UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS

TESIS DOCTORAL

Variación genética y fenotípica en "Psammodromus algirus" (lagartija colilarga: implicaciones ecológicas y evolutivas

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Alejandro Llanos Garrido

Directores

José Augusto Díaz González-Serrano
Javier Pérez Tris

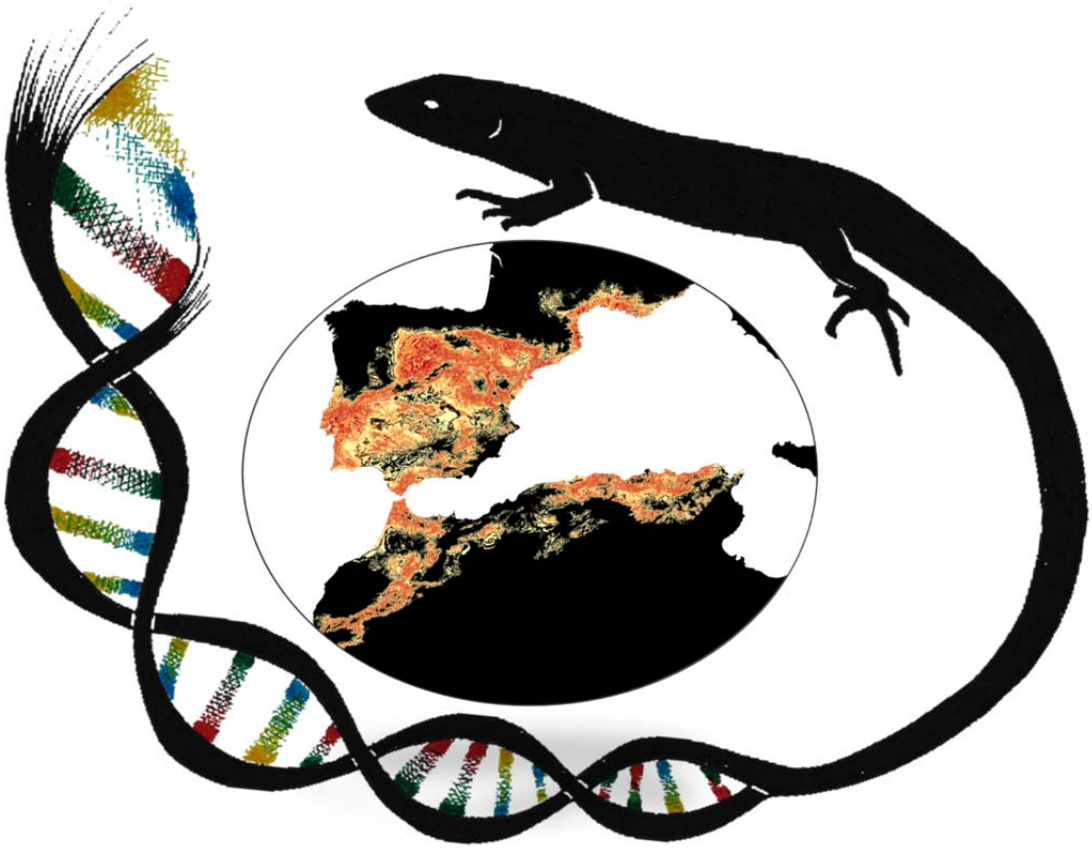
Madrid
Ed. electrónica 2019

Variación genética y fenotípica en *Psammodromus algirus* (lagartija colilarga): implicaciones ecológicas y evolutivas

Alejandro Llanos Garrido

Tesis Doctoral

Abril 2019



Departamento de Biodiversidad, Ecología y Evolución
Facultad de Ciencias Biológicas
Universidad Complutense de Madrid

La presente Tesis Doctoral ha sido financiada por un contrato de Formación de Personal Investigador (FPI, EEBB-I-17-12533) concedido al proyecto CGL2013-41642-P financiado por el Ministerio de Economía, Industria y Competitividad.

Universidad Complutense de Madrid

Facultad de Ciencias Biológicas

Departamento de Biodiversidad, Ecología y Evolución

Variación genética y fenotípica en *Psammodromus algirus* (lagartija colilarga): implicaciones ecológicas y evolutivas.

Memoria presentada por Alejandro Llanos Garrido para optar al grado de Doctor en Ciencias Biológicas, dirigida por los doctores José A. Díaz González-Serrano y Javier Pérez Tris, del Departamento de Biodiversidad, Ecología y Evolución de la Universidad Complutense de Madrid.

Madrid, abril de 2019

El Doctorando **VºBº de los Directores**

Alejandro Llanos Garrido **José A. Díaz González-Serrano** **Javier Pérez Tris**

AGRADECIMIENTOS

En primer lugar, estoy completamente obligado a agradecer, y mucho, a mis directores Javi y Pepe, sin ellos todas estas páginas en blanco estarían llenas de borrones mucho menos entretenidos. No sé dónde hubiera acabado yo si Pepe no me hubiera leído tan bien hace ya más de siete años, o si Javi no hubiera confiado en mí para llevar a cabo esta tarea, pero no me imagino un entorno mejor en el que haber madurado este pedazo de ser. Gracias por las risas cuando el cerebro ya no daba más de sí y por ayudarme a soportar el crítico escrutinio del suspicaz Dobzhansky mientras le dábamos forma a la tesis.

A mis amigos no becarios porque sin ellos hubiera perdido la cabeza mucho antes (o mucho más fuerte), lo que sin duda me hubiera incapacitado para acabar la tesis a tiempo. Gracias por sujetarme sin tan siquiera pretenderlo cuando me flaqueaban las piernas, habéis sido una pieza mucho más clave de lo que seguramente penséis. Gracias a mis amigos becarios también (los que fueron, son o serán becarios) por los estímulos intelectuales dentro y fuera del campo, seguramente habéis pasado más tiempo conmigo que muchas de las personas de mi entorno, muchas gracias por aguantar mis demonios con estoicismo.

A mi familia, por el orgullo mantenido aun después de tropezar más veces que piedras había en el camino. Por ofrecerme madurar en presencia tanto de la razón como de la emoción más pura, sin eso seguramente no hubiera llegado a querer dedicarme a pensar.

Gracias a Andrea por el ser y tiempo que le he robado, por obligarme a existir a punta de pistola y arañarme en la sien con precisión cirujana. A ti Andrea, por no llegar nunca ni pronto ni tarde, por llegar siempre cuando te lo has propuesto, siempre que era necesario.

ÍNDICE

RESUMEN

Esta tesis trata de, una vez explorados los rasgos fenotípicos que pueden interpretarse como adaptaciones locales, contestar a preguntas que tienen que ver con cómo responde el genoma a las presiones selectivas divergentes que subyacen a la diferenciación de dichos rasgos. En general, no se pretende descubrir la base genética de la adaptación a los gradientes ambientales que generan adaptación local; lo que se busca es desentrañar cómo la evolución provee la variación genética necesaria para afrontar diferentes retos ecológicos en una especie capaz de subsistir en una gran variedad de ambientes. Mediante escaneos genómicos a gran escala, conseguimos definir fracciones del genoma relacionadas con la adaptación a distintos ambientes, revelando patrones de convergencia y divergencia adaptativa gracias a procesos tanto de mantenimiento de diversidad genética ancestral como de innovación genética. Tras conseguir descifrar la huella de la selección natural en el genoma, nos preguntamos acerca de las dinámicas genéticas responsables de los umbrales adaptativos que configuran los límites de la distribución de la especie, como proceso que depende de forma directa de la adaptación local. Por último, realizamos un estudio de los efectos genéticos de la fragmentación del hábitat en una metapoblación de lagartijas, en la que el aumento de la homocigosidad en los fragmentos de menor tamaño dio lugar a una pérdida de aptitud (eficacia biológica o aptitud) en los individuos afectados. Además, nuestro trabajo nos permitió aumentar significativamente la potencia estadística de unos pocos marcadores neutrales (6 loci de microsatélites utilizados para sustentar las correlaciones entre fragmentación, homocigosis y aptitud), validándolos con una base de datos genómica con más de 73,000 variantes genéticas.

ABSTRACT

The aim of this PhD Thesis is to clarify the genomic basis of local adaptation, as documented by divergence in adaptive phenotypic traits. For that purpose, I address different questions all of them related with genetic responses to divergent selective pressures along environmental gradients on a genome-wide scale. My aim is not to uncover the genetic basis of particular adaptations, but rather to reveal how evolution provides the genetic variation that ecology demands in a widespread species that is able to survive in a large variety of environmental conditions. By means of genomic scans based on massive sequencing and different methods of outlier detection, I could define genome portions that could be related to adaptation along a well-defined environmental gradient. My results revealed both adaptive divergence and convergence, which were achieved by a combination of divergent selection, balancing selection, genetic innovation, and gene flow. After having defined the genomic markers linked to adaptation (i.e. outlier loci subject to selection), I explored the genetic dynamics that could underlie adaptive thresholds and, as a consequence, determine the location and shape of distribution range boundaries; my goal here was to model the overall distribution of a model species on the basis of a few genetic markers that lie at the core of local adaptation dynamics. Finally, in the last chapter I documented how increased individual homozygosity correlates with low fitness in a fragmented lizard population. Moreover, I substantially increased the power of the original markers used to support the homozygosity-fitness correlations (i.e. 6 microsatellite loci) by demonstrating the ability of such markers to predict homozygosity estimates on a genome-wide basis (with more than 70,000 SNPs).

