux, ces marqueurs nous donnent un aperçu de la différenciation neutre au sein du génome et des barrières présentes entre les différentes formes, notamment si les barrières agissent à l'échelle du génome (Feder *et al.* 2012). Dans un système d'étude comme *Z. borbonicus*, où on trouve une forte structuration géographique des phénotypes, des habitats très variés et de nombreuses zones hybrides (*i.e.* un isolement reproducteur partiel entre les différentes formes), il y a de fortes probabilités que les taux d'introgression entre les différentes formes soient très variables au sein du génome. Par ailleurs, la comparaison des patrons de variations des marqueurs génétiques neutres et des traits phénotypiques au travers des zones hybrides nous a déjà donné un aperçu de cette hétérogénéité. L'avènement des outils génomiques rend actuellement possible l'accès à un grand nombre de marqueurs génétiques répartis sur l'ensemble du génome et ce pour des coûts relativement réduits. Les marqueurs RAD (Restriction site-Associated DNA) développés pendant la thèse de Yann Bourgeois (Bourgeois 2013; Bourgeois *et al.* 2013 - Annexe 2) permettent déjà d'avoir un aperçu des zones du génome qui diffèrent entre formes de couleur. Ces données ont été obtenus pour douze localités : deux par forme de couleur de basse altitude et trois pour la forme de haute altitude. Le séquençage a été fait sur des mélanges d'ADN comprenant les extraits de 18 à 25 individus par localités. Cette approche appelée « pooling » a l'avantage d'être moins coûteuse que les approches individuelles. Les séquences obtenues ont ensuite été replacées sur le génome apparenté du Diamant mandarin (*Taeniopygia guttata*). Ces données ont, pour le moment, principalement servi à identifier les bases génomiques des différences de coloration entre les variantes de couleur de haute altitude (G et HBHB) (Bourgeois 2013). L'analyse de ces données sur l'ensemble des formes de couleur est encore en cours mais les premiers résultats suggèrent qu'il existe plusieurs îlots génomiques de différenciation entre les populations de haute et de basse altitude qui sont répartis à la fois sur les autosomes et les chromosomes sexuels (Bourgeois 2013). Néanmoins, ces données ont été obtenues sur des « pool » d'individus et ne permettent pas d'accéder à l'information individuelle. De plus, il est difficile de tirer des conclusions sur les barrières à la reproduction en se basant seulement sur des populations allopatriques. En effet, ces pics de divergence donnent finalement assez peu d'information sur la nature des effets sélectifs (Bierne *et al.* 2011) et la part relative des différentes régions génomiques dans la réduction du flux de gènes (Payseur 2010). Nous sommes donc actuellement en train de développer des marqueurs de type GBS (Genotyping-by-Sequencing) qui vont nous permettre d'obtenir un grand nombre de données SNPs au niveau individuel. Dans un premier temps, ces marqueurs seront séquencés pour un ensemble de populations au centre des aires de répartition des

formes de couleur et une population de l'espèce sœur *Z. mauritianus*. Dans un deuxième temps, le transect altitudinal du Nord-Ouest de l'île traversant la zone hybride entre les LBHB et la forme d'altitude (G - HBHB) (appelé T1 dans le chapitre 3), et le transect de basse altitude traversant la rivière des Galets (limite entre les LBHB et GHB) seront utilisés. Dans cette sous partie je m'attacherai à la description de quelques perspectives de recherche possibles grâce à l'utilisation de ces nouveaux marqueurs.

2.2.1 - Architecture génomique de la divergence

L'examen attentif des patrons d'introgression sur chacun des marqueurs obtenus sera particulièrement utile afin de comprendre comment la sélection affecte chacun des loci (Payseur 2010). L'utilisation de clines géographiques (cf. Encadré 1 en introduction), qui permet de caractériser les variations spatiales des fréquences alléliques, nous renseignera précisément sur les échelles spatiales en jeu et la force relative de la sélection et de la migration dans les différentes zones hybrides. L'examen des ces patrons pourra aussi se faire à l'aide de clines génomiques (Gompert & Buerkle 2009, 2010). Cette méthode permet de mettre clairement en évidence les patrons d'introgression différentiels au sein du génome et potentiellement d'inférer la nature des effets sélectifs sous jacent (Gompert & Buerkle 2009). Cette approche consiste à faire des régressions entre les fréquences alléliques de chaque locus et un indice hybride calculé sur l'ensemble des loci. En utilisant cette méthode, on s'attend à ce que les loci appartiennent à trois catégories distinctes : les SNPs qui ne s'écartent pas de l'indice hybride global, ceux qui sont significativement plus différenciés que l'indice hybride et enfin ceux qui sont significativement moins différenciés.

Les loci qui ne s'écartent pas de l'indice hybride obtenu avec l'ensemble des loci ne sont vraisemblablement pas soumis à la sélection et reflète l'introgression neutre. Au contraire, les loci qui sont significativement plus différenciés que l'indice hybride sont vraisemblablement sous sélection divergente. La proportion de ces loci dans le génome nous renseignera sur la force de la sélection globale agissant sur le génome. Il s'agira ensuite d'examiner la position de chacun de ces loci afin de déterminer la répartition dans le génome des barrières potentielles à la reproduction. Les fragments d'ADN générés par la méthode GBS sont plus courts que ceux obtenus avec l'approche RAD, et par conséquent l'étape de replacement de ces fragments sur le génome apparenté du Diamant mandarin sera probablement plus compliqué. Néanmoins, la bonne conservation de la synténie chez les

oiseaux (Derjusheva *et al.* 2004) nous permettra vraisemblablement de replacer une bonne partie des marqueurs. Enfin, dans le cas des loci qui sont significativement moins différenciés que la moyenne, la présence d'allèles bénéfiques pour les différentes entités qui s'hybrident est à envisager. En effet, bien que l'importance du passage d'allèles bénéfiques par les zones hybrides (*i.e.* introgression adaptative) soient débattue (Barton 2013; Servedio *et al.* 2013; Abbott *et al.* 2013), il a été démontré dans plusieurs cas que l'hybridation pouvait être une source d'introgression adaptative (Heliconius Genome Consortium 2012; Hedrick 2013). Abbott & Brennan (2014) suggèrent que, dans les zones hybrides localisées sur les gradients altitudinaux, l'introgression adaptative pourrait avoir de l'importance pour permettre aux populations de répondre aux changements climatiques. Néanmoins, les difficultés méthodologiques associées à la détection de ces loci sont nombreuses (*e.g.* variation des taux de recombinaison dans le génome ; Nachman & Payseur 2012). Il apparaît que l'utilisation de méthodes tenant compte de l'histoire des populations seraient nécessaires (Fraïsse *et al.* 2014).

Il sera aussi utile de comparer les données obtenues par la méthode GBS aux résultats des analyses de clines présentées dans les chapitres 3 et 4. Sur le gradient altitudinal de l'Ouest de l'île (zone hybride entre LBHB et la forme de haute altitude HBHB - G) , nous prédisons qu'il y aura une plus grande proportion des loci qui présenteront des clines plus différenciés que l'indice hybride moyen que sur le transect traversant la rivière des Galets (zone hybride entre LBHB et GHB). Dans le cas des zones hybrides de basse altitude (Chapitre 4), nous nous attendons à ce que les loci sélectionnés et participant à l'isolement reproducteur soit regroupés sur une petite portion du génome. Il est d'ailleurs possible que ces loci soient localisés sur les chromosomes sexuels. En effet, il a été montré dans plusieurs zones hybrides chez les oiseaux (*e.g.* Sæther *et al.* 2007; Carling *et al.* 2008; Storchová *et al.* 2010) et chez les mammifères (*e.g.* Macholán *et al.* 2007; Carneiro *et al.* 2014) que l'introgression était particulièrement freinée sur les chromosomes sexuels. Ceci est d'autant plus vraisemblable lorsque les barrières à la reproduction impliquées sont pré-zygotiques, comme escompté entre les formes de basse altitude de *Z. borbonicus*.

2.2.2 - Histoire évolutive

Les premières comparaisons des zones hybrides de basse altitude à celles localisés sur les gradients altitudinaux suggèrent que les barrières à la reproduction seraient plus

nombreuses et mieux réparties dans le génome entre la forme de haute altitude et les formes de basse altitude qu'entre les formes de basse altitude (Chapitre 3 et 4). Ceci restera néanmoins à confirmer grâce à l'utilisation des données génomiques (cf. 2.2.1). Selon Wu (2001), les génomes deviennent généralement de moins en moins perméables au cours du temps à cause de l'accumulation de barrières à la reproduction (Fig. 5). Nous pouvons donc supposer que la divergence altitudinale a précédé la divergence entre les formes de couleur de basse altitude. Plusieurs éléments semblent indiquer que la zone de contact entre la forme de haute altitude et les formes de basse altitude résultent d'un contact secondaire après une phase de divergence allopatrique. Par ailleurs, des analyses préliminaires sur des données récoltées au cours de ma thèse sur deux autres transects altitudinaux à l'Est de l'île traversant la zone hybride entre la forme GHB et celle d'altitude G/HBHB (données microsatellites et phénotypiques non présentées) semblent indiquer le même type de patrons de variation que sur les gradients altitudinaux de l'Ouest (Chapitre 2 et 3). De plus, l'absence d'association entre les phénotypes de basse altitude et l'habitat semble aussi aller à l'encontre d'un scénario de différenciation primaire pour les formes de basse altitude. Dès les années 1970, les scientifiques se sont rendu compte de la difficulté de séparer les scénarios de différentiation primaire des situations de contact secondaire (Endler 1977). L'une des difficultés provient du fait que les histoires de divergence peuvent être complexes avec, par exemple des contacts multiples entre les populations au cours du temps. Depuis les années 2000, de nouvelles méthodes sont en train d'émerger afin de tester ces différents scénarios sur la base de données génétiques. Les modèles de type IM (« Isolation with Migration » ; Nielsen & Wakeley 2001; Hey & Nielsen 2004) ont été les plus populaires afin de tenter de séparer ces deux catégories de scénarios. Plus récemment, la variabilité des taux d'introgression dans le génome, une caractéristique longtemps ignorée dans ce type de modèle, a été intégré dans des modèles de type ABC (« Approximate Bayesian Computation ») afin de tester ces différents de scénarios divergence et d'inférer les âges des différents contact entre les populations (Roux *et al.* 2013, 2014). Ces méthodes prometteuses doivent être préférentiellement appliqués à des marqueurs sous sélection ou liés aux loci sélectionnés qui impriment les événements de divergence sur de plus longues périodes. Ainsi le choix de marqueurs sélectionnés dans le jeu de données GBS sera très utile pour tester différents scénarios de différentiation au sein de l'espèce et estimer l'âge de la divergence entre chaque paire de formes.

Pour conclure, la détermination de la position, du nombre et de l'âge des différents îlots génomiques de différenciation existant entre les chaque paire de formes de couleur

permettra d'éclairer grandement les processus conduisant à la diversification *in situ* et le maintien de zones hybrides à une échelle spatiale extrêmement réduite.

Références

Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Abbott RJ, Brennan AC (2014) Altitudinal gradients, plant hybrid zones and evolutionary novelty. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **369**, 20130346.

Balloux F, Goudet J, Perrin N (1998) Breeding System and Genetic Variance in the Monogamous, Semi-Social Shrew, Crocidura russula. *Evolution*, **52**, 1230.

Barton NH (1996) Natural selection and random genetic drift as causes of evolution on islands. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **351**, 785–794; discussion 795.

Barton NH (2013) Does hybridization influence speciation? *Journal of Evolutionary Biology*, **26**, 267–269.

Bertrand J (2013) Causes de la différenciation génétique à une très petite échelle spatiale chez un oiseau insulaire (zosterops borbonicus). phd Thesis. Université de Toulouse, Université Toulouse III - Paul Sabatier.

Bertrand JAM, Bourgeois YXC, Delahaie B *et al.* (2014) Extremely reduced dispersal and gene flow in an island bird. *Heredity*, **112**, 190–196.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044–2072.

Bourgeois Y (2013) Génétique évolutive d'un cas extrême de polymorphisme de la coloration du plumage chez un oiseau insulaire, Zosterops borbonicus (Zosteropideae). Université Toulouse 3 Paul Sabatier, Toulouse.

Bourgeois YXC, Lhuillier E, Cézard T *et al.* (2013) Mass production of SNP markers in a nonmodel passerine bird through RAD sequencing and contig mapping to the zebra finch genome. *Molecular Ecology Resources*, **13**, 899–907.

Cadena CD, Kozak KH, Gómez JP *et al.* (2011) Latitude, elevational climatic zonation and speciation in New World vertebrates. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20110720.

Carling MD, Brumfield RT, Webster M (2008) Haldane's Rule in an Avian System: Using Cline Theory and Divergence Population Genetics to Test for Differential Introgression of Mitochondrial, Autosomal, and Sex-Linked Loci Across the Passerina Bunting Hybrid Zone. *Evolution*, **62**, 2600–2615.

Carneiro M, Albert FW, Afonso S *et al.* (2014) The Genomic Architecture of Population Divergence between Subspecies of the European Rabbit. *PLoS Genet*, **10**, e1003519.

Caro LM, Caycedo-Rosales PC, Bowie RCK, Slabbekoorn H, Cadena CD (2013) Ecological speciation along an elevational gradient in a tropical passerine bird? *Journal of Evolutionary Biology*, **26**, 357–374.

Colwell RK, Brehm G, Cardelús CL, Gilman AC, Longino JT (2008) Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. *Science (New York, N.Y.)*, **322**, 258–261.

Coyne JA, Price TD (2000) Little evidence for sympatric speciationin island birds. *Evolution*, **54**, 2166–2171.

Derjusheva S, Kurganova A, Habermann F, Gaginskaya E (2004) High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, **12**, 715–723.

Diamond JM (1981) Flightlessness and fear of flying in island species. *Nature*, **293**, 507–508.

Double MC, Peakall R, Beck NR, Cockburn A (2005) Dispersal, philopatry, and infidelity: dissecting local genetic structure in superb fairy-wrens (Malurus cyaneus). *Evolution; International Journal of Organic Evolution*, **59**, 625–635.

Dufresnes C, Bonato L, Novarini N *et al.* (2014) Inferring the degree of incipient speciation in secondary contact zones of closely related lineages of Palearctic green toads (Bufo viridis subgroup). *Heredity*, **113**, 9–20.

Endler JA (1977) *Geographic variation, speciation, and clines*. Princeton University Press, Princeton, New Jersey.

Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics*, **28**, 342–350.

Fraïsse C, Roux C, Welch JJ, Bierne N (2014) Gene flow in a mosaic hybrid zone: is local introgression adaptive? *Genetics*, genetics.114.161380.

Fuchs J, Fjeldså J, Bowie RC (2011) Diversification across an altitudinal gradient in the Tiny Greenbul (Phyllastrephus debilis) from the Eastern Arc Mountains of Africa. *BMC Evolutionary Biology*, **11**, 117.

Gill FB (1973) Intra-island variation in the Mascarene white-eye Zosterops borbonica. *Ornithological Monographs*, iii–66.

Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, **18**, 1207–1224.

Gompert Z, Buerkle AC (2010) introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*, **10**, 378–384.

Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, **105**, 795–809.

Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, **22**, 4606–4618.

Heliconius Genome Consortium (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, **487**, 94–98.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D. persimilis. *Genetics*, **167**, 747–760.

Kirschel ANG, Blumstein DT, Smith TB (2009) Character displacement of song and morphology in African tinkerbirds. *Proceedings of the National Academy of Sciences*, **106**, 8256–8261.

Kisel Y, Barraclough TG (2010) Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, **175**, 316–334.

Komdeur J, Piersma T, Kraaijeveld K, Kraaijeveld-Smit F, Richardson DS (2004) Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there?: Reduced island colonization by Seychelles Warbler. *Ibis*, **146**, 298–302.

Körner C (2007) The use of "altitude" in ecological research. *Trends in Ecology & Evolution*, **22**, 569–574.

Laurance WF, Carolina Useche D, Shoo LP *et al.* (2011) Global warming, elevational ranges and the vulnerability of tropical biota. *Biological Conservation*, **144**, 548–557.

Lawson Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology*, **16**, 1559–1578.

Macholán M, Munclinger P, Sugerková M *et al.* (2007) Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution; International Journal of Organic Evolution*, **61**, 746–771.

Milá B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: Zosterops borbonicus). *BMC Evolutionary Biology*, **10**, 158.

Nachman MW, Payseur BA (2012) Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 409–421.

Nadeau NJ, Martin SH, Kozak KM *et al.* (2013) Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology*, **22**, 814–826.

Nielsen R, Wakeley J (2001) Distinguishing Migration From Isolation: A Markov Chain Monte Carlo Approach. *Genetics*, **158**, 885–896.

Painter JN, Crozier RH, Poiani A, Robertson RJ, Clarke MF (2000) Complex social organization reflects genetic structure and relatedness in the cooperatively breeding bell miner, Manorina melanophrys. *Molecular Ecology*, **9**, 1339–1347.

Payseur BA (2010) Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Molecular Ecology Resources*, **10**, 806–820.

Poelstra JW, Vijay N, Bossu CM *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, **344**, 1410–1414.

Roux C, Fraïsse C, Castric V *et al.* (2014) Can we continue to neglect genomic variation in introgression rates when inferring the history of speciation? A case study in a *Mytilus* hybrid zone. *Journal of Evolutionary Biology*, **27**, 1662–1675.

Roux C, Tsagkogeorga G, Bierne N, Galtier N (2013) Crossing the species barrier: genomic hotspots of introgression between two highly divergent Ciona intestinalis species. *Molecular Biology and Evolution*, mst066.

Sæther SA, Sætre G-P, Borge T *et al.* (2007) Sex Chromosome-Linked Species Recognition and Evolution of Reproductive Isolation in Flycatchers. *Science*, **318**, 95–97.

Secondi J, Bretagnolle V, Compagnon C, Faivre B (2003) Species-specific song convergence in a moving hybrid zone between two passerines. *Biological Journal of the Linnean Society*, **80**, 507–517.

Servedio MR, Hermisson J, van Doorn GS (2013) Hybridization may rarely promote speciation. *Journal of Evolutionary Biology*, **26**, 282–285.

Slabbekoorn H, Smith TB (2002) Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **357**, 493–503.

Storchová R, Reif J, Nachman MW (2010) Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. *Evolution; International Journal of Organic Evolution*, **64**, 456–471.

Warren BH, Bermingham E, Prys-Jones RP, Thébaud C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. *Molecular Ecology*, **15**, 3769–3786.

Woxvold IA, Adcock GJ, Mulder RA (2006) Fine-scale genetic structure and dispersal in cooperatively breeding apostlebirds. *Molecular Ecology*, **15**, 3139–3146.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Annexes

Ces annexes correspondent à deux articles publiés au cours des thèses de Joris Bertrand et Yann Bourgeois sur lesquels j'ai pu collaborer.

Annexe 1 :

Bertrand JAB, Bourgeois YXC, Delahaie B., Duval T., Garcia-Jimenez R, Cornuault J, Heeb P, Milá B, Pujol B. and Thébaud C. 2014. Extremely reduced dispersal and gene flow in an island bird. *Heredity* **112**: 190-196.

Annexe 2 :

Bourgeois, YXC, Lhuilier E, Cézard T. Bertrand JAM, Delahaie B, Cornuault J, Duval T, Bouchez O, Milá B and Thébaud C. 2013. Mass production of SNP markers in a nonmodel passerine bird through RAD sequencing and contig mapping to the zebra finch genome. *Molecular Ecology Resources* **13**: 899-907.

Heredity (2013), 1–7

www.nature.com/hdy

ORIGINAL ARTICLE

Extremely reduced dispersal and gene flow in an island bird

JAM Bertrand[1], YXC Bourgeois[1], B Delahaie[1], T Duval[2], R García-Jiménez[3], J Cornuault[1], P Heeb[1], B Milá[3], B Pujol[1] and C Thébaud[1]

The Réunion grey white-eye, *Zosterops borbonicus*, a passerine bird endemic to Réunion Island in the Mascarene archipelago, represents an extreme case of microgeographical plumage colour variation in birds, with four distinct colour forms occupying different parts of this small island (2512 km^2). To understand whether such population differentiation may reflect low levels of dispersal and gene flow at a very small spatial scale, we examined population structure and gene flow by analysing variation at 11 microsatellite loci among four geographically close localities (<26 km apart) sampled within the distribution range of one of the colour forms, the brown-headed brown form. Our results revealed levels of genetic differentiation that are exceptionally high for birds at such a small spatial scale. This strong population structure appears to reflect low levels of historical and contemporary gene flow among populations, unless very close geographically (<10 km). Thus, we suggest that the Réunion grey white-eye shows an extremely reduced propensity to disperse, which is likely to be related to behavioural processes.
Heredity advance online publication, 2 October 2013; doi:10.1038/hdy.2013.91

Keywords: Mascarene islands; Réunion; white-eye; *Zosterops*; microsatellites; population differentiation

INTRODUCTION

The spatial scale of population differentiation varies widely among organisms, mostly as a result of lineage-specific variation in the potential for dispersal and gene flow (Slatkin, 1987). Thus, organisms with shorter dispersal distances are able to differentiate at smaller spatial scales than those with longer dispersal distances, leading to increased opportunities for allopatric or parapatric divergence within a given area (Kisel and Barraclough, 2010). The exact impact of gene flow on population divergence depends upon several factors, including the spatial context of selection and also the balance between the strength of divergent selection pressures and the between-population migration rates (Endler, 1973; Lenormand, 2002). However, reductions in gene flow between adjacent populations that are sufficient to enable divergence seem more likely at larger than at smaller spatial scales for organisms with a given dispersal ability.

In relatively mobile organisms like birds, the strength of gene flow is thought to retard or prevent differentiation between neighbouring populations to such an extent that geographic barriers to dispersal and long isolation times are often considered necessary for genetic differentiation to take place (Mayr and Diamond, 2001; Price, 2010). Although many empirical studies provide support to this idea, with low or non-significant differentiation found even in bird species distributed over a broad geographic scale (Kekkonen *et al.*, 2011; Procházka *et al.*, 2011, for some recent examples), there are also a few striking cases of phenotypic and genetic differentiation among passerine and some other bird populations at relatively small spatial scales (for example, De Léon *et al.*, 2010; Milá *et al.*, 2010).

While it is tempting to invoke a role for strong divergent selection in the face of gene flow to explain such cases of population divergence at small spatial scales, it is often difficult to rule out the possibility that dispersal is reduced or absent without gene flow data because the ability to disperse at given distances does not easily predict the efficiency of dispersal movements, that is, the realized gene flow (Slatkin, 1987; Mallet, 2001), even in birds. For instance, the Seychelles Warbler (*Acrocephalus sechellensis*) shows locomotory structures similar to those found in closely related species that can sustain flight over long distances, but does not manage to disperse successfully to islands with suitable habitats just outside its distribution range (Komdeur *et al.*, 2004). Thus, to investigate the causes of population divergence at a small spatial scale, it is important to be able to tease apart the effects of gene flow from those due to selection and drift at the relevant spatial scale.

One approach that could be potentially useful relies upon comparing populations of a species in which phenotypic and genetic differentiation occur at a small spatial scale relative to dispersal ability over a range of geographical distances, while minimising the strength of divergent selection pressures by sampling populations experiencing a similar and continuously distributed environment. This should enable, in principle, estimations of dispersal and gene flow independently of the effects of geographic barriers and ecological differences on the patterns of genetic structure.

In this study, we use this approach in combination with genetic indirect approaches to investigate the patterns of population genetic structure at a small spatial scale in the Réunion grey white-eye (*Zosterops borbonicus*). This species complex, endemic to the small island of Réunion (2512 km^2), represents an extreme case of microgeographical variation in birds, with parapatrically distributed plumage colour forms restricted to different parts of the island

[1]Laboratoire Evolution et Diversité Biologique, UMR 5174 Centre National de la Recherche Scientifique (CNRS)—Université Paul Sabatier—Ecole Nationale de Formation Agronomique, Toulouse, France; [2]Société Calédonienne d'Ornithologie Nord, Nouvelle-Calédonie, France and [3]Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, Spanish Research Council (CSIC), Madrid, Spain
Correspondence: JAM Bertrand, Laboratoire Evolution et Diversité Biologique, UMR 5174 Centre National de la Recherche Scientifique (CNRS) and Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse, France.
E-mail: jorisbertrand@gmail.com
Received 8 March 2013; revised 2 July 2013; accepted 15 August 2013

(Gill, 1973; Figure 1). A previous study has shown substantial genetic differentiation among localities distributed across the island, including pairs of localities sampled within the range of the different forms (overall F_{ST} analogue for dominant AFLP markers: $\Phi_{PT} = 0.148$) (Milá *et al.*, 2010). However, little is known about the evolutionary mechanisms underlying phenotypic and genetic divergence, and no direct or indirect measures of dispersal movements are available for this species. Here we aim to test if restriction of gene flow could have played a role in generating the patterns of genetic differentiation, which have been observed at a very small spatial scale in the Réunion grey white-eye. In order to control for the effects of geographic barriers and ecological differences on population differentiation, we obtained estimates of gene flow from measures of genetic differentiation among localities sampled within the distribution range of one of the colour form, the brown-headed brown form (see description in the studies by Gill (1973) and Milá *et al.* (2010)), which is entirely restricted to the lower western slopes of Réunion.

MATERIALS AND METHODS

Bird samples

We sampled a total of 67 individuals at four sites (Figure 1), where the brown-headed brown form of the Réunion grey white-eye is abundant and broadly distributed across the entire area from low to intermediate elevations. All four sites were located in the central part of the form's distribution range in a habitat type classified as semi-dry sclerophyllous forest (Thébaud *et al.*, 2009). They were fairly close to one another (mean = 16.3 km, ranging from 8.8 to 25.2 km), with no obvious physical barriers to gene flow between them. Field procedures and authorizations have been described elsewhere (Milá *et al.*, 2010).