1. INTRODUCCIÓN

1.1. Marco conceptual

1.1.1. *¿Qué rasgos fenotípicos ponen de manifiesto procesos de adaptación local?*

La adaptación local es el proceso mediante el cual los individuos de una especie presentan una mayor aptitud en sus ambientes locales que los individuos que provienen de otras poblaciones (Williams 1966). Este proceso tiene lugar cuando la selección natural actúa en base a diferencias ambientales entre poblaciones de la misma especie, dando lugar a diferentes fenotipos. Por esta razón, la adaptación local se encuentra en el núcleo de la explicación del origen y mantenimiento de la variación fenotípica dentro de las especies, además de dar respuesta a una pregunta completamente central en ecología evolutiva: ¿por qué la variabilidad fenotípica presenta patrones espaciales heterogéneos a lo largo del área de distribución de una especie? (Endler 1977, Hereford 2009, Savolainen et al. 2013, Schluter 2000).

Una de las demostraciones empíricas de que la variabilidad fenotípica observada es el resultado de la adaptación local consiste en realizar experimentos de trasplante recíproco, en los que se trasladan individuos entre los distintos sitios en los que aparentemente están sujetos a presiones selectivas distintas, y a continuación se miden los cambios en el rasgo de interés junto con los cambios en la aptitud de los individuos trasplantados (Clausen et al. 1948). Otra posibilidad, consiste en realizar estudios de ambiente común (*common-garden*) en los que individuos de distintas procedencias con distintos fenotipos son mantenidos en las mismas condiciones ambientales, de manera que también se puede medir la respuesta plástica al cambio de ambiente y contraponerla a la varianza fenotípica independiente de dicho cambio (Turesson 1922).

Sin embargo, estas pruebas empíricas de la existencia de adaptación local

presentan numerosas limitaciones para su puesta en marcha. En muchas ocasiones, los organismos son difíciles de trasplantar, y cuando esto se consigue, muchas veces resulta difícil obtener medidas globales de aptitud (capaces de abarcar una porción representativa del ciclo de vida de la especie). De hecho, no es de extrañar que estos experimentos fueran concebidos para realizarse con plantas herbáceas, de crecimiento muy rápido y ciclos cortos, de manera que se podía cuantificar la relación entre variabilidad fenotípica y aptitud en menos de un año (Anderson et al. 2015, Toräng et al. 2015). En cualquier caso, estos experimentos se han aplicado con éxito en el estudio de la adaptación local entre poblaciones animales permitiendo así controlar el componente ambiental que configura el fenotipo de los individuos que las constituyen (véase un ejemplo con la especie modelo de esta tesis en Iraeta et al. 2006, Iraeta et al. 2013). En estos estudios, se atiende normalmente a características fenotípicas concretas que resulten ser clave en el ciclo de vida de la especie, como la longevidad, fecundidad o tasa de crecimiento, de manera que no resulta completamente necesario mantener el experimento durante ciclos vitales enteros para relacionar dichos fenotipos con la aptitud de los individuos (Schluter 2001, Linn et al. 2003). Como contrapartida, estas medidas locales de aptitud (*sensu* Roff 1992) son generalmente incapaces de capturar fielmente la contribución de los individuos al acervo genético de la siguiente generación (su aptitud global), ya que en general la existencia de compromisos entre los rasgos del ciclo vital (longevidad, fecundidad, viabilidad de los descendientes, etc.) implica que la maximización de cualquiera de esos rasgos tenga un coste en términos de decremento de otros (Stearns 1989).

1.1.2. *Determinantes eco-inmunológicos potencialmente relacionados con los patrones de selección divergente*

La selección divergente que sustenta cualquier proceso de adaptación local tiene

lugar gracias a diferencias ambientales entre las poblaciones, incluyendo estructura de hábitat, clima, recursos, presencia de parásitos, etc. (Schluter 2001, Schluter 2005, Shaner et al. 2015). La contribución de estas diferencias ambientales a la divergencia adaptativa es un proceso muy bien documentado gracias a las numerosas demostraciones empíricas que se han llevado a cabo (Schluter 2001, Linn et al. 2003), en las que incluso se ha podido comprobar que ambientes distintos pueden generar aislamiento reproductivo en simpatría cuando la especie segrega en microhábitats (Rice y Salt 1990).

Algunas de las diferencias ecológicas entre poblaciones que podrían afectar de forma directa al aislamiento reproductivo y que promueven la adaptación local son aquellas que constituyen presiones selectivas que afectan directamente a la evolución de ornamentos sexuales (Endler y Houde 1995, Giery y Layman 2015). Esta idea se basa en la hipótesis del hándicap propuesta por Zahavi (1975) en la que se establece que la evolución y mantenimiento de ornamentos sexuales como señales honestas de calidad genética tiene que pasar por la existencia de *tradeoffs* entre dichas señales y diferentes aspectos relacionados con la aptitud de los individuos que las portan (riesgo de depredación, eficiencia locomotora, sistema inmune, etc.; Giery y Layman 2015, Moller y Lope 1994, Hamilton y Zuk 1982). Por lo tanto, si las distintas poblaciones de una especie presentan diferentes presiones selectivas involucradas en dichos compromisos, el posible aislamiento reproductivo que promueva la adaptación local podría verse mediado por la evolución de ornamentos sexuales (Endler y Houde 1995). Por ejemplo, en el caso de que las señales ornamentales provean información acerca de la calidad de los portadores a la hora de resistir parasitismo, las diferencias en la carga de parásitos entre poblaciones pueden dar lugar a que los individuos puedan permitirse desviar más recursos inmunitarios a la formación del ornamento cuando esta presión selectiva está

ausente (Hamilton y Zuk 1982, Folstad y Karter 1992). En estos casos, la recepción de la señal se podría volver más exigente en los contextos inmunológicos más favorables (desbocando así la evolución de la señal), o la propia ornamentación podría evolucionar hacia otras vías de señalización en los contextos en los que no se pudieran dedicar tantos recursos a la generación de dicha señal (Endler y Houde 1995, Giery y Layman).

1.1.3. *El papel de la información genómica para entender dinámicas de adaptación local*

Las técnicas de secuenciación masiva, junto con la disponibilidad de nuevos análisis de escaneo genómico, han permitido que los estudios de adaptación local hayan pasado de ser mayoritariamente fenotípicos a poder también profundizar en la arquitectura genética que subyace a este proceso, incluyendo la identificación de loci responsables de las adaptaciones divergentes (Hoban et al. 2016, Delph 2018). En este tipo de análisis genómico se buscan congruencias espaciales entre la variabilidad fenotípica que manifiesta adaptaciones locales (aptitud de los fenotipos locales incrementada con respecto a la de los fenotipos de otras localidades) y los genotipos de las poblaciones que las presentan (Tigano y Friesen 2016). Esta manera de explorar las dinámicas de adaptación local basadas en secuenciación masiva ofrece una perspectiva temporal que las aproximaciones experimentales descritas anteriormente (trasplante recíproco y *common garden*) no pueden conseguir. Además, estos análisis no están restringidos por las dificultades logísticas que representan dichas demostraciones empíricas, por lo que se pueden realizar en cualquier tipo de organismo.

El potencial de los escaneos genómicos para descubrir los genes que subyacen a la adaptación local normalmente no se basa en la búsqueda de genes candidatos a priori. Para ello, los estudios genómicos de adaptación local se sustentan en el hecho de que la diferenciación neutral entre dos poblaciones de una misma especie se distribuye de

forma aproximadamente homogénea a lo largo del genoma, mientras que la selección se manifiesta en forma de picos de diferenciación que sobresalen en el escaneo genómico y que se corresponden con el patrón de adaptación local esperado (*genomic landscapes*; Poelstra et al. 2014). Una vez que se identifican los genes sujetos a selección, se puede inferir la variabilidad fenotípica en la que se manifiesta esta variación genética simplemente atendiendo a la función de dichos genes (aproximación a posteriori; Lee et al. 2008). Sin embargo, este tipo de aproximaciones están muy restringidas en la medida en que dependen del conocimiento de la función de los genes. Esta información depende de la anotación de genes derivada de escaneos genómicos en especies modelo y esto puede suponer un problema a la hora de inferir la funcionalidad de la variación genética en especies filogenéticamente distantes de cualquier modelo bien conocido (Tiffin y Ross-Ibarra 2014). Debido a esto, las conclusiones a las que se llega en estudios con especies no modelo suelen resultar sobre-interpretativos, ya que resulta fácil encontrar genes de interés biológico que pueden ser interpretados a la luz de fuerzas selectivas conocidas (Pavlidis et al. 2012). Este conocimiento incompleto de las funciones génicas sirve, en algunos casos, para reforzar ideas preconcebidas acerca de los rasgos y presiones selectivas que ponen de manifiesto procesos de adaptación local (Barrett y Hoekstra 2011, Tiffin y Ross-Ibarra 2014).