Molecular procedures

We extracted genomic DNA from blood samples using the DNeasy Blood & Tissue Kit (Qiagen, Venlo, Netherlands). All 67 individuals were genotyped at 12 polymorphic microsatellite loci previously isolated in the study species (Bertrand *et al.*, 2012). PCR amplifications were performed in three 10-µl multiplexes (see Supplementary Table S1), each containing ~5–30 ng of DNA, 0.2 mM dNTPs, 0.5 µM of each primer and 0.25 U *Taq* polymerase in 1 × manufacturer's buffer (2 mM $MgCl_2$). PCR thermal profiles were as follows: initial denaturation at 94 °C for 3 min, followed by 35 cycles at 94 °C for 30 s, locus-specific annealing temperature (see Supplementary Appendix 1) for 30 s, 72 °C for 30 s and a final elongation step at 72 °C for 10 min. Fluorescently labelled PCR products were mixed with formamide. Fragment analysis was carried out on an ABI PRISM 3730 DNA analyser (Applied Biosystems, Foster City, CA, USA) with GeneScan-500(LIZ) size standard. Genotyping profiles were scored using GENEMAPPER v.4.0 software (Applied Biosystems).

Basic genetic analyses and within-population diversity

The presence of null alleles was tested with MICRO-CHECKER v.2.2.3 (van Oosterhout *et al.*, 2004) by running 10 000 Monte Carlo simulations and calculating 95% confidence intervals. The probability of null alleles was negligible for all loci except one (Z16), so we excluded this locus from all analyses. We used GENEPOP v4.0 (Rousset, 2008) to test whether each locus significantly deviated from Hardy–Weinberg equilibrium or showed linkage disequilibrium. Genetic diversity was characterised by calculating the mean number of alleles per locus (A), expected and observed heterozygosity (H_E and H_O) and F_{IS} values in GENODIVE v2.0 (Meirmans and Van Tienderen, 2004). The allelic richness (A_R) corrected for sample size was estimated in FSTAT v2.9.3.2 (Goudet, 2001). To account for potential bias due to family sampling, we examined each population sample for the presence of full-sibs with the COLONY v2.0.4.4 programme (Jones and Wang, 2010). The procedure consisted in four 'long runs' with mating system models allowing for polygamy (for both males and females) and inbreeding. We made no further assumption about sibship prior.

Testing for departures from mutation/drift and migration drift equilibriums

Departures from mutation/drift and migration/drift equilibriums might indicate the action of a particular phenomenon (for example, population size variations or restriction in gene flow) so we used two different methods to investigate the likelihood of such departures in our data set. In order to evaluate the possibility for 'recent' bottleneck events, departure from mutation/drift equilibrium was tested by comparing levels of observed and expected heterozygosities with the programme BOTTLENECK v1.2.02 (Cornuet and Luikart, 1996; Piry *et al.*, 1999). As the mutation model underlying our microsatellite markers was uncertain, we considered two alternative mutational models: the Stepwise Mutational Model and the Two-phase Model. As

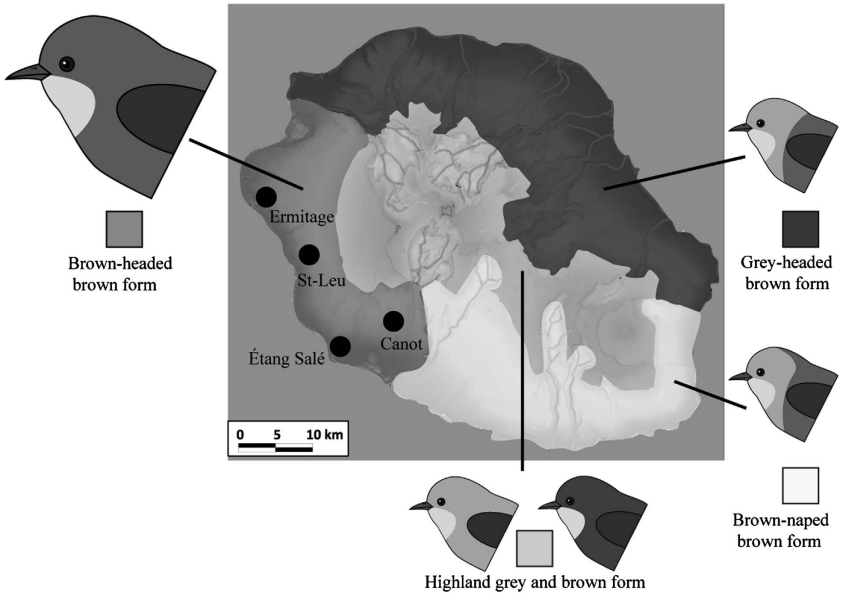


Figure 1 Map of Réunion, with sample localities (black dots) and the geographical distribution of the four Réunion grey white-eye plumage colour forms. The geographical coordinates of sample localities are given in Supplementary Table S5. A full color version of this figure is available at the *Heredity* journal online.

recommended by the authors we assumed 95% of single-step mutation for the latter model (Piry *et al.*, 1999). Calculation was run for 10 000 iterations. Wilcoxon tests were then used to estimate whether potential heterozygote excess or deficit were significantly associated with a recent reduction in effective population size. To appreciate the contribution of neutral genetic drift (associated with a low level of migration) in this system, deviation from migration/drift equilibrium was tested by comparing the relative probabilities of a 'gene flow/drift' and a 'drift only' model with the programme 2MOD (Ciofi *et al.*, 1999). We ran four independent runs setting the MCMC to 1 000 000 iterations (100 000 states burned-in). This method assumes that no allele appeared by mutation since the current population was founded and compares the relative probabilities of the two alternative models.

Population genetic differentiation and gene flow

Data from 11 microsatellite loci were used to estimate levels of genetic differentiation between pairs of populations and among all populations. To obtain indirect estimates of gene flow and compare these estimates with those obtained in other species, we calculated Wright's fixation indices (F_{ST}) using Weir and Cockerham's estimators θ_{ST} (1984) as well as a nearly bias-corrected estimator: $\theta_{RH'}$ (Raufaste and Bonhomme, 2000), which is particularly suited to weakly differentiated populations. Computations were performed in GENETIX v.4.03 (Belkhir *et al.*, 2004). To enable comparisons with the literature, we also computed G_{ST} (Nei, 1987) as well as G''_{ST} (Meirmans and Hedrick, 2011) as implemented in GENODIVE v2.0 (Meirmans and Van Tienderen, 2004). The alternative D_{est} (Jost, 2008) (also implemented in GENODIVE v2.0) was also computed because its assumptions differ from F_{ST} estimators that are under debate in the literature.

We also used the Bayesian multilocus genotyping method implemented in the software BAYESASS 1.3 (Wilson and Rannala, 2003) to detect recent gene flow (over the last several generations) among populations. Three runs of 10 000 000 generations (with a burn-in period of 2 500 000 states) were conducted with all other parameters set to default. This method has been shown to perform well for low migration rate (<33% of migrant individuals per generation) and under moderate genetic differentiation ($F_{ST} \geqslant 0.05$) (Faubet *et al.*, 2007).

Clustering analyses

Bayesian clustering analyses were performed with STRUCTURE 2.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003) to infer the likelihood of $K = 1$–5 populations. MCMC iterations were set to 500 000 (burn-in period was set to 100 000) with 20 replicates per K. The optimal value of K was evaluated by considering the highest mean likelihood value of K, that is, $L(K)$, as well as the ΔK method (Evanno *et al.*, 2005). The optimal alignment of the 20 replicates was determined with the *Greedy* algorithm implemented in CLUMPP (Jakobsson and Rosenberg, 2007). All analyses were run using the admixture model (and correlated allele frequencies), which provides us with estimates of admixture proportions for each individual among clusters. We also used the LOCPRIOR model implemented in STRUCTURE 2.3, as it is more efficient at detecting genetic structure at lower levels of divergence than previous STRUCTURE models. This model makes use of information about sampling locations, but it does not tend to detect any sub-structure when none is present and will ignore sample group information when the ancestry of individuals does not correlate with sampling locations (Hubisz *et al.*, 2009).

RESULTS

Within-population genetic diversity and equilibriums

All microsatellite loci were in Hardy–Weinberg equilibrium. No significant linkage disequilibrium was found across all pairs of loci after correcting for multiple comparisons with the sequential Bonferonni procedure (Rice, 1989). All populations pre-

Table 1 Microsatellite diversity for the four populations across the 11 loci with sample size (n), average number of alleles per locus (A), Allelic richness (A_R) Observed and expected heterozygosity (H_O and H_E), and inbreeding coefficient (F_{IS})

Locality	*Ermitage* (n = 16)				*St–Leu* (n = 21)				*Étang Salé* (n = 20)				*Canot* (n = 10)				*Overall* (n = 67)			
Locus	A	A_R	H_O	H_E	F_{IS}	A	A_R	H_O	H_E	F_{IS}	A	A_R	H_O	H_E	F_{IS}	A	A_R	H_O	H_E	F_{IS}	A	A_R	H_O	H_E	F_{IS}
Z1	5	4.89	0.81	0.78	−0.05	4	3.98	0.71	0.75	0.05	4	3.825	0.75	0.69	−0.09	4	3.90	0.60	0.65	0.08	5	4.305	0.72	0.72	−0.01
Z2	5	4.53	0.81	0.75	−0.08	4	3.42	0.62	0.67	0.08	5	4.624	0.85	0.74	−0.15	4	3.90	0.70	0.70	0.00	6	4.245	0.75	0.72	−0.04
Z3	9	7.60	0.87	0.88	0.01	9	7.78	0.95	0.87	−0.09	8	6.275	0.79	0.79	0.00	7	6.80	1.00	0.86	−0.16	11	7.621	0.90	0.85	−0.06
Z4	11	9.01	0.87	0.91	0.05	11	7.61	0.86	0.86	0.00	12	8.788	0.95	0.89	−0.07	8	7.69	1.00	0.87	−0.15	22	9.912	0.92	0.88	−0.04
Z5	7	6.16	0.75	0.77	0.02	6	4.84	0.76	0.78	0.02	5	4.744	0.90	0.74	−0.22	5	4.80	0.50	0.71	0.29	9	5.761	0.73	0.75	0.03
Z7	12	9.32	0.94	0.91	−0.03	16	11.27	0.91	0.94	0.03	12	9.068	0.80	0.89	0.10	12	11.20	1.00	0.93	−0.07	26	11.46	0.91	0.92	0.01
Z15	11	8.81	0.75	0.89	0.15	15	10.80	0.95	0.93	−0.02	17	12.19	0.90	0.96	0.06	10	10.00	1.00	0.92	−0.08	28	13.04	0.90	0.92	0.03
Z22	5	4.78	0.69	0.76	0.10	4	3.97	0.76	0.74	−0.03	5	4.606	0.70	0.69	−0.01	4	4.00	0.67	0.68	0.02	5	4.773	0.70	0.72	0.02
Z24	6	5.28	0.69	0.76	0.09	7	5.76	0.76	0.81	0.06	10	7.711	0.70	0.88	0.20*	6	5.80	0.80	0.79	−0.01	13	7.00	0.74	0.81	0.09
Z28	4	3.99	0.67	0.73	0.09	6	5.19	0.81	0.77	−0.05	5	4.406	0.70	0.74	0.05	3	3.00	0.60	0.69	0.14	6	4.535	0.69	0.73	0.05
Z31	5	4.87	0.67	0.76	0.12	7	4.96	0.76	0.75	−0.02	5	4.073	0.45	0.61	0.27	6	5.88	0.80	0.76	−0.05	8	5.284	0.67	0.72	0.07
Mean	7.27	6.29	0.77	0.81	0.04	8.09	6.33	0.81	0.81	0.00	8	6.392	0.77	0.78	0.02	6.27	6.09	0.79	0.78	−0.01	12.63	7.085	0.78	0.79	0.01

* indicate a significant value ($P<0.05$).

Table 2 Population-level excess of heterozygote genotypes in the four populations sampled. Wilcoxon-based levels of significance are shown for two mutations model: Stepwise (SMM) and Two-Phase (TPM)

Sample locality	*TPM*	*SMM*
Ermitage	0.61768	0.81738
St–Leu	0.41553	0.61768
Étang Salé	0.20654	0.55078
Canot	0.18262	0.36523

Table 3 Species-level likelihood of 'migration/drift equilibrium' and 'drift only' models for the four runs

Run	*Migration-drift equilibrium model*	*Drift model*	*Bayes factor*
1	20010	159990	7.996
2	20342	159658	7.849
3	20123	159877	7.945
4	20923	159077	7.603

Bayes factors are calculated as ratios of most to least likely models.

sented similar levels of intra-population polymorphism (with A ranging from 6.27–8.09 and A_R ranging from 6.09–6.39) (Table 1). Within-population tests of mutation/drift equilibrium provided no evidence of heterozygote excess ($P>0.18$) for both the Two-Phase model and the Stepwise Mutational model (Table 2). Bayes factors comparing the models of 'pure drift' versus 'gene flow-drift' models were close to 8, strongly supporting the 'pure drift' over the 'gene flow/drift' model (Table 3). This result alone suggests that gene flow is significantly restricted between populations.