Por otro lado, la genómica de poblaciones puede ayudar también a identificar cuáles son las variables ecológicas más importantes que derivan en procesos de adaptación ecológica (de nuevo *landscape genomics*; Fumagalli et al. 2011). La manera de hacerlo consiste normalmente en correlacionar la variabilidad de determinadas fracciones del genoma sujetas a selección con la variabilidad ambiental que experimentan distintas poblaciones. Una vez conseguido esto, es posible explorar cuáles son los loci que gobiernan la adaptabilidad de una especie a un gradiente ambiental

concreto, lo que permite modelizar qué ambientes dentro de dicho gradiente son aptos para según qué combinaciones de alelos en esos loci (Lotterhos y Whitlock 2015, Rellstab et al. 2015). Además, el aumento de potencia estadística que suponen los escaneos genómicos a gran escala, permite discernir entre la importancia relativa de la distancia geográfica con respecto a la distancia ambiental a la hora de generar estructura genética entre poblaciones muy cercanas (Sexton et al. 2014). Sin embargo, estos estudios correlacionales se enfrentan al desafío que supone identificar las regiones del genoma sujetas a selección, donde realmente estriba la capacidad de los individuos de establecerse y prevalecer en un ambiente determinado (explicado en más detalle más adelante; Dudaniec et al. 2018).

1.1.4. *Limitantes de las técnicas de secuenciación masiva para explorar la base genética de la adaptación local*

Para localizar en el genoma la base genética de los procesos de adaptación local es conveniente la secuenciación y ensamblaje de fragmentos relativamente largos del genoma (Poelstra et al. 2014, Lawson y Petren 2017). Dicho ensamblado debe ser mapeado físicamente en los cromosomas de las especies mediante mapas de ligamiento, de manera que se puedan después "re-secuenciar" las muestras dentro de las regiones de interés. De esta forma, se obtiene un paisaje genómico donde se pueden identificar regiones con altos niveles de diferenciación entre fenotipos o poblaciones con respecto a las posiciones contiguas en el genoma (Campbell et al. 2018). Para conseguir esto, son necesarios altos niveles de calidad de los datos genéticos de partida, algo que con los escaneos genómicos masivos constituye un verdadero reto (Tiffin y Ross-Ibarra 2014). El problema suele estribar en que, aunque obtener una gran colección de secuencias de un individuo es relativamente rápido y sencillo, pasar de esas secuencias cortas a una representación precisa de la variación a lo largo de todo el genoma es muy complicado.

La manera más común de hacerlo es utilizando mapas de ligamiento como se ha explicado anteriormente, pero en muchas ocasiones no se tienen genomas de referencia de especies cercanas con los que alinear dichas secuencias cortas o *reads* (como en el caso de esta tesis). En estos casos, se suelen desarrollar ensamblajes *de novo* con los que las secuencias cortas se alinean entre sí, simplemente con el objetivo de obtener variabilidad genética homóloga comparable a través de las distintas muestras del estudio. Esto hace que la longitud de los ensamblados sea relativamente corta (< 100 pb) y como consecuencia, que sean difíciles de anotar cuando la variabilidad que se encuentra en ellos parece subyacer a algún proceso adaptativo de interés (por lo que muchas veces no se consigue conocer la base genética de dicha adaptación). Además, cuando se trabaja con especies no modelo y no se espera conseguir grandes ensamblajes que anotar en genomas de referencia, se suele optar por datos de baja representación genómica (como el *Genotyping by Sequencing,* GBS, método utilizado en esta tesis). Estos métodos consisten en un muestreo aleatorio de variación genética a lo largo del genoma y resultan muy útiles para ganar potencia estadística con respecto a métodos clásicos de medición de variabilidad genética, además de servir para desvelar patrones de diferenciación genética subyacentes a procesos adaptativos sin atender a la identificación de su base genética (Bradburd et al. 2013).

Otra limitación importante en este tipo de estudios con variabilidad genética deslocalizada pero supuestamente relacionada con procesos adaptativos es la propia interpretación (o incluso correcta identificación) de dicha variabilidad. Esto se debe a que la detección de las fracciones genómicas implicadas utiliza predicciones teóricas acerca de los efectos de la selección natural en las frecuencias alélicas para detectar regiones en el genoma que portan mutaciones adaptativas (Kim y Stephan 2002, Przeworski 2002). Sin embargo, el paisaje genómico de divergencias varía dependiendo

de las historias evolutivas de las poblaciones cuyos patrones de adaptación local se quieren estudiar, lo que supone un reto a la detección de islas de selección a lo largo del genoma. Uno de los escenarios evolutivos que puede dificultar esta tarea consiste en la existencia de divergencia poblacional profunda, de manera que el patrón de diferenciación se acumula durante periodos largos de tiempo. En estos casos, las mutaciones adaptativas se pueden confundir fácilmente con alelos que se han ido fijando a lo largo del tiempo en cada una de las ramas filogenéticas (Rhode et al. 2017, Schield et al. 2017).

Un posible método para identificar loci bajo selección en estos casos es buscar aquellos cuyas frecuencias alélicas se alejan significativamente de lo esperado bajo modelos demográficos concretos (Bohomme et al. 2010). Otra posibilidad, es comparar las frecuencias alélicas de los loci candidatos (definidos a priori) con las de loci de referencia que se sabe que se encuentran bajo selección neutral (Berg y Coop 2014). Cuando esta información no se conoce por carecer de un genoma de referencia, es posible utilizar el mismo fundamento teórico haciendo escaneos a lo largo del genoma para encontrar loci cuyas distancias genéticas entre grupos predefinidos son mayores a las esperadas dado el grado de diferenciación genómico de fondo (cuya inmensa mayoría se espera que constituya variación genética neutral) (Lewontin y Krakauer 1975). De esta forma, se pueden intentar definir los polimorfismos genéticos sobre los que pueda recaer la selección divergente que promueve las diferencias adaptativas observadas. La potencia a la hora de detectar picos de diferenciación genética es menor que en el caso de los paisajes genómicos, ya que no se conoce la posición física de cada variante y resulta imposible encontrar islas de selección con respecto a posiciones adyacentes. El problema de realizar estos análisis con variabilidad genética deslocalizada reside en el hecho de que el grado de diferenciación genética se distribuye

heterogéneamente a lo largo del genoma y a veces resultan indiferenciados loci cuyo grado de divergencia es bajo comparado con todo el genoma, pero que se mostrarían muy divergentes si se pudieran comparar con regiones adyacentes (Campagna et al. 2015). Este escenario es especialmente común en las regiones codificantes, donde reside la base genética de la diversidad fenotípica, por lo que muchas veces resulta imposible identificar la base genética de un determinado fenotipo utilizando aproximaciones con variabilidad genética deslocalizada (Tigano et al. 2017).

En resumen, la aproximación más básica para detectar polimorfismos genéticos bajo selección a lo largo del genoma consiste en determinar la distancia genética debida a la existencia de poblaciones y la contribución parcial a esa distancia de diferentes polimorfismos (Foll y Gaggiotti 2008). Además, es posible encontrar trazas de selección independientes de la configuración filogenética de las poblaciones, como muestran distintos análisis que detectan aquellos polimorfismos cuyas frecuencias alélicas se desvían de lo esperado bajo deriva genética en un escenario filogenético preestablecido (Bonhomme et al. 2010). Si los loci detectados por estas aproximaciones alternativas no solapan, es posible discernir entre la contribución relativa de la innovación genética y de la variación ancestral en los procesos de adaptación local (Barrett y Schluter 2008). De esta manera, se pueden investigar las huellas genéticas de las contingencias pasadas y presentes que configuran la adaptación local como un proceso dinámico e incompleto, no sólo estudiando cómo interactúa con las dinámicas genéticas actuales, sino también infiriendo cómo dicho proceso ha sido favorecido, restringido y/o desdibujado por la historia evolutiva del sistema de estudio (Rundle y Nosil 2005, McEntee et al. 2018).

1.1.5. *Papel de la adaptación local definiendo áreas de distribución: ¿qué puede aportar la genómica?*

Los procesos de adaptación local que rigen la diferenciación adaptativa de las

distintas poblaciones de una especie también subyacen a la configuración del área de distribución de la misma (Kirkpatrick y Barton 1997). Las teorías más aceptadas que explican la presencia de límites a la expansión de dicha área implican saturación de nicho o efectos centro-borde en los que el hábitat o las condiciones demográficas se deterioran hacia los márgenes de la distribución de la especie, hasta llegar a un punto en el que las poblaciones no son viables (Hutchinson 1957, Brown 1984). Sin embargo, los límites del área de distribución de las especies suelen ser más abruptos de lo que se espera en un proceso de adaptación a cambios ambientales graduales (Sexton et al. 2009). Es más, normalmente las especies están adaptadas a rangos ambientales mucho más amplios que lo que supondría dar un paso más allá del límite (Kirkpatrick y Barton 1997). Una posible explicación a este fenómeno es que el flujo genético desde el centro del área de distribución hacia los bordes, contrarreste procesos de adaptación local en las poblaciones marginales, provocando maladaptación y reduciendo las posibilidades de que se produzca una expansión de área (*gene swamping*; Haldane 1956). El estudio de todos estos modelos es crucial para entender las causas adaptativas que configuran la forma y el tamaño de las áreas de distribución (Lee-Waw et al. 2018).