Among-population genetic differentiation and clustering

The programme COLONY 2.0.4.4 identified two fullsib pairs in one locality (St–Leu) and a group of four individuals in a second locality (Canot). We ran the analyses twice with either the full data set or with a reduced data set from which one randomly chosen individual from each fullsib pair and three randomly chosen individuals within the group of four were excluded. The results we obtained in the two sets of analyses are qualitatively and quantitatively similar (see Table 4 and Supplementary Table S3 and S4), and so we present only the results using the full data set. Overall estimates of among-population genetic differentiation were low but significant for all indices ($\theta_{ST} \approx G_{ST} = 0.03$, $\theta_{RH'} = 0.22$, $G''_{ST} = 0.14$, $D_{est} = 0.12$, $P<0.05$) (see Supplementary Table S1). All pairwise comparisons were also significant (Table 4), independent of the geographic distance between sample localities. Although there were differences in the magnitude of genetic differentiation between indices, the global pattern of pairwise differentiation was consistent across indices (Table 4). We also found that the most likely number of genetic clusters is three ($K=3$, $L(K) = -2902.635$, $\Delta K = 9.640$), with all individuals unambiguously assigned to the different clusters (Figure 2, see also Supplementary Table S2). The first two clusters matched exactly two sample localities, whereas the third one consisted of individuals from the two sample localities in close geographic proximity to one another. Thus, sample localities are broadly differentiated from one another at a very small spatial scale ($<100\,km^2$), likely as a result of limited gene exchange among them. However, even at that scale, reductions in gene flow between adjacent sample localities appear more likely at larger than at smaller spatial scales.

Table 4 Pairwise geographic distances and estimates of genetic differentiation based on 11 microsatellite markers surveyed in four sample localities

	St–Leu	*Étang Salé*	*Canot*	*Overall*
Geographic distances (km)				
Ermitage	9.6	23.1	25.2	—
St–Leu	—	14.3	15.7	—
Étang Salé	—	—	8.8	—
θ_{ST}				
Ermitage	**0.024**	**0.030**	**0.042**	**0.030**
St–Leu	—	**0.025**	**0.028**	—
Étang Salé	—	—	**0.026**	—
$\theta_{RH'}$				
Ermitage	**0.193**	**0.151**	**0.293**	**0.219**
St–Leu	—	**0.190**	**0.222**	—
Étang Salé	—	—	**0.130**	—
G_{ST}				
Ermitage	**0.024**	**0.030**	**0.042**	**0.030**
St–Leu	—	**0.025**	**0.028**	—
Étang Salé	—	—	**0.026**	—
G''_{ST}				
Ermitage	0.124	0.144	0.208	**0.143**
St–Leu	—	0.124	0.139	
Étang Salé	—	—	0.121	
	—	—	—	
Jost's D_{est}				
Ermitage	0.103	0.118	0.172	**0.117**
St–Leu	—	0.101	0.114	—
Étang Salé	—	—	0.097	—

Various indices are θ_{ST} (Weir and Cockerham, 1984), $\theta_{RH'}$ (Raufaste and Bonhomme, 2000), G_{ST} (Nei, 1987), G''_{ST} (Meirmans and Hedrick, 2011) and Jost's D (Jost, 2008). Figures in bold face indicate significance at $P<0.05$.

Contemporary gene flow

Most sample localities have low proportion of migrant (<9%) and high proportions of non-migrant (>88%) individuals per generation, with the exception of the sites situated at close geographic proximity to its nearest neighbour (8.8 km), which seems to have a large expected proportion of migrants (22%). We found no evidence of asymmetry in migration between adjacent localities and among all localities (Table 5). Thus, these results are consistent with the idea that gene flow and dispersal are extremely reduced between adjacent localities, unless very close (<10 km).

DISCUSSION

Levels of population genetic differentiation implies reduced dispersal and gene flow

We found a significant signal of population differentiation across four sample localities situated within the distribution range of the brown-headed brown form of the Réunion grey white-eye. Accordingly, we demonstrated the presence of population structure by differentiating three clusters of individuals within the data set, with the two sample localities situated within 10 km of each other forming one of the clusters. This shows that population structuring in the Réunion grey white-eye can occur at a scale of 10–20 km even in a broadly

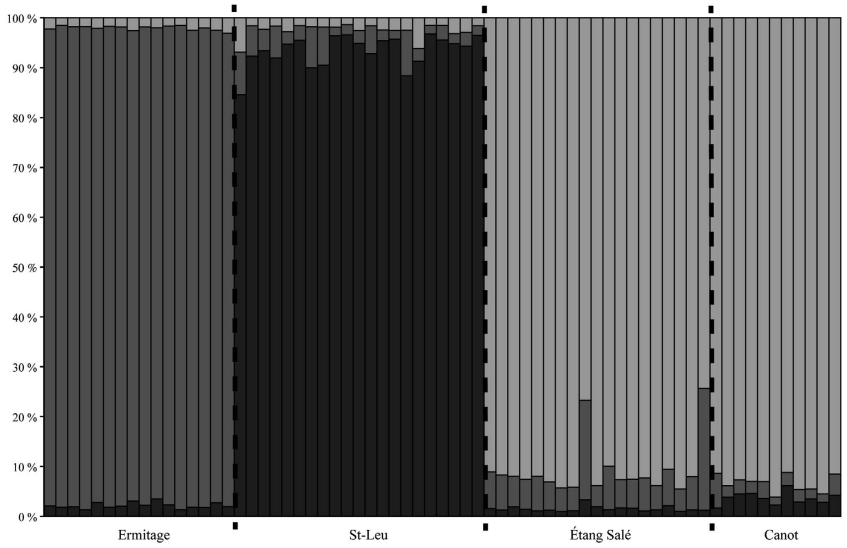


Figure 2 Admixture proportions as inferred from genetic clustering. Each bar represents an individual. Each colour reflects the likelihood of belonging to one of the inferred genetic clusters. Dashed lines delimit sample localities.

Table 5 Mean estimates of the distribution of recent migration rates (*m*) calculated using BAYESASS and given as the proportion of migrant individuals per population per generation

Ermitage	*St–Leu*	*Étang Salé*	*Canot*
0.881	0.018	0.008	0.045
0.017	**0.952**	0.008	0.028
0.089	0.024	**0.974**	0.224
0.012	0.008	0.011	**0.701**

Columns represent the incoming migration rates and rows represent outgoing migration rates. Bold values represent the proportion of non-migrant individuals in a population.

continuous habitat. The levels of genetic differentiation observed in this study are unexpectedly high, given the spatial scale considered. While comparing the magnitude of genetic differentiation found in different studies is not devoid of problems (for example, Meirmans and Hedrick, 2011; Whitlock, 2011), our estimates ($\theta_{ST} \approx G_{ST} = 0.03 =$, $\theta_{RH'} = 0.22$, $G''_{ST} = 0.14$, $D_{est} = 0.12$) are close to average F_{ST} values reported in previous studies (0.049 in the study by Barrowclough (1983) and 0.048 in the study by Evans (1987)). It is also striking that these estimates are comparable to those found at the scale of much larger regions in other passerines such as the House Sparrow (*Passer domesticus*) ($F_{ST} = 0.004$ across Finland; Kekkonen *et al.*, 2011) or the Eurasian Reed Warblers (*Acrocephalus scirpaceus*) ($F_{ST} = 0.013$, $G'_{ST} = 0.078$, $D_{est} = 0.063$ across Europe; Procházka *et al.*, 2011). Thus, we suggest that the levels of genetic differentiation found in our study are not just unexpected, they are exceptionally high for birds at such a small spatial scale.

Both significant levels of genetic differentiation and population structure are therefore consistent with extremely reduced gene flow between populations. It is noteworthy that our comparison of 'pure drift' and the 'gene flow/drift' models also supports the idea that drift is prevalent relative to migration in explaining differentiation. Our analysis of contemporary gene flow also suggests that current levels of gene flow among populations are very low, in agreement with our suggestion that historical gene flow must have been extremely reduced to explain present-day patterns of genetic differentiation.

Genetic differentiation with no geographic and ecological transitions

We compared sample localities that are very close to one another, occupy a similar habitat type and are not separated by any obvious physical or ecological discontinuities. Thus, our results likely reflect reduced dispersal and gene flow at a very small spatial scale, independently of the effects of divergent selection pressures and geographic barriers to gene flow.

There are several explanations for such an unusual pattern of genetic differentiation at a small scale. First, we assumed that the populations were experiencing similar environments, using habitat type as a proxy. Environmental factors, including for example, altitude, temperature or rainfall, often vary within a habitat type, especially on islands with a very rugged topography (Whittaker and Fernandez-Palacios, 2007). In addition, biotic factors such as parasites and pathogens may also vary between localities within a same habitat type. Populations could then diverge in response to heterogeneous natural selection, and this may keep gene flow at low levels if immigrants have reduced fitness relative to residents. This was found in wild populations of great tits (*Parus major*) separated by distances of <3 km (Garant *et al.*, 2005; Shapiro *et al.*, 2006; Björklund *et al.*, 2010), and also in song sparrows (*Melospiza melodia*) with five subspecies coexisting in a restricted area made of various micro-habitats (Chan and Arcese, 2003). However, the latter results were not associated with significant neutral genetic differentiation, which suggests that genetic drift was not strong enough to generate neutral genetic differentiation (but also refer to the studies by Senar *et al.* (2006), Lee *et al.* (2010) and Rutz *et al.* (2012)). There is no evidence from previous studies that populations belonging to one particular colour form of the Réunion grey white-eye and being as geographically close as those used in this study show any sign of niche differentiation (Gill, 1971, 1973), suggesting that this explanation is unlikely.

Subtle, undetected geographic barriers could also account for reduced gene flow between populations. However, this seems also unlikely as effective geographic barriers for the Réunion grey white-eye are conspicuous physiographic features such as major river beds and extensive lava flows (Gill, 1973; Milá *et al.*, 2010), none of which

occur in our study area. A more plausible explanation is that the Réunion grey white-eye, like some other island birds including several species of white-eyes, could show a reduced propensity to disperse, perhaps as a result of selection against long-distance dispersal (Komdeur *et al.*, 2004; Moyle *et al.*, 2009). As the species is clearly capable of sustained flight over hundreds of metres of land, it may express the phenomenon of 'behavioural flightlessness', that is, a behavioural reluctance to move away from its source locality (Diamond, 1981). Very high recapture rates on long-term study sites could be consistent with this idea (Milá and Thébaud, unpublished data).

The extent to which social behaviour could also influence the spatial genetic structure of populations is largely unknown (Painter *et al.*, 2000), but substantial levels of genetic differentiation was found at relatively small spatial scales in lekking (Höglund and Shorey, 2003; Bouzat and Johnson, 2004) or cooperative breeding bird species (Painter *et al.*, 2000; Double *et al.*, 2005; Temple *et al.*, 2006; Woxvold *et al.*, 2006). Strong social behaviours such as allopreening, huddling and cooperative breeding are common in the Réunion grey white-eye, with no apparent territorial behaviour throughout the breeding season (Gill, 1971; Gill, 1973), and may further contribute to reducing gene exchange among populations located at very short distances (for example, through social structuring and/or strong philopatry). Clearly, more work needs to be done to understand if such behavioural processes can be associated with reductions in gene flow among populations.

Another possible explanation relies on the idea that vocal micro-geographic variation in the form of song 'dialects' may easily arise in birds that learn their songs (Catchpole and Slater 2008), such as white-eyes (Baker 2012). This could potentially contribute to reductions in gene flow among geographically close populations. If neighbouring populations differ in their song types, with young males and females preferentially learning local song types and then, later in life, preferring these songs while discriminating against non-local variants, then interpopulation matings could be reduced relative to intrapopulation matings, causing a restriction in gene flow (MacDougall-Shackleton and MacDougall-Shackleton, 2001). This seems especially likely in resident species, but it will ultimately depend upon natal dispersal distances. In the case of the Réunion grey white-eye, whether vocal dialects may contribute to population differentiation has yet to be tested.

We have shown that populations of the Réunion grey white-eye can exhibit spatial genetic structure and differentiation at a very small scale ($<100\,km^2$), even in the absence of any obvious geographic barrier and/or change in habitat attributes. This strong population structure appears to reflect low levels of historical and contemporary gene flow among populations, unless very close geographically ($<10\,km$). Thus, the Réunion grey white-eye seems to show an extremely reduced propensity to disperse, which is likely to be related to behavioural processes, because the birds show no sign of wing reduction or of a reduced power of flight. Besides the fact that the pattern seen here reveals levels of genetic differentiation at a small spatial scale, which are exceptionally high for birds that are good flyers, our findings also have implications for how the different colour forms found in the Réunion grey white-eye have likely been shaped by the interplay of natural selection, genetic drift and reduced gene flow.

DATA ARCHIVING

Data deposited in the Dryad repository: doi:10.5061/dryad.1652m.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the editor and two anonymous referees for their insightful and constructive comments that improved an earlier version of this manuscript. Fieldwork was facilitated by the outstanding efforts of Guillaume Gélinaud, Dominique Strasberg, Juli Broggi, Ben Warren, Magali Thierry, René-Claude Billot, Jean-Michel Probst, Isabelle Henry and Vincent Leconte. We gratefully thank the Réunion National Park for permission to conduct fieldwork. Marc Salamolard and Benoît Lequette provided valuable help with fieldwork and logistics. JB, YB, BD and JC were supported by MESR (Ministère de l'Enseignement Supérieur et de la Recherche) PhD scholarships. The research was supported by Agence Française pour le Développement grants to CT, the Fondation pour la Recherche sur la Biodiversité (FRB) through its Centre for Synthesis and Analysis of Biodiversity (CESAB), the 'Laboratoire d'Excellence' TULIP (ANR-10-LABX-41) and the SYNTHESYS Project (http://www.synthesys.info/), which was funded by the European Community Research Infrastructure Action under the FP7 'Capacities' Programme at the Museo Nacional de Ciencias Naturales (CSIC) of Madrid, Spain.

Baker MC (2012). Silvereyes (*Zosterops lateralis*) song differentiation in an island-mainland comparison: analyses of a complex cultural trait. *Wilson J Ornithol* **124**: 454–466.
Barrowclough GF (1983). Biochemical studies of microevolutionary processes. In: Brush AH, Clark GA Jr (eds) *Perspectives in Ornithology*. University of Cambridge Press: Cambridge, UK pp 223–261.
Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire Génome Populations Interactions CNRS UMR5* **117**: 5000.
Bertrand JAM, García-Jiménez R, Bourgeois Y, Duval T, Heeb P, Thébaud C *et al.* (2012). Isolation and characterization of twelve polymorphic microsatellite loci for investigating an extreme case of microgeographical variation in an island bird (*Zosterops borbonicus*). *Conservation Genet Res* **4**: 323–326.
Björklund M, Ruiz I, Senar JC (2010). Genetic differentiation in the urban habitat: the great tits (*Parus major*) of the parks of Barcelona city. *Biol J Linn Soc* **99**: 9–19.
Bouzat J, Johnson K (2004). Genetic structure among closely spaced leks in a peripheral population of lesser prairie-chicken. *Mol Ecol* **13**: 499–505.
Catchpole CK, Slater PJB (2008). *Bird Songs: Biological Themes and Variations*. Cambridge University Press: Cambridge, UK.
Chan Y, Arcese P (2003). Morphological and microsatellite differentiation in *Melospiza melodia* (Aves) at a microgeographic scale. *J Evol Biol* **16**: 939–947.
Ciofi C, Beaumont MA, Swingland I R, Bruford MW (1999). Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. *Proc R Soc B Biol Sci* **266**: 2269–2274.
Cornuet JM, Luikart G (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.
De Léon LF, Bermingham E, Podos J, Hendry AP (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.
Diamond JM (1981). Flightlessness and fear of flying in island species. *Nature* **293**: 507–508.
Double MC, Peakall R, Beck NR, Cockburn A (2005). Dispersal, philopatry, and infidelity: dissecting local genetic structure in superb fairy-wrens (*Malurus Cyaneus*). *Evolution* **59**: 625–635.
Endler JA (1973). Gene flow and population differentiation. *Science* **179**: 243–250.
Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.
Evans PGH (1987). Electrophoretic variability of gene products. In: Cooke F, Buckley PA (eds) *Avian Genetics*. Academic Press: New York, NY, USA, pp 105–162.
Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.
Faubet P, Waples RS, Gaggiotti OE (2007). Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Mol Ecol* **16**: 1149–1166.
Garant D, Kruuk LEB, Wilkin TA, McCleery RH, Sheldon BC (2005). Evolution driven by differential dispersal within a wild bird population. *Nature* **433**: 60–65.
Gill FB (1971). Ecology and evolution of the sympatric Mascarene white-eyes, *Zosterops borbonica* and *Zosterops olivacea*. *Auk* **88**: 35–60.
Gill FB (1973). Intra-island variation in the Mascarene White-eye *Zosterops borbonica*. *Ornithol Monogr* **12**: 1–66.
Goudet J (2001). FSTAT, a program to estimate and test gene diversities and fixation indices. Version 2: 3. *J Heredity* **86**: 485–486.

Höglund J, Shorey L (2003). Local genetic structure in a white-bearded manakin population. *Mol Ecol* **12**: 2457–2463.
Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Res* **9**: 1322–1332.
Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.
Jones OR, Wang J (2010). COLONY: a program for parentage and sibship inference from multilocus genotypes data. *Mol Ecol Res* **10**: 551–555.
Jost L (2008). G_{ST} and its relatives do not measure differentiation. *Mol Ecol* **17**: 4015–4026.
Kekkonen J, Seppä P, Hanski IK, Jensen H, Väisänen RA, Brommer JE (2011). Low genetic differentiation in a sedentary bird: house sparrow population genetics in a contiguous landscape. *Heredity* **106**: 183–190.
Kisel Y, Barraclough TG (2010). Speciation has a spatial scale that depends on level of gene flow. *Am Nat* **175**: 316–334.
Komdeur J, Piersma T, Kraaijeveld K, Kraaijeveld-Smit F, Richardson DS (2004). Why Seychelles warblers fail to recolonize nearby islands: unwilling or unable to fly there? *Ibis* **146**: 298–302.
Lee JW, Simeoni M, Burke T, Hatchwell BJ (2010). The consequences of winter flock demography for genetic structure and inbreeding risk in vinous-throated parrotbills, *Paradoxornis webbianus*. *Heredity* **104**: 472–481.
Lenormand T (2002). Gene flow and the limit to natural selection. *Trends Ecol Evol* **17**: 183–189.
MacDougall-Shackleton EA, MacDougall-Shackleton SA (2001). Cultural and genetic evolution in mountain White-crowned Sparrows: song dialects are associated with population structure. *Evolution* 55: 2568–2575.
Mallet J (2001). Gene flow. In: Woiwod IP, Reynolds DR, Thomas CD (eds) *Insect Movement: Mechanisms and Consequences*. CAB Internationalpp 337–360.
Mayr E, Diamond JM (2001). *The Birds of Northern Melanesia: Speciation, Ecology, and Biogeography*. Oxford University Press: Oxford, UK.
Meirmans PG, Hedrick PW (2011). Assessing population structure: F_{ST} and related measures. *Mol Ecol Res* **11**: 5–18.
Meirmans PG, Van Tienderen PH (2004). Genotype and Genodive: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* **4**: 792–794.
Milá B, Warren BH, Heeb P, Thébaud C (2010). The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evol Biol* **10**: 158.
Moyle RG, Filardi CE, Smith CE, Diamond J (2009). Explosive Pleistocene diversification and hemispheric expansion of a 'great speciator'. *Proc Natl Acad Sci USA* **106**: 1863–1868.
Nei M (1987). *Molecular Evolutionary Genetics,* MacIntyre RJ (eds) Columbia University Press: New York, NY, USA.
Painter J, Crozier R, Poiani A, Robertson R, Clarke M (2000). Complex social organization reflects genetic structure and relatedness in the cooperatively breeding bell miner, *Manorina melanophrys*. *Mol Ecol* **9**: 1339–1347.
Piry S, Luikart G, Cornuet JM (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* **90**: 502–503.
Price TD (2010). *Speciation in Birds*. Roberts and Company Publishers: Greenwood Village, CO, UK.
Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
Procházka P, Stokke BG, Jensen H, Fainová D, Bellinvia E, Fossøy F *et al*. (2011). Low genetic differentiation among reed warbler *Acrocephalus scirpaceus* populations across Europe. *J Avian Biol* **42**: 103–113.
Raufaste N, Bonhomme F (2000). Properties of bias and variance of two multiallelic estimators of F_{ST}. *Theor Popul Biol* **57**: 285–296.
Rice WR (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
Rousset F (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Res* **8**: 103–106.
Rutz C, Ryder TB, Fleischer RC (2012). Restricted gene flow and fine-scale population structuring in tool using New Caledonian crows. *Naturwissenschaften* **99**: 313–320.
Senar JC, Borras A, Cabrera J, Cabrera T, Björklund M (2006). Local differentiation in the presence of gene flow in the citril finch *Serinus citrinella*. *Biol Lett* **2**: 85–87.
Shapiro BJ, Garant D, Wilkin TA, Sheldon BC (2006). An experimental test of the causes of small-scale phenotypic differentiation in a population of great tits. *J Evol Biol* **19**: 176–183.
Slatkin M (1987). Gene flow and the geographic structure of natural populations. *Science* **236**: 787–792.
Thébaud C, Strasberg D, Warren BH, Cheke A (2009). Mascarene islands, biology. In: Gillespie RG, Clague DA (eds) *Encyclopedia of Islands*. University of California Press: Berkeley, CA, USA pp 612–619.
Temple HJ, Hoffman JI, Amos W (2006). Dispersal, philopatry and intergroup relatedness: fine-scale genetic structure in the white-breasted thrasher, *Ramphocinclus brachyurus*. *Mol Ecol* **15**: 3449–3458.
Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**: 535–538.
Weir B, Cockerham CC (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
Whitlock MC (2011). G'_{ST} and *D* do not replace F_{ST}. *Mol Ecol* **20**: 1083–1091.
Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**: 1177–1191.
Whittaker RJ, Fernandez-Palacios JM (2007). *Island biogeography: ecology, evolution and conservation*. Oxford University Press: Oxford, UK.
Woxvold IA, Adcock GJ, Mulder RA (2006). Fine-scale genetic structure and dispersal in cooperatively breeding apostlebirds. *Mol Ecol* **15**: 3139–3146.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2013) **13**, 899–907 doi: 10.1111/1755-0998.12137

Mass production of SNP markers in a nonmodel passerine bird through RAD sequencing and contig mapping to the zebra finch genome

YANN X. C. BOURGEOIS,* EMELINE LHUILLIER,†‡ TIMOTHÉE CÉZARD,§ JORIS A. M. BERTRAND,* BORIS DELAHAIE,* JOSSELIN CORNUAULT,* THOMAS DUVAL,¶ OLIVIER BOUCHEZ,‡** BORJA MILÁ†† and CHRISTOPHE THÉBAUD*

**Laboratoire Évolution et Diversité Biologique, UMR 5174 CNRS - Université Paul Sabatier – ENFA, 118 route de Narbonne, Bâtiment 4R1, F-31062 Toulouse Cedex 9, France, †INRA, UAR 1209 Département de Génétique Animale, INRA Auzeville, F-31326 Castanet-Tolosan, France, ‡GeT-PlaGe, Genotoul, INRA Auzeville, F-31326 Castanet-Tolosan, France, §The GenePool, Ashworth Laboratories, The University of Edinburgh, The King's Building, Edinburgh EH9 3JT, UK, ¶Société Calédonienne d'Ornithologie Nord, BP 236, F-98822 Poindimié, Nouvelle Calédonie, France, **INRA, UMR 444 Laboratoire de Génétique Cellulaire, INRA Auzeville, F-31326 Castanet-Tolosan, France, ††Museo Nacional de Ciencias Naturales, CSIC, José Gutiérrez Abascal 2, Madrid 28006, Spain*

Abstract

Here, we present an adaptation of restriction-site-associated DNA sequencing (RAD-seq) to the Illumina HiSeq2000 technology that we used to produce SNP markers in very large quantities at low cost per unit in the Réunion grey white-eye (*Zosterops borbonicus*), a nonmodel passerine bird species with no reference genome. We sequenced a set of six pools of 18–25 individuals using a single sequencing lane. This allowed us to build around 600 000 contigs, among which at least 386 000 could be mapped to the zebra finch (*Taeniopygia guttata*) genome. This yielded more than 80 000 SNPs that could be mapped unambiguously and are evenly distributed across the genome. Thus, our approach provides a good illustration of the high potential of paired-end RAD sequencing of pooled DNA samples combined with comparative assembly to the zebra finch genome to build large contigs and characterize vast numbers of informative SNPs in nonmodel passerine bird species in a very efficient and cost-effective way.

Keywords: next-generation sequencing, passerine, pooled DNA, SNP detection, zebra finch genome, *Zosterops*

Received 26 March 2013; revision received 24 May 2013; accepted 4 June 2013

Introduction

The rapid development of new sequencing technologies has brought much hope to identify the allelic variants that underlie phenotypic variation and divergence in natural populations of nonmodel species (Davey *et al.* 2011; Radwan & Babik 2012; but see Bierne *et al.* 2011; Rockman 2012; Travisano & Shaw 2013). However, unravelling the molecular causes of phenotypic changes, reconstructing the demographic histories of multiple populations or quantifying fine-scale gene flow require genome-wide sequence data from multiple individuals, something which remains difficult to achieve for most nonmodel species. Strategies based on genome reduction are thus of great interest when attempting to link genetic and phenotypic variants, as they can provide substantial population genomic data for a large number of individuals at a reasonable cost (Davey *et al.* 2011).