Si el hábitat apto para la especie se mantiene más allá de los límites de distribución, deben existir barreras geográficas, efectos de *gene swamping,* o demográficos que restrinjan la expansión del área (Kirkpatrick et al. 1997, Bridle y Vines 2007, Charlesworth 2009, Peterson et al. 2011). En estos casos, la caracterización genómica de poblaciones que se encuentren distribuidas entre el centro y el borde del área de distribución, puede aclarar las causas de la restricción a la expansión (Lee-Waw et al. 2018). En cuanto al *gene swamping,* por ejemplo, es posible realizar modelos de alta resolución que infieran el grado de flujo genético entre el centro y el borde de la distribución de la especie, de manera que se pueda determinar si esa presión migratoria

se correlaciona con la homogeneización de la diversidad genética sujeta a selección (Beerli y Felsenstein 2001). Por otra parte, si existieran efectos demográficos que restringieran la viabilidad de las poblaciones del margen, esperaríamos encontrar un tamaño efectivo (y una diversidad genética) en dichas poblaciones inversamente proporcional a su distancia al centro de distribución (Delser et al. 2018).

Por otra parte, cuando la saturación de nicho es la mayor causa detrás de la existencia de límites a la expansión, una ampliación del rango de ambientes a los que la especie está adaptada debe preceder a cualquier expansión geográfica (Hutchison 1957). Esa expansión de nicho se encuentra limitada por la magnitud de la varianza genética aditiva asociada con los gradientes ambientales a los que la especie está adaptada (Lande y Shannon 1996). Como se ha explicado anteriormente, resulta relativamente fácil conocer qué partes del genoma se asocian mejor con los gradientes ambientales que experimentan diferentes poblaciones de una especie (Rellstab et al. 2015). Utilizando modelos de asociación genotipo-ambiente, se puede identificar esa varianza genética aditiva y relacionarla con el área de distribución de una especie, de manera que es posible saber qué celdas geográficas son aptas para qué combinaciones de alelos en esos loci (Dudaniec et al. 2018). De esta forma, se pueden llegar a conocer profundamente las dinámicas de variabilidad genómica que subyacen a los umbrales adaptativos responsables de la existencia de los límites de distribución de las especies.

1.1.6. *La fragmentación del hábitat y la erosión de la diversidad genética*

La fragmentación del hábitat no solo se traslada a las poblaciones naturales reduciendo su tamaño efectivo, sino que también disminuye la capacidad de movimiento de los individuos entre las subpoblaciones resultantes que quedan desconectadas, lo que las vuelve vulnerables ante la falta de inmigrantes (Nathan et al. 2017). Estas poblaciones pueden verse afectadas por la pérdida de diversidad genética

asociada a la deriva que sucede a cuellos de botella, un efecto que se ve agravado en las situaciones en las que la inmigración es infrecuente y los individuos tienen que reproducirse con parientes cercanos (Young et al. 1996). Esto acaba desembocando en la disminución de la aptitud media de la población (depresión por endogamia), un efecto muy relacionado con el riesgo de extinción en este tipo de situaciones (Frankham 2005, Wootton y Pfister 2013). Cuanto mayor es la diferencia entre la tasa de fijación de alelos y el rescate genético por migración o la purga genética por selección, mayores son las consecuencias de la depresión por emdogamia causadas por la fragmentación (Hedrick y García-Dorado 2016).

El conocimiento que se tiene acerca de estos procesos suele estar basado en evidencias indirectas (Dixo et al. 2009, Yates y Fraser 2014), por lo que todavía se necesitan observaciones directas de la pérdida de aptitud asociada a la erosión genética causada por la fragmentación de hábitat. Aunque estos efectos han sido identificados en poblaciones naturales que han sufrido fragmentación, en la mayoría de los casos resulta difícil discernir entre el aislamiento vinculado a la historia evolutiva de las poblaciones y el atribuible únicamente al evento de fragmentación reciente (Burbrink 2010, Reed et al. 2011). Para estos casos, resulta especialmente conveniente el aumento de potencia estadística que ofrecen los métodos de secuenciación masiva. Estos métodos pueden llegar no sólo a definir estructura genética entre poblaciones muy cercanas, sino también a inferir (por coalescencia) diferentes tasas de migración entre las poblaciones fragmentadas en distintos puntos de su historia evolutiva (Beerli y Felsenstein 2001). De esta forma se consigue atribuir a la fragmentación reciente del hábitat su efecto preciso en las dinámicas genéticas que subyacen a la depresión por endogamia. Además, gracias a los métodos de detección de variantes genéticas sujetas a selección (sitios con valores de Fst atípicamente altos, u *outliers*) explicados anteriormente, la posibilidad de

definir los loci sujetos a selección hace mucho más fácil la identificación de las regiones genómicas donde recae la pérdida de aptitud por acumulación de mutaciones deletéreas, de manera que se puede relacionar la pérdida de diversidad en esas regiones con una reducción de la tasa de intercambio genético entre poblaciones como consecuencia de la fragmentación de hábitat (Hedrick y García-Dorado 2018).

Por otra parte, se necesitan cinco líneas de evidencia para lograr una verificación integradora de los efectos genéticos negativos de la fragmentación reciente del hábitat: 1) estructura genética entre poblaciones aisladas en fragmentos, 2) oportunidad reducida de cruzamiento con individuos de otros fragmentos, 3) aumento de la homocigosidad individual en los fragmentos, 4) correlaciones negativas entre aptitud y diversidad genética y 5) efectos demográficos negativos relacionados con la homocigosidad. Para explorar estas líneas de evidencia se han venido utilizando diferentes marcadores neutrales que sirven para medir la diversidad genética de los individuos, como las secuencias repetidas en tándem ofrecidas por los microsatélites (Keller y Waller 2002, Johansson et al. 2007). La utilidad de los microsatélites ha sido validada con bases de datos genómicos tanto para la búsqueda de correlaciones entre aptitud y homocigosidad, como para la modelización de estructura genética (Kolbe et al. 2012, Putman y Carbone 2014). Sin embargo, la inferencia de propiedades a escala genómica basada en unos pocos loci de microsatélites puede resultar difícil, ya que los procesos que generan su diversidad son a menudo inciertos, dependen muchas veces de su contexto genómico (pueden estar sesgados por la distribución heterogénea de la variabilidad genética a lo largo del genoma), y además resulta muy común en estudios con especies no modelo que el número de microsatélites sea bajo, lo que supone una falta de potencia estadística que en ocasiones puede resultar grave (Putman y Carbone 2014). Todos estos problemas pueden diagnosticarse -y en su caso descartarse- validando bases de datos de

marcadores clásicos como los microsatélites con bases de datos genómicos, de manera que se pueda comprobar que la información que representan esos marcadores clásicos se correlaciona con la inferida a través de muestreos aleatorios del genoma a gran escala, por tanto no sesgados y con una potencia estadística muy alta (Putman y Carbone 2014).

1.2. Sistema de estudio: la lagartija colilarga (*Psammodromus algirus*)

La lagartija colilarga *Psammodromus algirus* (L.) 1758 es un lacértido (Fam. Lacertidae) de tamaño mediano (longitud cabeza-cloaca del adulto: 60-90 mm; peso: 6–16 g) ampliamente distribuido ocupando un rango de condiciones ambientales extraordinariamente amplio en la península Ibérica (excluida la franja eurosiberiana septentrional), sur de Francia y norte de África. La diversidad genética de la especie está estructurada en dos linajes mitocondriales monofiléticos separados entre hace 3.6 y 3 millones de años (Carranza et al. 2006; Verdú-Ricoy et al. 2010, Fitze et al. 2011). El clado oriental se encuentra sólo dentro de la península Ibérica y sur de Francia, mientras que el occidental está presente además de en la Península, en el norte de África. Estos linajes mitocondriales presentan diferencias fenotípicas heredables congruentes con la heterogeneidad ambiental presente a lo largo de su área de distribución: en general, las lagartijas del Este presentan un patrón dorsal rayado y habitan ambientes predominantemente herbáceos y abiertos mientras que las occidentales presentan un patrón dorsal liso y habitan ambientes boscosos, con mayor cobertura vegetal (Díaz et al. 2017). El carácter adaptativo de estas diferencias fenotípicas radica en el hecho de que los patrones de coloración dorsal parecen resultar más crípticos ante depredadores orientados visualmente en sus respectivos hábitats (Díaz et al. 2017).