The first protocols of genome reduction have included the development of Reduced Representation Libraries (RRLs) or RNA sequencing (Altshuler *et al.* 2000; Wang *et al.* 2009). One problem with the RRL approach is that it usually requires extensive testing prior to identifying orthologous single-copy loci that can be compared between different individuals or populations (van Tassell *et al.* 2008; van Bers *et al.* 2010). RNA sequencing can yield much information about coding mutations and relative expression levels, especially in combination with other approaches (Wang *et al.* 2009; Hawkins *et al.* 2010). However, it can be difficult to implement in nonmodel organisms because obtaining and preserving RNA can be challenging, especially in field conditions. It also excludes noncoding regions that can be of interest.

In this context, Restriction-site-Associated DNA sequencing or RAD sequencing has become a method of

Correspondence: Yann X. C. Bourgeois, Fax: +33 (0)5 61 55 73 27; E-mail: yann.x.c.bourgeois@gmail.com

choice for high-density single-nucleotide polymorphism (SNP) discovery and genotyping across many individuals in many populations (Davey & Blaxter 2010). RAD sequencing allows reducing genome complexity by sequencing the same loci across the genome in several individuals, facilitating among-individual comparisons and limiting sequencing investment. The approach consists in cleaving double-stranded genomic DNA with a restriction enzyme chosen to obtain an appropriate sequencing depth. Then, DNA fragments are randomly sheared to a specific length that varies depending on which next-generation sequencing (NGS) platform is used. Using specific adapters, it is thus possible to selectively amplify the regions flanking the restriction sites. Paired-end reads for each RAD tag can then be assembled into long contiguous sequences (Hohenlohe *et al.* 2011). While RAD sequencing was initially designed for microarray (Miller *et al.* 2007), it has been quickly adapted to NGS technology (Baird *et al.* 2008) and has opened up a wide range of applications in evolutionary genomics (e.g. Emerson *et al.* 2010; Gagnaire *et al.* 2012; Hess *et al.* 2012; Keller *et al.* 2012; Takahashi *et al.* 2012; Wang *et al.* 2012), most notably in species that have a reference genome.

Here, we present an efficient and cost-effective protocol for developing RAD markers by adapting previous protocols (Baird *et al.* 2008) to the Illumina Hiseq2000 technology, which allows the sequencing of 150 to 180 million paired-end reads per lane for a reduced price per base pair, compared with the 40 million produced by a Genome Analyzer IIx. Technological advances now allow sequencing at substantial depth tens to hundreds of libraries per sequencer run (Davey & Blaxter 2010) and have been shown to be of interest in most recent studies using RAD sequencing (Davey *et al.* 2012; Keller *et al.* 2012; Peterson *et al.* 2012; Wagner *et al.* 2013).

The present study aimed at generating a very large number of SNP markers in a small passerine bird endemic to Réunion (Mascarene Islands, southwestern Indian Ocean), the Réunion grey white-eye (*Zosterops borbonicus*). To deal with small genomic DNA quantities that are typically obtained from field-collected blood samples, while estimating allele frequencies at a genome-wide scale and keeping cost as low as possible, we sequenced multiple libraries, each corresponding to pools of individuals. Pooling has been shown to be a very cost-effective approach to estimate allele frequencies at a large number of SNPs for many individuals in multiple populations (Futschik & Schlötterer 2010), when using whole genome sequencing (Kolaczkowski *et al.* 2011; Turner *et al.* 2011; Boitard *et al.* 2012) or reduced representation libraries (van Tassell *et al.* 2008; Pérez-Enciso & Ferretti 2010). Combining pooling and reduced representation is thus an affordable way to obtain a large number of informative SNPs.

Because information about distribution of SNPs across the genome is important for further population genomic analyses, mapping contigs to a reference genome remains critical. By taking advantage of the high degree of genome stability in birds (Backström *et al.* 2008; Griffin *et al.* 2008; Warren *et al.* 2010), we were able to build and map white-eye contigs to the zebra finch (*Taeniopygia guttata*) genome (Warren *et al.* 2010). In this note, we describe our protocol from library construction to contig mapping, with an emphasis on how to fully use information from paired-end reads, and evaluate its performance with regard to mass-producing SNP markers in our study species.

Methods

Library construction

DNA was extracted from individual blood samples using the QIAGEN DNeasy® Blood and Tissue kit, following manufacturer's instructions. Six pools were prepared, each including genomic DNA from 18 to 25 individuals and representing two replicates with different individuals from three distinct localities (named 'Bois Ozoux', 'Tévelave' and 'Pas de Bellecombe'). To minimize the risk of high variance in the number of reads per individual within a pool, we assessed double-stranded DNA concentration for each individual sample using the Quant-iT™ dsDNA Assay Kit (Invitrogen) and made the necessary adjustments to bring each individual DNA in the pool to equal molar concentration. Each library was built with equally represented samples for a total amount of 3 μg of total genomic DNA in a final volume of 75 μL.

To prepare RAD libraries, genomic DNA was digested with *EcoRI*, a widely used enzyme, cutting frequently enough so that a good coverage of the *Z. borbonicus* genome (around 300 000 restriction sites) could be obtained without compromising sequencing depth per locus. Each 75 μL-library was digested using the Promega *EcoRI* restriction reagents. For each reaction, 12 μL of H buffer, 30 μL of pure water and 3 μL (36 units) of enzyme were added (final volume: 120 μL). The reaction mix was divided into three 40 μL aliquots, and digestion was performed at 37 °C overnight, ending with a 20-min deactivation step at 65 °C. The three aliquots from each library were then progressively cooled at 4 °C and pooled. Based on the protocol by Baird *et al.* (2008) and customizing the sequences given by Illumina (oligonucleotide sequences © 2007-2012 Illumina, Inc., all rights reserved), we built new P1 and P2 adapters compatible with the paired-end technology currently supported by the Illumina HiSeq2000 system (whereas the adapters by

Baird and collaborators were designed for a previous version of single-read technology on the Genome Analyzer IIx system). Our P1 adapter, which includes the *EcoRI* restriction site, contains reverse amplification and Illumina sequencing primer sites, as well as a six base-long barcode sequence for sample identification, following the design of the TruSeq indexed adapters of Illumina (Table 1). Barcodes differed by at least four nucleotides to avoid misidentification of samples. This design allows the barcode to be read independently of the two reads of the genomic insert, in contrast to previous protocols in which the barcode was read together with the genomic fragment (*e.g.* Etter *et al.* 2011), avoiding subsequent barcode trimming from the raw reads. Barcodes were chosen among the 24 Illumina TruSeq Barcodes in order to be used with the Illumina TruSeq PCR Kit and to be easily read by the HiSeq2000 during the barcode read of the run. Our P2 adapter contains forward amplification and Illumina sequencing primer sites, following the design of the Illumina TruSeq Universal Adapter. As previously described (Baird *et al.* 2008), an asymmetric design of the P2 adapter ensured that only P1-ligated fragments could be amplified during the final amplification step (Table 1).

For each adapter, stocks of nonannealed oligonucleotides (Table 1) were diluted at 100 μM in 1× elution buffer (10 mM Tris-Cl, pH 8.5). Then, the pairs of forward and reverse oligonucleotides for each adapter were combined at 10 μM in 1× AB buffer (10× AB: 50 mM NaCl, 10 mM Tris-Cl, pH 8.0). Each stock of adapters was denatured 2 min at 95 °C and slowly cooled for 45 min at room temperature to obtain double-stranded adapters. P1 adapters were then diluted at a final concentration of 100 nM in 1× AB buffer. P2 adapters were used at a 10 μM concentration. 40 μL ligations were performed using the Promega High Concentration T4 DNA ligase, adding 18 μL of T4 DNA ligase Buffer (10×), 6 μL of NaCl (500 mM) and 18 μL purified water. Salt was added to ensure double-stranded adapters stability. After a 15 min incubation step at room temperature, 15 μL of 100 nM P1 adapters and 30–60 units of HC T4 DNA ligase (3 μL) were added to the mix. The reaction mix was divided into 60 μL aliquots and incubated at 22 °C for three hours, then deactivated for 10 min at 70 °C. For each library, aliquots were then pooled, purified using the QiaQuick PCR purification kit and eluted in a final volume of 100 μL.

Fragmentation of digested DNA was performed by sonication on a Bioruptor (Diagenode), using 10 cycles of 30 s on and 90 s off, in high mode. Control of sonication was done by checking fragment sizes with Bioanalyzer. Purification on Agencourt AMPure XP beads (Beckman-Coulter) was then performed, and the DNA of each library was eluted in 25 μL Resuspension Buffer (Illumina). Fragments around 500 bp (± 150 bp) were selected on an E-Gel system (E-Gel® CloneWell 0.8% SYBR Safe™ gel, Life Technologies) and retrieved in 25 μL Resuspension Buffer. Fragment end repair and adenylation were performed using Illumina TruSeq DNA Sample Preparation kit and guidelines. After another AMPure purification and elution in 45 μL EB buffer, the P2 ligation was performed by adding 5.8 μL of T4 DNA ligase buffer (10×), 5.8 μL of NaCl (500 mM), 1 μL of P2 adapter (10 μM) and 10 units of high

Table 1 Modified Illumina© adapters (a) used in this study (Oligonucleotide sequences © 2007-2012 Illumina, Inc., all rights reserved). In the P1 oligos: underlined nucleotides correspond to the overhanging end of the *EcoRI* restriction products; '**XXXXXX**' ('**YYYYYY**' for reverse strand) refers to the index-sequence or barcode. The sequences of barcodes used are given in (b) with corresponding localities. P2 adapter is designed to allow the amplification only of P1-linked DNA fragments. [PHO] designs the addition of a phosphate group in 5', * indicates the addition of a phosphorothioate bond to enhance nuclease resistance

Oligonucleotide	Sequence
(a)	
P1 forward	[PHO]AATTAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC**XXXXXX**ATCTCGTATGCCGTCTTCTGCTTG
P1 reverse	CAAGCAGAAGACGGCATACGAGAT**YYYYYY**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
P2 forward	[PHO]GATCGGAAGAGCGTCGTG
P2 reverse	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T

Barcode	Locality	Latitude (degrees)	Longitude (degrees)	Number of individuals
(b)				
ATCACG	Bois Ozoux	−21.198	55.647	18
TTAGGC	Bois Ozoux	−21.198	55.647	25
ACTTGA	Tévelave	−21.169	55.387	24
GATCAG	Tévelave	−21.169	55.387	20
TAGCTT	Pas de Bellecombe	−21.217	55.688	25
GGCTAC	Pas de Bellecombe	−21.217	55.688	25

concentration ligase (0.5 μL). After two other AMPure purifications and concentration in 20 μL Resuspension Buffer, an aliquot was kept as a control on Bioanalyzer.

An underappreciated issue of RAD sequencing concerns the amount of genomic material that is needed for library construction because *ca*. 1 μg of good quality genomic DNA is typically required per library (Baird *et al.* 2008). Using the Illumina HiSeq2000 technology requires even greater DNA quantities (3 μg per library) in order to satisfy quality standards. While this could be seen as an inconvenience, it makes possible to reduce the number of PCR cycles needed to obtain usable libraries, thereby reducing PCR amplification biases that can be especially problematic when dealing with samples with low DNA concentration. Thus, an enrichment step was performed for each library, consisting in just 12 cycles of PCR amplification, using TruSeq Sample Prep PCR Kit and guidelines from Illumina. After AMPure purification, a final step of size selection was then performed on the libraries to remove the remaining adapters, using the E-Gel system again.

Library profiles were controlled on a BioAnalyzer High Sensitivity chip. Finally, quantities of usable material for each of the six libraries were estimated by qPCR (KAPA Library Quantification Kit–Illumina Genome Analyzer-SYBR Fast Universal) and then normalized and pooled. The quality of the pool was then checked using qPCR and immediately followed by sequencing on the HiSeq2000 platform (Plateforme Génomique - Genopole Toulouse Midi-Pyrénées), using TruSeq PE Cluster Kit v3 (2 × 100 pb) and TruSeq SBS Kit v3.

Assembling contigs and SNP detection

Large contigs are required to perform accurate alignments to a related genome. In the case of the white-eye, the divergence time to the zebra finch is estimated to be 40 million years (Barker *et al.* 2004), which prevents the use of short-read alignment tools such as BWA (Li & Durbin 2009). Therefore, we used a pipeline aimed at assembling large contigs (300–500 bp), using information from both paired-end reads. First ustacks (version 0.9995) and then cstacks (Catchen *et al.* 2011) were used to group reads immediately flanking restriction sites (reads 1) for each of the six libraries, allowing up to 3 mismatches between stacks (ustacks options: -m 4 -M 3). Loci not found in at least 5 of the 6 libraries were discarded. Both reads 1 and reads 2 were aligned with BWA (version 0.7.0) on this catalogue of stack, forcing the alignment of the second read by using a custom python script (RAD_assign_reads_to_consensus.py, all scripts available at the address https://github.com/tcezard/RADmapper). The resulting .bam file was then translated into several .fastq files, each corresponding to one stack, using another python script (RAD_bam_to_fastq.py). For each locus, a consensus of reads 2 was then assembled with the consensus of reads 1, using a third python script (RAD_assemble_read2.py) making use of the IDBA_UD assembler (Peng *et al.* 2012; version 1.0.9), which is a fast assembler originally designed for single-cell assembly, not relying upon an even coverage and using several k-mer lengths. Reads 1 and reads 2 consensuses were merged into a single contig with EMBOSS (version 6.4.0.0) merger (Rice *et al.* 2000). Consensuses that did not overlap were forced into one contig by adding ten ambiguous bases ('N') between them. Contigs with the best assembly score were then extracted and used as a reference for mapping back reads and eliminating PCR duplicates with SAMTOOLS (Li *et al.* 2009, version 0.1.19). Options for the java script MarkDuplicates were as follows: VALIDATION_STRINGENCY=LENIENT, MAX_FILE_HANDLES_FOR_READ_ENDS_MAP=100, CREATE_INDEX=true, AS=true.