Sin embargo, este patrón de diferenciación fenotípica no es completamente congruente con la pertenencia a los diferentes linajes, de manera que parece que el

proceso de adaptación local subyacente es lo suficientemente fuerte como para imponerse al efecto de la historia evolutiva (Díaz et al. 2017). Este hecho implica que la adaptación local reciente en este sistema puede desdibujar los efectos del aislamiento ancestral, promoviendo al mismo tiempo tanto la divergencia adaptativa dentro de un mismo linaje como la convergencia adaptativa con respecto al otro. Por ejemplo, la población de Lerma ocupa un ambiente boscoso fragmentado similar al que se encuentra en la mayoría del área de distribución del clado occidental (Díaz et al. 2005; Santos et al. 2008, 2009; Telleria et al. 2011). Además, las lagartijas de esta población presentan un fenotipo parecido al de las lagartijas del Oeste, sin patrón dorsal rayado, lo que es de esperar en un proceso de adaptación local a los ambientes boscosos donde viven. Sin embargo, el análisis de ADN mitocondrial localiza a las lagartijas de Lerma dentro de un subclado oriental (E2) que está constituido casi exclusivamente por individuos rayados (Díaz et al. 2017). Por tanto, la población de Lerma representa una oportunidad excepcional para investigar cómo los procesos de adaptación local pueden verse afectados tanto por contingencias históricas como por dinámicas genéticas actuales.

Además de su interés desde el punto de vista de la adaptación local, la población de lagartijas localizada en Lerma persiste en un hábitat fragmentado relativamente cerca del borde norte de la distribución de la especie dentro de la península Ibérica (Santos et al. 2008). En la región que ocupa esta población, se encuentran numerosos parches de hábitat teóricamente aptos para el establecimiento de poblaciones de lagartija colilarga pero que sin embargo se encuentran desprovistos de lagartijas (Díaz et al. 2000). Además, a pesar de que en los fragmentos habitados los ciclos vitales de las lagartijas deberían estar adaptados a condiciones ambientales norteñas (o de montaña) (Díaz et al. 2007, Iraeta et al. 2013), los individuos de estas poblaciones muestran un éxito

reproductivo bajo, sobre todo en los fragmentos pequeños (Díaz et al. 2005). Por lo tanto, esta metapoblación también ofrece un buen sistema de estudio en el que preguntarse por qué la fragmentación del hábitat perjudica a la capacidad reproductiva de la lagartija colilarga. De esta forma, se podría explicar la baja aptitud de los individuos instalados en los retazos del bosque original como una consecuencia de la erosión genética por depresión endogámica vinculada al aislamiento de las poblaciones acantonadas en los fragmentos de menor tamaño.

Dentro del linaje mitocondrial occidental, se ha descrito también variabilidad fenotípica que en este caso podría asociarse con gradientes altitudinales en los que las diferencias en temperatura media anual, precipitación y estructura del hábitat podrían generar procesos de adaptación local dentro de este clado. Estas diferencias se encuentran en poblaciones de lagartijas que presentan muy poca diferenciación genética (El Pardo, una población de llanura, y Navacerrada, una población de montaña; Verdú-Ricoy et al. 2010, Díaz et al. 2017) y consisten en una gran variedad de rasgos fenotípicos adaptativos. Uno de estos rasgos, el tamaño corporal de los individuos adultos, está muy relacionado con su aptitud, ya que determina tanto la supervivencia y el atractivo de los machos como la fecundidad de las hembras (Díaz et al. 2005; Martín y Forsman 1997; Iraeta et al. 2013). La ornamentación sexual también está muy relacionada con el éxito reproductivo (y por tanto con la aptitud) y consiste en que los machos de lagartija presentan una coloración rojiza durante los meses de celo que puede ir desde marcas sublabiales poco conspicuas en machos subadultos a ocupar toda la cabeza en los machos más grandes y dominantes. Esta señalización visual se complementa con señalizaciones químicas a través de poros femorales, cuya actividad aumenta a medida que avanza la época de celo, sobre todo en los machos (Díaz et al. 1994). Otros rasgos en los que se diferencian estas poblaciones incluyen: inversión

reproductiva (las hembras de El Pardo ponen menos huevos pero más grandes), tasa de crecimiento (las lagartijas crecen más rápido a mayor altitud, pero en experimentos de *common garden*, las de llanura crecen más rápido; Iraeta et al. 2006 y 2013), tasa de actividad, supervivencia (sobreviven más tiempo las lagartijas de Navacerrada; Iraeta, Salvador, y Díaz 2008), tácticas de escape anti-depredatorias (las hembras de la población de El Pardo escapan haciendo carreras más cortas; Iraeta et al. 2010), señalización química (los machos de la población de llanura tienen más poros femorales y además éstos son más grandes; Iraeta et al. 2011), ornamentación visual (los machos de Navacerrada presentan una coloración rojiza más saturada mientras que los de El Pardo presentan mayor superficie coloreada de rojo; Capítulo 1: Llanos-Garrido et al. 2017) y cargas de ectoparásitos (a pesar de que en ambas poblaciones la presencia de ectoparásitos es alta, sólo infestan a *P. algirus* en Navacerrada; Capítulo 1: Llanos-Garrido et al. 2017). Todas estas diferencias fenotípicas adaptativas (comprobadas en muchos casos empíricamente en experimentos de trasplante recíproco y *common garden*), sumadas a una baja diferenciación genética, hacen de estas poblaciones un muy buen sistema en el que estudiar los patrones genéticos que subyacen a los procesos de adaptación local. Además, el hecho de que una de las diferencias ecológicas entre las poblaciones pudiera afectar de manera directa al desarrollo de ornamentos sexuales (presencia de ectoparásitos), hace que este escenario sea el apropiado para estudiar la posible divergencia en señalización sexual que pueda generar aislamiento reproductivo en este contexto de adaptaciones locales entre poblaciones estrechamente relacionadas.

1.3. Objetivos

Capítulo 1

En este capítulo, se intentan desentrañar qué rasgos de un ornamento sexual

pueden ser empleados como señales honestas de la calidad inmune en dos poblaciones con diferencias eco-inmunológicas muy marcadas: las lagartijas de Navacerrada muestran una elevada prevalencia de ectoparásitos (garrapatas), que no se encuentran parasitando a las de El Pardo a pesar de ser muy abundantes en esa localidad. Este estudio abunda en la caracterización de las diferencias fenotípicas adaptativas entre poblaciones ibéricas de lagartija colilarga. La hipótesis de partida es que si la señalización por medio de un determinado aspecto se encuentra limitada por la presencia de parásitos, otros aspectos ornamentales deberían ser seleccionados para señalizar la calidad inmune.

Capítulo 2

En este capítulo, se pretende desentrañar los patrones de diferenciación genética que subyacen a las diferencias fenotípicas entre las poblaciones de El Pardo y Navacerrada. Aunque estas poblaciones presentan muy poca diferenciación genética, difieren en una gran variedad de rasgos fenotípicos adaptativos (incluidos los documentados en el Capítulo 1), muchos de los cuales se han comprobado estables en experimentos de trasplante recíproco y ambiente común. Por lo tanto, debe haber una base genética que determine esas diferencias fenotípicas. Este capítulo pretende encontrar un patrón de diferenciación a lo largo del genoma que sea congruente con los procesos de adaptación local que mantienen esas dos poblaciones fenotípicamente diferenciadas, a pesar de estar geográfica y filogenéticamente tan próximas.

Capítulo 3

En este capítulo, exploramos cómo tanto la selección equilibradora (que depende del mantenimiento de la variación genética ancestral) como la innovación genética subyacen a los procesos de adaptación local. Lo que se pretende es investigar la fuente

de las variantes genéticas que se seleccionan localmente en poblaciones que se enfrentan, ya sea a retos ecológicos congruentes con su historia filogeográfica, ya sea a otros que no concuerdan con dicha historia. El objetivo en este caso no es identificar la base genética de las adaptaciones existentes, sino entender los medios por los que la evolución provee la variación genética necesaria para que una especie pueda afrontar diferentes retos ecológicos cuando sus poblaciones se encuentran repartidas a lo largo de un gradiente ambiental muy heterogéneo y con extremos muy contrastados.

Capítulo 4

En este capítulo, nos valemos de la distribución geográfica de las combinaciones alélicas de loci sujetos a selección para realizar un modelo de regresión genoma-ambiente y contestar las siguientes preguntas: (1) ¿es posible inferir el área de distribución completa de una especie a partir de ese modelo de regresión genotipo-ambiente? (2) ¿Cómo de importante es la capacidad de dispersarse de una especie en la definición de los límites de distribución de una especie? (3) ¿Hay menos genotipos adaptados a las condiciones marginales del área que al centro de distribución de la especie? Y (4) ¿existen umbrales de adaptabilidad, determinados por la cantidad de varianza genética bajo selección disponible, que limiten la expansión del área?

Capítulo 5

En este capítulo, se pretende documentar los efectos genéticos negativos asociados con la fragmentación antropogénica de un hábitat originalmente continuo. Para ello, intentamos demostrar que la bajada de aptitud de los individuos más homocigóticos de una metapoblación de nuestra especie modelo es atribuible a la erosión genética causada por el aislamiento genético y el aumento de la endogamia en los fragmentos de menor tamaño.