The mpileup2sync.jar java script (version 1.201) from POPOOLATION2 (Kofler *et al.* 2011b) was used to construct files with allele counts (synchronized 'pileup' files) with the following options: –fastq-type sanger –min-qual 20 –threads 8. To call biallelic SNPs, we applied a standard quality threshold of 20 (corresponding to an error probability of less than 1%), a global minimum allele count (MAC) of 2 or 3, a minimum sequencing depth of 10× or 20× by library (60× or 120× overall) and a maximum depth of 500×.

We checked that contigs did not contain an excess of SNPs because poor assembly and bad alignment of the reads can lead to a heterogeneous distribution of SNPs over contigs. Finally, we further assessed whether those SNPs could be due to massive sequencing errors by testing whether the same SNPs could be found in multiple libraries and in the two replicates from each locality. We performed SNPs calling by applying a MAC of three and a minimal sequencing depth of 20×.

LASTZ alignment

We assessed the quality of contig assembly by testing whether the contigs that were obtained could be mapped easily onto the zebra finch genome. To this end, all the contigs assembled from paired-end reads were aligned against the zebra finch genome (version July 2008, assembly WUGSC v.3.2.4). Alignment was performed using LASTZ (Harris 2007), an improved version of BLASTZ (Schwartz *et al.* 2003) using default parameters, except for 'ambiguous=n' and ydrop=7000, mainly to allow large gaps (up to 220 bp) inside contigs because consensuses from paired reads sometimes did not overlap. A minimum identity of 60% and coverage of 70% were required.

Results

Mapping of reads on a related genome

More than 154 million usable paired reads (2 × 100 bp) were generated from the six libraries using only one Illumina HiSeq2000 lane, with a mean sequencing depth of 27× per library at the end defined by the restriction site (SD between libraries = 2.83). From these reads, we generated a total of 606 725 contigs, 99% ranging in size from 101 to 612 bp (mean= 396 bp, median = 384 bp; number of contigs larger than 100 bp = 592 712; number of contigs larger than 300 bp = 582 193; N50 = 402 bp). Eighty-six per cent of these contigs could be mapped to the zebra finch genome (Table 2). After excluding all sites associated with a chromosome but with no known position ('random' chromosomes) and nonmapped repetitive sequences ('Unknown' chromosome), we found that 398 793 contigs were unique hits, and 386 841 (63.8% of all contigs) could be positioned unambiguously. A fraction of the contigs (14.5%) mapped to only two distinct locations in the zebra finch genome. Nearly two-thirds of these contigs mapped onto a known chromosome and also onto the 'Unknown' chromosome, suggesting a possible location in repetitive regions.

Having started with a relatively low-identity requirement (60.0%), we finally obtained a large set of contigs unambiguously aligned displaying between 80.0 and 100.0% identity (mean: 89.9%; SD: 3.5%) with the corresponding zebra finch sequence. In birds, substitution rates in autosomes have been estimated at 3.6×10^{-9} substitution per year per site (Axelsson *et al*. 2004). Using a divergence time of ca. 40 MYA (Barker *et al*. 2004), we would expect a sequence divergence of 14.0% between zebra finch and *Zosterops*, a figure that is consistent with our data.

Table 2 Counts of contigs mapped onto the zebra finch genome. Repartition of double hits and unique hits onto the zebra finch genome are detailed

Category of hits	Count	Percentage
Total	606 725	100
Mapping onto zebra finch	523 744	86.32
More than two hits	37 113	6.12
Total of double hits	87 838	14.48
Double hits with one single hit on unknown chromosome	60 539	9.98
Total of unique hits	398 793	65.73
Known position	386 841	63.76
Known chromosome (random)	9521	1.57
Unknown chromosome	2431	0.40

Distribution of contigs across chromosomes

We observed a strong correlation between chromosome size and the number of contigs mapped unambiguously (Fig. 1, $R^2 = 0.987$, P-value <0.0001). We did not observe any obvious outliers and were able to map contigs even on the smallest assembled chromosomes, such as chromosomes 16 or 1B. This indicates that contigs and associated SNPs were evenly distributed across the white-eye genome.

SNP calling

The number of reads that qualified as PCR duplicates accounted for less than 23.0% of the whole data set, a figure that remains low compared to some other RAD-sequencing studies (M. Gautier, personal communication). This is probably due to the fact that we used relatively large amounts of genomic DNA (~3 μg) and performed no more than 12 amplification cycles. Also, we used several stringency criteria to call SNPs (Table 3). Using a MAC of three and a minimum sequencing depth of 20× in each library after quality and duplicate filtering, we were able to identify 133 958 SNPs. Of these, 81 246 SNPs (60.7%) could be mapped unambiguously, which is consistent with the proportion of contigs with a unique hit at a known position (63.8%) on the zebra finch genome.

Minor allele frequencies (MAF) for the whole data set were mainly below 0.2, with a mean frequency of 0.103 (SD: 0.116) when using the most stringent conditions for SNP calling. Reducing sequencing depth to at least 10× per library did not drastically change the MAF distribution (mean frequency: 0.114, SD: 0.118). When

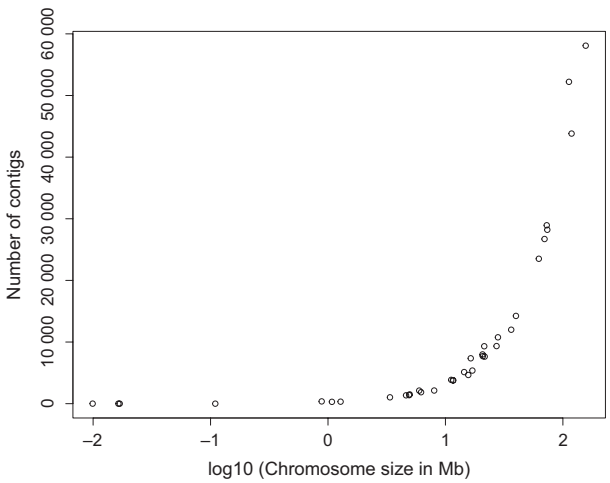


Fig. 1 Correlation between chromosome size and number of unambiguous contig hits.

considering each library, no obvious differences in MAF distributions could be observed between replicates from the same locality (Fig. 2a). Differences in allele frequencies were low (means of 0.061, 0.067 and 0.066 for 'Bois Ozoux', 'Tévelave' and 'Pas De Bellecombe' libraries, respectively) and in the range of sampling noise (Fig. 2b). A total of 118 683 (89%) SNPs were found in a minimum of two libraries (Table 4), and most loci found to be polymorphic in a given replicate were also polymorphic in the other replicate from the same locality. On average, each polymorphic contig contained 2 SNPs (Fig. 3; median = 1, SD = 2.36), even when considering SNPs sampled at a lower depth (mean = 2.54, median = 2, SD = 2.72). This further suggests that contigs were correctly assembled and that multiple SNPs on contigs were not due to poor alignment, but rather to contig length.

Table 3 Number of SNPs called following several stringency criteria

Minimal sequencing depth by library	MAC	
All SNPs	2	3
10×	397 969	327 705
20×	163 086	133 958
SNPs at a known position		
10×	244 384	199 955
20×	99 821	81 246

MAC, minimum allele count required to call a SNP.

Discussion

There has been an increasing interest in the use of next-generation sequencing to address evolutionary questions in nonmodel species (Davey *et al.* 2011). However, the lack of a reference genome, the costs associated with the sequencing of many individuals and the need to compare homologous sequences across individuals have remained limiting when trying, for example, to associate SNPs to genes and traits of interest.

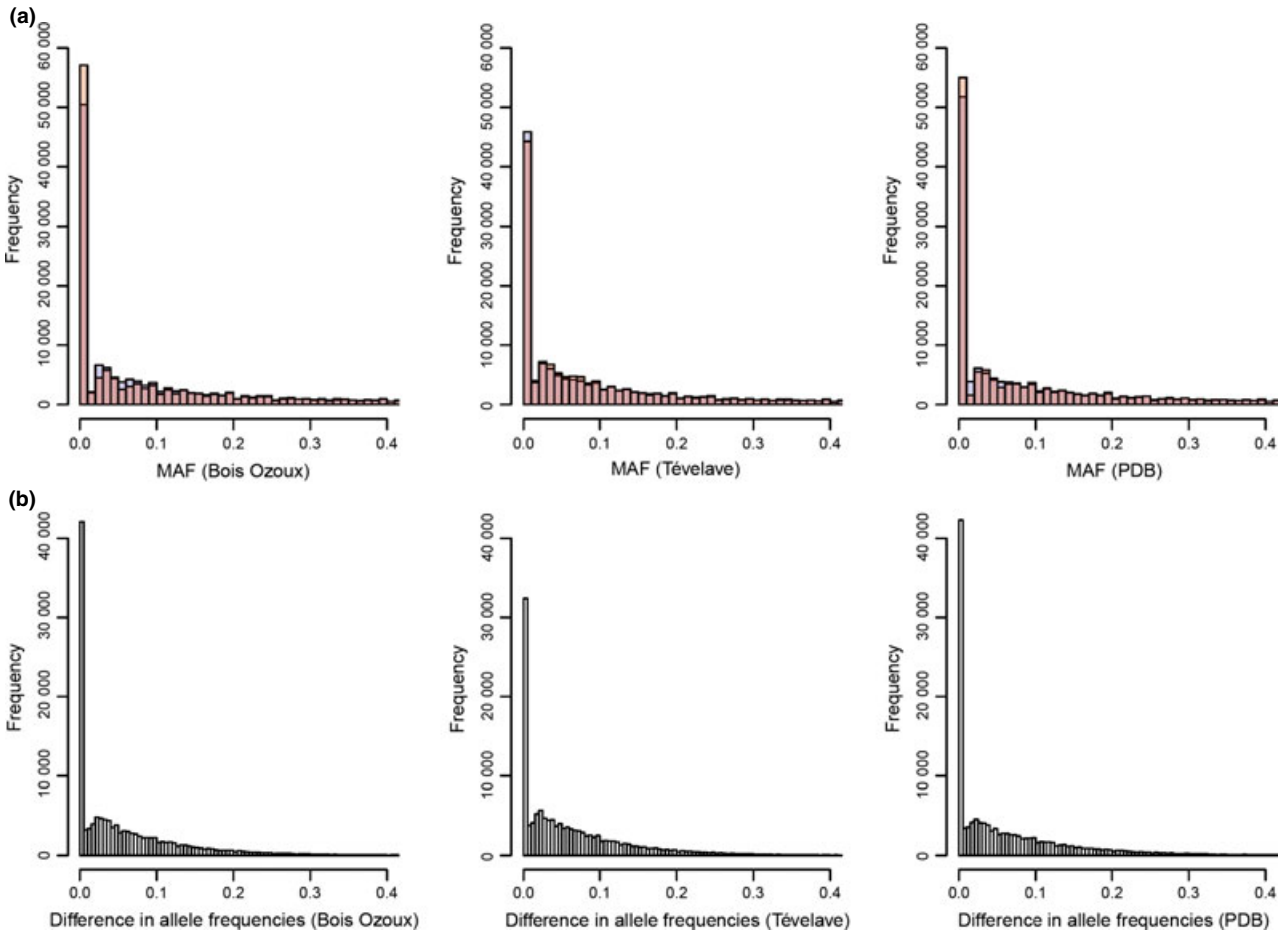


Fig. 2 Distribution of minimum allele frequencies (MAF) for each library used in this study, grouped by population (a). Changes in allele frequencies between libraries from the same population are also plotted (b). PDB: Pas de Bellecombe.

Table 4 Comparison of allele frequencies and count of shared polymorphisms in two replicates for each of the three populations. Estimates are based on SNPs obtained using a minimum sequencing depth of 20× and a MAC of 3

Population	Mean difference in allele frequencies	Median difference in allele frequencies	Standard deviation	Number of SNPs polymorphic in one library	Number of SNPs polymorphic in both libraries
Bois Ozoux	0.061	0.036	0.076	95 404	65 504
Tévelave	0.067	0.043	0.076	105 852	72 727
PDB	0.066	0.036	0.083	95 090	66 726

Our HiSeq2000-based RAD sequencing protocol, by making use of the very large and cost-efficient production of paired-end sequences for pools of individuals, has enabled us to build a very large number of 300–500-bp contigs that could be mapped unambiguously onto the zebra finch genome. This led to the discovery of more than one hundred thousand SNPs across 137 individuals sampled in three relatively close localities, separated by less than 35 km.

Through our experiment, we also confirm the usefulness of the RAD sequencing approach to identify the genomic position of markers in passerine birds, even when dealing with nonmodel species. This had been previously suggested by van Bers *et al.* (2010) in their study of the great tit (*Parus major*). However, comparing their results to ours, we note that we detected six times more SNPs and were able to map nearly 20 times more SNPs to unique locations distributed over the zebra finch genome. This was mostly due to the fact that we obtained many more contigs and that a much larger proportion of these contigs were greater than 100 bp (3.5% versus 97.7%). Thus, our approach based on pooled DNA samples, which keeps the cost of library construction and adapter preparation to a minimum, and HiSeq2000 technology provides a cost-effective strategy for SNP detection and mapping in a passerine bird species for which a sequenced genome is currently lacking.

While haplotype information is obviously not available, DNA pools produce a high number of informative SNPs that are ideal for characterizing variation in population samples or can subsequently be assayed in further experiments. Because synteny is remarkably conserved in birds (Derjusheva *et al.* 2004; Griffin *et al.* 2008), a large number of these SNPs can be readily identified as orthologous of known genes in the zebra finch genome, providing unprecedented opportunities to describe the molecular background of nonmodel passerine birds. Depending on the pooling strategy, the approach provides a wide range of applications that will enable students in ecology and evolution to have access to genome-wide allele frequency estimates, to compare patterns of differentiation on a genomic scale, to

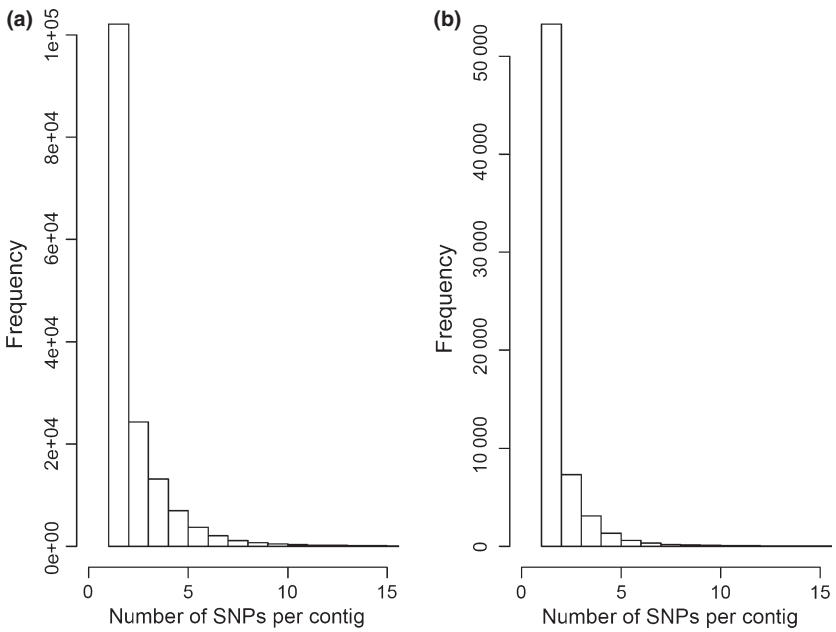


Fig. 3 Distribution of the number of SNPs per polymorphic contig for a sequencing depth of 10× and a minimum allele count (MAC) of 2 (a) and for a sequencing depth of 20× and a MAC of 3 (b).

characterize the demographic history of differentiated populations and detect selective sweeps, or even to investigate the genetic basis of ecologically significant traits using genome-wide association mapping (e.g. Boitard *et al.* 2012; Rubin *et al.* 2010; Willing *et al.* 2011; Kofler *et al.* 2011a; Futschik & Schlötterer 2010; Zhu *et al.* 2012; Gautier *et al.* 2013). This may pave the way to other approaches based on genotyping by sequencing that will then allow to get individual genotypes and to characterize molecularly precise variants within a population (see e.g. Garroway *et al.* 2013; Hagen *et al.* 2013).

Acknowledgements

Ben Warren, Guillaume Gélinaud, Dominique Strasberg, Juli Broggi, Magali Thierry, René-Claude Billot, Jean-Michel Probst, Isabelle Henry, Vincent Leconte, Marc Salamolard, Benoît Lequette, We gratefully acknowledge the Réunion National Park for permission to conduct fieldwork. We thank Mathieu Gautier for his insights about pooling experiments. Ulrich Knief and an anonymous reviewer provided valuable comments that greatly improved an earlier version of this manuscript. This work was supported by Institut Français de la Biodiversité (IFB), Agence Française pour le Développement (AFD) and ANR Biodiversity Program grants to CT, the Génopole Toulouse Midi-Pyrénées, the National Geographic Society and the 'Laboratoire d'Excellence' TULIP (ANR-10-LABX-41). YB, JB and BD were supported by MESR (Ministère de l'Enseignement Supérieur et de la Recherche) PhD scholarships.

References

Altshuler D, Pollara VJ, Cowles CR *et al.* (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature*, **407**, 513–516.

Axelsson E, Smith NGC, Sundström H, Berlin S, Ellegren H (2004) Male-biased mutation rate and divergence in autosomal, z-linked and w-linked introns of chicken and Turkey. *Molecular Biology and Evolution*, **21**, 1538–1547.

Backström N, Fagerberg S, Ellegren H (2008) Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Molecular Ecology*, **17**, 964–980.

Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, **3**, e3376.

Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J (2004) Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 11040–11045.

van Bers NEM, van Oers K, Kerstens HHD *et al.* (2010) Genome-wide SNP detection in the great tit *Parus major* using high throughput sequencing. *Molecular Ecology*, **19**(Suppl 1), 89–99.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044–2072.

Boitard S, Schlötterer C, Nolte V, Pandey RV, Futschik A (2012) Detecting selective sweeps from pooled next-generation sequencing samples. *Molecular Biology and Evolution*, **29**, 2177–2186.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci *de novo* from short-read sequences. *Genes, Genomes, Genetics*, **1**, 171–182.

Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, **9**, 416–423.

Davey JW, Hohenlohe PA, Etter PD *et al.* (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

Davey JW, Cezard T, Fuentes-Utrilla P *et al.* (2012) Special features of RAD Sequencing data: implications for genotyping. *Molecular Ecology*, **22**, 3151–3164.

Derjusheva S, Kurganova A, Habermann F, Gaginskaya E (2004) High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. *Chromosome Research*, **12**, 715–723.

Emerson KJ, Merz CR, Catchen JM *et al.* (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 16196–16200.

Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA (2011) Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *PLoS One*, **6**, e18561.

Futschik A, Schlötterer C (2010) The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics*, **186**, 207–218.

Gagnaire P-A, Normandeau E, Pavey SA, Bernatchez L (2012) Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish (Coregonus clupeaformis). *Molecular Ecology*, **22**, 3036–3048.

Garroway CJ, Radersma R, Sepil I *et al.* (2013) Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. *Evolution*, doi: 10.1111/evo.12121.

Gautier M, Foucaud J, Gharbi K *et al.* (2013) Estimation of population allele frequencies from next-generation sequencing data: pooled versus individual genotyping. *Molecular Ecology*, **22**, 3766–3779.

Griffin DK, Robertson LB, Tempest HG *et al.* (2008) Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics*, **9**, 168.

Hagen IJ, Billing AM, Rønning B *et al.* (2013) The easy road to genome-wide medium density SNP screening in a non-model species: development and application of a 10 K SNP-chip for the house sparrow (Passer domesticus). *Molecular Ecology Resources*, **13**, 429–439.

Harris RS (2007) *Improved pairwise alignment of genomic DNA*. PhD Thesis, Pennsylvania State University, University Park, Pennsylvania.

Hawkins RD, Hon GC, Ren B (2010) Next-generation genomics: an integrative approach. *Nature Reviews Genetics*, **11**, 476–486.

Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2012) Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Molecular Ecology*, **22**, 2898–2916.

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources*, **11** (Suppl 1), 117–122.

Keller I, Wagner CE, Greuter L *et al.* (2012) Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology*, **22**, 2848–2863.

Kofler R, Orozco-terWengel P, de Maio N *et al.* (2011a) PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One*, **6**, e15925.

Kofler R, Pandey RV, Schlötterer C (2011b) PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, **27**, 3435–3436.

Kolaczkowski B, Kern AD, Holloway AK, Begun DJ (2011) Genomic differentiation between temperate and tropical Australian populations of *Drosophila melanogaster*. *Genetics*, **187**, 245–260.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Miller M, Dunham J, Amores A *et al*. (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, **17**, 240–248.

Peng Y, Leung HCM, Yiu SM, Chin FYL (2012) IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, **28**, 1420–1428.

Pérez-Enciso M, Ferretti L (2010) Massive parallel sequencing in animal genetics: wherefroms and wheretos. *Animal Genetics*, **41**, 561–569.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*, **7**, e37135.

Radwan J, Babik W (2012) The genomics of adaptation. *Proceedings of the Royal Society of London Series B. Biological Sciences*, **279**, 5024–5028.

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. *Trends in Genetics*, **16**, 2–3.

Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, **66**, 1–17.

Rubin C-J, Zody MC, Eriksson J *et al*. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, **464**, 587–591.

Schwartz S, Kent W, Smit A *et al*. (2003) Human–mouse alignments with BLASTZ. *Genome Research*, **13**, 103–107.

Takahashi T, Sota T, Hori M (2012) Genetic basis of male colour dimorphism in a Lake Tanganyika cichlid fish. *Molecular Ecology*, **22**, 3049–3060.

van Tassell C, Smith T, Matukumalli L *et al*. (2008) SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nature Methods*, **5**, 247–252.

Travisano M, Shaw RG (2013) Lost in the map. *Evolution*, **67**, 305–314.

Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM (2011) Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genetics*, **7**, e1001336.

Wagner CE, Keller I, Wittwer S *et al*. (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**, 57–63.

Wang N, Thomson M, Bodles WJA *et al*. (2012) Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. *Molecular Ecology*, **22**, 3098–3111.

Warren WC, Clayton DF, Ellegren H *et al*. (2010) The genome of a songbird. *Nature*, **464**, 757–762.

Willing EA, Hoffmann M, Klein JD, Weigel D, Dreyer C (2011) Paired-end RAD-seq for de-novo assembly and marker design without available reference. *Bioinformatics*, **27**, 2187–2193.

Zhu Y, Bergland AO, González J, Petrov DA (2012) Empirical validation of pooled whole genome population resequencing in Drosophila melanogaster. *PLoS One*, **7**, e41901.

Y.B., B.M. and C.T. planned the project and wrote the article. B.M. and C.T. obtained funding and organized sample collection. T.C. provided scripts for data analysis. Y.B., J.B., B.D., T.D., J.C., C.T. and B.M. provided samples and performed fieldwork. Y.B., E.L. and O.B. prepared libraries.

Data accessibility

Raw sequence information has been submitted as bam files to the European Nucleotide Archive (ENA) repository at the accession number ERP002555 (available at http://www.ebi.ac.uk/ena/data/view/ERP002555). All scripts used for contig reconstruction are available at https://github.com/tcezard/RADmapper. Files with allelic counts and afasta file with contigs have been deposited on DRYAD (http://dx.doi.org/10.5061/dryad.755b5).

AUTEUR : Boris DELAHAIE

TITRE : Spéciation, zones hybrides et gradients environnementaux : le cas du Zostérops des Mascareignes

DIRECTEURS DE THESE : Christophe Thébaud et Borja Milá

LIEU ET DATE DE SOUTENANCE : Université Paul Sabatier le 13 mars 2015

RESUME

Les îles fournissent de bonnes opportunités pour étudier l'émergence de la biodiversité de part leur contexte spatial facilement appréhendable. Nous avons étudié une espèce de passereau endémique de l'île de la Réunion : le Zostérops des Mascareignes, *Zosterops borbonicus*. Cette espèce présente une extraordinaire variabilité de la couleur de son plumage à une échelle spatiale rarement documentée chez les oiseaux. L'analyse des patrons de variations génétiques et phénotypiques le long de gradients altitudinaux et au travers des zones hybrides séparant les différentes formes de couleur de l'espèce a permis de mettre en évidence le rôle de différents facteurs (sélectifs, historiques et neutres) dans l'émergence et le maintien de cette diversité.

MOTS-CLES : spéciation, sélection naturelle, isolement reproducteur, gradients environnementaux, zones hybrides, Zosterops.

DISCIPLINE ADMINISTRATIVE : ECOLOGIE

INTITULE ET ADRESSE DU LABORATOIRE :
Laboratoire Evolution & Diversité Biologique (EDB) - UMR 5174 (CNRS/UPS)
Université Paul SABATIER, 118 route de Narbonne 31062 TOULOUSE CEDEX 9 - France
Bâtiment 4R1.