2. MATERIALES Y MÉTODOS GENERALES

2.1. Muestreo de campo y análisis fenotípicos

En esta tesis se han analizado individuos de *P. algirus* capturados en los años 2001, 2002, 2005, 2012, 2014, 2015 y 2016. Durante los meses de máxima actividad de la especie (Marzo-Junio) se capturaron un total de 504 lagartijas colilargas distribuidas entre las cinco poblaciones de estudio de esta tesis (Tabla 1). Las capturas se realizaron con lazo o a mano y las lagartijas fueron llevadas al laboratorio para realizarles distintas medidas morfométricas (longitud cabeza-cloaca, longitud y anchura de la cabeza, longitud de las patas traseras, número de poros femorales, peso) y tomarles fotografías del dorso y de la garganta en posición ventral de las que, una vez procesadas, extraer información sobre coloración dorsal y ornamentación sexual. Para los estudios en los que se requería información sobre parasitismo, el conteo de ninfas de garrapatas (*Ixodes ricinus*) se realizó en el campo, ya que las lagartijas perdían toda su carga parasitaria en el laboratorio.

Para extraer información sobre ornamentación sexual o coloración dorsal de las fotografías de lagartijas, seguimos siempre el siguiente procedimiento. Colocamos a los individuos bajo un cristal antirreflectante inmovilizados contra una esponja blanda. En esa posición, tomamos fotografías de la garganta (ornamentación sexual) o de la espalda (coloración dorsal) en una sala oscura, utilizando un soporte fijo con dos fuentes de luz blanca colocadas a ambos lados de la lagartija, siempre a 25 cm del individuo y con un ángulo de incidencia de la luz de 45º. La cámara, provista de un mecanismo de compensación de la exposición y con una velocidad de obturación de 0.5 ms, se colocó a 35 cm del cristal anti-reflectante.

Una vez realizadas las fotografías, utilizamos Adobe Photshop CS6 (Adobe

Systems, 2002) para procesar las imágenes: estandarizamos el área de análisis utilizando la herramienta de 'lazo magnético' para delimitar la superficie aproximadamente triangular comprendida entre las líneas definidas por el extremo anterior de la cabeza y el extremo posterior de las escamas sublabiales (ornamentación sexual), o la superficie aproximadamente rectangular que incluye la vista dorsal del cuerpo a lo largo de 5 cm desde la línea de los hombros en sentido anteroposterior (coloración dorsal) (véase la Fig. S1 del Capítulo 1). Medimos el área coloreada utilizando la herramienta de la 'varita mágica' y la opción 'similar', con un 30% de tolerancia para la ornamentación sexual y un 10% para la coloración dorsal. En el caso de la coloración sexual de la garganta, medimos la saturación en rojo como la proporción del canal rojo dentro de la tarjeta gráfica RGB, mediante la expresión S = R/(R+G+B), donde S es la saturación y R, G y B son los canales de rojo, verde y azul. Para medir la oscuridad de la raya dorsal, utilizamos el inverso de la luminosidad de la capa coloreada (255 = negro; 0 = blanco).

Tabla 1. Poblaciones estudiadas en cada capítulo, con indicación de los tamaños muestrales correspondientes a cada año.

	POBLACIÓN	AÑO (N)
Capítulo 1	El Pardo	2005 (74) y 2014 (24)
	Navacerrada	2005 (83), 2014 (19) y 2015 (20)
Capítulo 2	El Pardo	2015 (20)
	Navacerrada	2015 (20)
Capítulos 3 y 4	Lerma	2012 (18) y 2016 (18)
	Aranjuez	2015 (20)
	Brihuega	2015 (18)
	El Pardo	2015 (19)
	Navacerrada	2015 (20)
Capítulo 5	Lerma	2001 y 2002 (131) y 2012 (18)

Estas variables fenotípicas se utilizaron en todos los capítulos de la tesis, menos las relativas a ornamentación sexual (solo en los Capítulos 1 y 2) y la coloración dorsal (Capítulo 3).

2.2. Extracción de ADN, secuenciación y genotipado (*variant calling*)

Para los estudios genómicos, obtuvimos muestras de tejido de los 2 cm distales de la cola de las lagartijas. Estas muestras se guardaron en etanol absoluto a 4 ºC hasta la extracción de ADN, que se purificó para la preparación de librerías utilizando el kit de extracción *Speedtools Tissue DNA Extraction kit* ® de Biotools. En dicha purificación, se alargó el paso de lisis celular hasta 24 horas y la resuspensión de ADN se realizó utilizando agua libre de DNAsas calentada a 60 ºC.

Utilizamos la enzima de restricción *Pst1* para la preparación de librerías de GBS (*Genotyping by Sequencing*). El proceso de secuenciación se llevó a cabo en un secuenciador Illumina ® HiSeq2500. Para genotipar los polimorfismos de un solo nucleótido (SNPs) de las secuencias utilizamos UNEAK, una cadena de programas implementada en TASSEL v3.0 (Bradbury et al. 2007) específicamente diseñada para muestras sin genoma de referencia. Las secuencias se alinearon entre ellas formando redes de secuencias en las que cada nodo estaba constituido por una secuencia y cada punta de la red representaba una diferencia entre dos secuencias de una única base. Estas redes se filtraron utilizando umbrales de tasa de error convencionales para eliminar posibles errores de secuenciación. La base de datos resultante estuvo constituida por 83,648 SNPs con una cobertura media de 5.96 y una tasa de secuenciación fallida de 0.49. Después de descartar loci cuyas frecuencias del alelo menos frecuente fueran menores que 0.01 o que no hubieran podido secuenciarse en, al menos, el 10% de los individuos, obtuvimos una base de datos de 73,291 SNPs, con un único SNP por locus, una cobertura media de 6.60 y una tasa de secuenciación fallida de 0.42. Esta fue la base de datos que se analizó para caracterizar la variación genética neutral a lo largo del genoma de los Capítulos 2, 3 y 5 de esta tesis.

2.3. Estructura genética

Antes de realizar los análisis de estructura genética de los Capítulos 2 y 3, usamos el programa PLINK v.1.9. (Purcell et al. 2007) para filtrar la base de datos de SNPs eliminando todos las variantes ligadas entre sí (basándonos en el parámetro de desequilibrio de ligamiento), de acuerdo con los coeficientes de correlación entre variantes. La necesidad de efectuar este filtrado radica en el hecho de que el análisis de estructura genética no tiene en cuenta el desequilibrio de ligamiento, y por lo tanto, los SNPs ligados entre sí pueden sesgar los modelos de agrupamiento resultantes. Utilizamos el programa Admixture v.1.3. (Alexander et al. 2009) para la estimación por máxima verosimilitud de las ancestrías individuales. Además, identificamos el número de grupos (K-value) para el que el modelo tenía más precisión predictiva utilizando el error de validación cruzada (*cross-validation errors*; cv-error). Sin embargo, tanto en el capítulo 2 como en el 3, exploramos modelos de agrupamiento con valores más altos de cv-error para intentar ganar resolución en la detección de estructura genética entre las poblaciones (hasta llegar a K = 2 en el capítulo 2 y a K = 5 en el capítulo 3, el número de poblaciones implicadas en cada estudio).

2.4. Análisis de detección de variantes genéticas sujetas a selección (*outliers*)

Para minimizar la tasa de falsos positivos en los análisis de outliers, aplicamos otro paso de filtrado a nuestra base de datos de SNPs. Con este objetivo, descartamos los loci con frecuencias del alelo menos frecuente menores que 0.05 en cada población, de manera que excluimos los alelos exclusivos de cada población. Además, eliminamos los loci que no fueron secuenciados con éxito en al menos el 75% de los individuos de cada población. De esta manera obtuvimos una base de datos de SNPs de 6,421 loci.

Una vez hecho esto, utilizamos dos análisis de outliers distintos para detectar

variantes genéticas sujetas a selección: un análisis de outliers básico o en bruto (Bayescan v.2.1; Foll y Gaggiotti 2008) y un análisis de outliers independiente de la coancestría de las poblaciones, o filogenéticamente independiente (FLK; Bonhomme et al. 2010). El análisis de Bayescan es una aproximación conservadora, muy poco propensa a los falsos positivos y que además resulta especialmente útil cuando el número de poblaciones es bajo (como en el caso del capítulo 2, donde es el único método utilizado para detectar outliers entre dos poblaciones). Bayescan utiliza un modelo de regresión logística para partir los coeficientes de distancia genética (F_{ST}) en un efecto parcial específico de la población (β) y otro específico del locus (α). Seleccionamos loci con un valor de $\alpha > 0$, lo que sugiere selección positiva, y una tasa de falsos positivos corregida por la realización de múltiples tests de $q \leq 0.05$. Para obtener estos parámetros, utilizamos el algoritmo MCMC implementado en el programa utilizando 20 cadenas piloto de 5,000 iteraciones cada una, seguidas de 100,000 iteraciones (de las que se obtienen los resultados) con un *burn-in* de 50,000.

Por otra parte, el análisis de FLK (independiente de coancestría) es una extensión del test de Lewontin-Krakauer que infiere el patrón de diferenciación histórico entre poblaciones a partir de una matriz de distancias entre poblaciones. Lo que hace después es identificar los outliers a partir de las desviaciones de F_{ST} respecto de lo esperado en un escenario en el que las tasas de fijación de alelos fueran proporcionales a la longitud de las ramas de la genealogía de poblaciones. Seleccionamos loci con un estadístico FLK significativo, con un umbral de significación muy restrictivo de $p \leq 0.001$. Utilizamos los outliers inferidos con Bayescan en el capítulo 2 (detección de patrones de diferenciación genómica entre poblaciones adaptadas localmente) y los que infirieron ambos métodos tanto en el capítulo 3 (para discernir entre la selección equilibradora y la innovación genética

independiente de la ancestría en los procesos de adaptación local) como en el Capítulo 4 (modelos de regresión genotipo-ambiente solo con la variación genética sujeta a selección).

La metodología específica utilizada en cada estudio de esta tesis se explica en el capítulo correspondiente, sin perjuicio de que la descripción de algunos de estos métodos generales pueda repetirse (al menos parcialmente) allí donde dichos métodos se aplican, en aras de la claridad e independencia discursiva de cada uno de los capítulos.

CAPÍTULO 1. Variación en los ornamentos de los machos de dos poblaciones de lagartijas con diferentes cargas de parásitos.[1]

Resumen

En el contexto de la hipótesis del hándicap de inmunocompetencia (que interpreta los ornamentos sexuales dependientes de andrógenos como señales honestas basadas en el efecto inmunosupresor de éstos), exploramos cómo las diferencias en la carga de parásitos afectan a la forma en la que los ornamentos sexuales codifican información acerca de la calidad individual de sus portadores. Estudiamos la variación de las señales sexuales entre dos poblaciones ibéricas de lagartija colilarga (*Psammodromus algirus*), una especie en la que los machos sexualmente activos muestran una coloración rojiza en la cabeza. En una de las poblaciones, los machos siempre se encontraron libres de parasitismo por ninfas de garrapata, mientras que en la otra población, todos los machos estaban parasitados (media de 12.7 ninfas de garrapata por individuo). Al comienzo de la estación reproductiva, la superficie coloreada fue mayor en la población no parasitada, mientras que en la población parasitada la saturación del rojo fue mayor. Gracias a la simulación experimental de una infección bacteriana (mediante la inyección intraperitoneal de lipopolisacárido) pudimos examinar los efectos de la activación inmune en la expresión de este ornamento sexual. En la población no parasitada, este tratamiento causó una reducción en la superficie coloreada de los machos experimentales, mientras que en la población parasitada causó una disminución en la saturación en rojo de la coloración sexual. En la población parasitada,

[1] Este capítulo transcribe íntegramente el siguiente artículo:

Llanos-Garrido, A., Díaz, J. A. , Pérez-Rodríguez, A. y Arriero, E. 2017. Variation in male ornaments in two lizard populations with contrasting parasite loads. Journal of Zoology 303, 218–225.

El estudio se inspiró en los resultados del trabajo de Fin de Máster del Doctorando ('Compromiso entre ornamentación sexual y sistema inmune en Psammodromus algirus'; Máster UCM en Biología Evolutiva, septiembre de 2014) e incluyó materiales del Trabajo de Fin de Máster de Héctor Rivero ('Ornamentación sexual, carga parasitaria e inmunocompetencia en la lagartija colilarga (*Psammodromus algirus*)'; Máster UCM en Zoología, noviembre de 2015).

los machos que mostraban coloración sexual fueron más grandes y tuvieron menos parásitos, y la respuesta inflamatoria a la inyección con lipopolisacárido en la palma de la pata delantera se correlacionó negativamente con la saturación en rojo, pero no con la extensión coloreada. Por lo tanto, sugerimos que los parásitos no sólo restringen la expresión de los ornamentos sexuales, sino que además pueden llegar a cambiar las propiedades específicas de la señal que informan sobre la calidad de sus portadores (en nuestro ejemplo, saturación del color en la población parasitada vs. superficie del mismo en la no parasitada).

Abstract

In the context of the immunocompetence handicap hypothesis, we explored how differences in parasite load affect the way in which sexual ornaments codify information about individual quality. We studied variation of sexual signals in two Iberian populations of the lizard Psammodromus algirus, a species in which sexually active males display a red head coloration. In one population, males were free of tick nymphs, whereas in the other one all males were tick-infested (mean of 12.7 tick nymphs/individual).At the onset of the breeding season, the red-colored surface was larger in the non-parasitized population than in the parasitized one, whereas the opposite was true for color saturation. We experimentally simulated a bacterial infection (by intraperitoneal injection of lipopolysaccharide) to examine the effects of immune activation on the expression of this sexual ornament. In the non-parasitized population, our treatment caused a reduction of the red-colored surface of experimental males, whereas in the parasitized population it caused a decrease in color saturation. In the parasitized population, males that displayed sexual coloration were larger, and had fewer parasites, than uncolored ones, and inflammatory response to lipopolysaccharide injection in the palm of the hind paw was negatively correlated with color saturation,

but not with color extension. Thus, we suggest parasites not only constrained the expression of sexual ornaments, but they also changed the signal properties that conveyed useful information about the quality of their bearers.

Introduction

The handicap hypothesis was proposed by Zahavi (1975) to explain the evolution and maintenance of costly sexual ornaments as honest signals of genetic quality. Since then, many researchers have expanded good-gene models to show how exaggerated sexual signals advertise overall heritable fitness (Iwasa et al 1991, Maynard-Smith 1991, Hill 1991). The basic idea is that the development of sexual ornaments may be involved in trade-offs with different aspects of organismal performance such as predation risk (Giery and Layman 2015), locomotor efficiency (Møller and Lope 1994), or immune system activation (Hamilton and Zuk 1982), in such way that only high quality individuals would be able to express the ornaments without compromising their fitness. Concerning the hypothesis that male ornaments are reliable signals of genetic resistance to parasites and diseases, testosterone was subsequently proposed as a link between sexually selected traits and associated immune costs: only immunocompetent males would be able to increase their testosterone levels, thereby deviating resources from immunity towards the production of an attractive signal, while controlling their pathogen load (Folstad and Karter 1992).

Previous research has supported the immunocompetence handicap hypothesis in a wide variety of organisms and signals: tail length in barn swallows (Møller 1998), song rate of starlings (Casagrande et al. 2015), chemical signals of burying beetles (Chemnitz et al. 2015) and rock lizards (Martín and López 2014), or sexual coloration of fish (Milinski and Bakker 1990), lizards (Salvador et al. 1996), or birds (Blount et al.

2003, Zuk et al. 1990). Most of these studies are based either on correlational approaches showing negative associations between immune system and ornamentation, or on experiments that manipulate either the intensity of sexual ornaments or the activation of the immune system (Zuk et al. 1990, Salvador et al. 1996, Møller 1998; Casagrande et al. 2015). However, few studies have taken advantage of ecological differences among populations of the same species to investigate the selective pressures that affect the evolution of sexual ornaments (Endler and Houde 1995, Giery and Layman 2015).

Here, we examine inter-population differences in sexual signals potentially mediated by parasite load. We test the hypothesis that differences between a non-parasitized and a heavily parasitized population can change the signal features that convey useful information about immune quality. In our model organism, the lacertid lizard *Psammodromus algirus*, males display a red coloration in the head during the breeding season whose degree of enlargement has been related to body size (Díaz 1993), testosterone levels (Díaz et al. 1994, Salvador et al. 1996), and reproductive success (Díaz 1993, Martín and Forsman 1999). We searched for potentially correlated differences in parasite load and sexual signals between two populations separated by a 600-m elevational gradient. At low-elevation, males were free of *Ixodes sp*. tick nymphs, and they showed a more conspicuous sexual coloration. At high-elevation, parasite load was high, and sexual coloration was less conspicuous. The reason why only high-elevation lizards were parasitized, despite ticks being also present at low-elevation, remains unknown, but it might be related to differences in the availability of alternative intermediate hosts (Casher et al. 2002). Both immunocompetence handicap and the importance of parasitism by tick nymphs of the genus *Ixodes* were previously shown in *P. algirus* by Salvador et al. (1996), in an experiment in which testosterone-

implanted males increased the conspicuousness of the red coloration of the head, but suffered a higher parasite load (associated with higher mortality) than control males.

In this study, we elicited an immune response in male lizards from both populations by intraperitoneal injection of a bacterial lipopolysaccharide solution (hereafter LPS), which simulates a bacterial infection. We aimed to unravel what particular traits of the sexual ornament could be employed as an honest signal of immune quality in each population. We discuss our results in the light of the hypothesis that, if signaling by means of the extension of colored area is limited by parasites, then other aspects of the ornament should be selected to signal immune quality.

Material and Methods

Study species and study areas

Psammodromus algirus (Linnaeus, 1758) is a medium-sized (snout-vent length 60-90 mm; mass 6-16 g) lacertid lizard that occupies shrub and woodland Mediterranean habitats. Males display a reddish coloration on the head during the breeding season (April-June; Veiga and Salvador 2001), which ranges from relatively inconspicuous and restricted to the posterior supralabial or infralabial scales in subordinate lizards, to a brilliant patch on the sides of head, mental scutes and throat in dominant, large individuals (Díaz 1993).

Our first study area was located at 'El Pardo' (Spain: 40º31'N, 03º47'W; 650 m elevation). This site is a holm oak (*Quercus ilex* L.) forest in which offshoots of *Q. ilex* dominate the shrub layer together with open areas covered by annual herbs. The second site was located at Navacerrada (Spain: 40º44' N, 4º00' W; 1300 m elevation), at a linear distance of 32 km from the first one. This area is a deciduous Pyrenean oak (*Quercus pyrenaica* Willd.) forest with a high cover of shrub patches. Lizards from both

populations show no apparent genetic differentiation (Díaz et al. 2016), but they differ in phenotypic traits such as escape-tactics, sexual dimorphism, and life history (Iraeta et al. 2010, 2011 and 2013).

Collection and husbandry of animals

Sample size consists of 39 males from high elevation, captured in 2014 and 2015, and 24 males from low elevation captured in 2014. Lizards were transferred to the lab, measured (snout-vent length [SVL], head length, head width, and average length of hindlimbs), weighed (body mass), and individually caged in terraria (40 x 60 cm and 30 cm high) in a room with natural photoperiod. Heat was supplied by a spotlight bulb, which created a thermal gradient allowing lizards to thermoregulate within their preferred temperature range. Food (crickets *Acheta domestica*, sprinkled with a commercial diet supplement) and water were supplied *ad libitum*.

Tick infestation

We used a database containing information on 157adults of both sexes captured in 2005 to analyze interpopulation, seasonal, and sexual variation in the prevalence and intensity of infestation by nymph ticks (identified as *Ixodes* spp. by Salvador et al. 1996). This database, which includes two sampling periods, one in the early (March 30 - April 26) and the other one in the mid-late breeding season (May 11 - June 7), has already been used in previous studies (Iraeta et al. 2006, 2011 and 2013). However, tick data have not been published before. Prevalence was defined as the percentage of individuals with at least one tick on their body surface, and parasite intensity was defined as the number of ticks per infested individual.

Immune system activation and measurement of inflammatory response

In 2014, we performed an experiment to examine the effects of immune system activation on the expression of sexual coloration at the peak (early May) and end (late

June) of the breeding season. Two subgroups (experimental and control) were generated at random within each of the two populations. Lizards in the experimental subgroups (N = 8 and 11 for high- and low-elevation, respectively) were intraperitoneally injected a dose of 2.5 mg of LPS of *Escherichia coli* serotype O111: B4, diluted in 0.05 ml saline solution per gram of body mass. Control individuals (N = 6 and 8 for high- and low-elevation, respectively) were injected with the same volume of saline solution. This procedure is similar to those previously employed with other lizards (López et al. 2009; Uller et al. 2006). Experimental and control males did not differ significantly in colored surface, color saturation, structural size, or physical condition, even after controlling for the effects of population of origin or date of capture (all *P*'s > 0.12). To quantify the effects of immune system activation on sexual coloration we examined two digital images (see 'Sexual coloration' below) taken immediately before and two weeks after injection of the antigen.

In 2015, lizards were photographed upon arrival to the lab to measure the colored surface and color saturation, as explained in the following section. We subsequently estimated inflammatory response by subcutaneously injecting all males with 0.1 mg of LPS diluted in 0.01 ml of serum in the palm of the left hind paw (see Zamora-Camacho et al. 2014 for a similar procedure). We used a caliper to measure the thickness of the palm just before injection of LPS and four hours later, i.e. at the moment of the expected peak of immune system activation. In 19 of 20 males, the thickness of the inoculated palm increased 4 h after the subcutaneous administration of LPS (repeated measurements ANOVA: $F_{1,19} = 52.72$, $P < 0.001$). Previous studies have shown that the thickness of the inflammation is associated with the strength of the immune response (Parmentier et al. 1998; Zamora-Camacho et al. 2014).

Sexual coloration

Coloration data were obtained from digital images of the ventral view of the head, throat and neck of each male, with all individuals immobilized in the same position. Pictures were taken in a dark room with a table supplemented by two side light sources at the same distance from the lizard and a holder to fix the camera at a standardized position. Although this method does not allow detecting ultraviolet components that could be present in the sexual signal, we chose to obtain coloration data from digital images instead of performing live measures with a spectrophotometer because of the lack of repeatability of spectrophotometry in non-static surfaces and the difficulties in characterizing irregularly distributed colour patches (Stevens et al. 2007).

We used Adobe Photoshop CS6 for image processing (see Fig. S1 in Supplementary Material). We standardized the analyzed area using the 'magnetic lasso' tool to delimit the surface comprised within the lines defined by the snout and the posterior edges of the infralabial scutes. The extent of red colored surface was measured with the 'magic wand' tool (at 30% tolerance) after randomly clicking at the middle of the red area. We subsequently used the 'similar' option of the magic wand tool (at 30% tolerance) to select all areas with similar RGB values, and we measured colored surface as the percentage of colored pixels in the analyzed area. We calculated color saturation of the colored surface using the ratio R/(R+G+B), where R, G and B are the red, green and blue channels of the graphics card; red saturation is 100% if such ratio is equal to 1. Color measurements were taken blindly with respect to population of origin and treatment.

Statistical analyses

We checked the assumptions of parametric tests and, when necessary, we log- or arcsine-transformed the variables. We analyzed the 2014 data using ANOVAs with the

difference between color measurements before and two weeks after treatment as the dependent variable, and treatment, population and date as the categorical predictors. We used contrast analysis (i.e. planned comparisons) to test the significance of a predicted specific effect within our larger statistical design, namely that males from both populations should differ in their responses to LPS injection while controlling for the effects of date (in other words, that the effects of the interaction between treatment and population on color variation should be significant).

We estimated the structural size of each male by computing its score on a Principal Component (hereafter PCsize) that combined all size measurements (SVL, head length, head width, and average hind limb length; retained variance = 0.785). Physical condition was estimated using the residuals of the regression of log-body mass on log-SVL.

The effect of parasite load on the sexual coloration of high elevation males was analyzed using a model selection approach based on the Akaike information criterion. Because 12 of 20 males had no sexual coloration at all, we used a logistic regression (generalized linear model with logit as the link function) with presence or absence of sexual coloration as the binomial response variable, and structural size, body condition, and parasite load as continuous predictors.

Finally, we employed a partial least squares regression (PLS) to examine variation in the inflammatory response. Extension of colored surface, color saturation, together with parasite load and morphological traits (size and condition) were used as predictors. All statistical analyses were executed with the Statistica (StatSoft) software package.

Results

Tick infestation levels and associated ecoimmunological scenarios

Tick nymphs were detected only on high-elevation lizards (data from the 2005 sample: overall prevalence of 78.3% and 0.0% at high- and low-elevation, with 95% confidence intervals of 67.6 – 86.3% and 0 – 6.1%, respectively; χ^2 = 98.9, df = 1, P < 0.001). Parasite load was on average 7.9 nymphs per infested individual. Prevalence increased during the breeding season, from 62.8% in March-April to 95.0% in May-June; χ^2 = 12.7, df = 1; P < 0.001). Parasite load was higher for males than for females, particularly at the end of the breeding season, when there was a dramatic increase in the mean number of nymphs carried by males (Fig. 1).

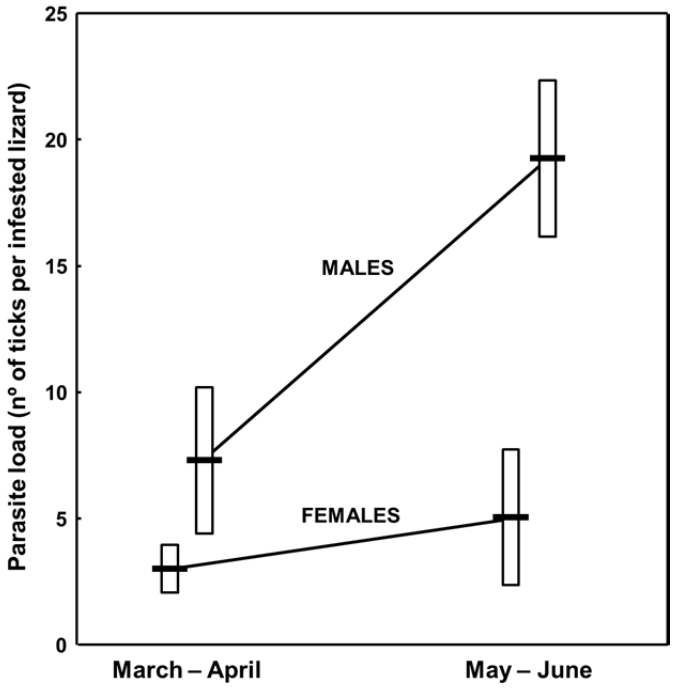


Figure 1. Parasite load (mean number of ticks per infested lizard ± 95% confidence interval) for male and female lizards of the parasitized, high-elevation population at the beginning (March-April) and end (May-June) of the breeding season.

In 2015, on a sample of 20 males captured in late June, tick prevalence was of 100%, but load (8.2 mites / individual) was lower than in May-June 2005. By the end of the breeding season, and pooling together the 2005 and 2015 samples, overall parasite

intensity was 12.7 tick nymphs per high-elevation male (*vs* no tick nymphs at low elevation). Occasional visits between 2005 and 2015 confirmed that whereas low-elevation lizards never carry tick nymphs, most lizards are parasitized at high elevation.

Inter-population differences

In May, the red colored area was greater in males of the non-parasitized population than in males of the parasitized one ($F_{1,18}$ = 5.63, P = 0.029; non-parasitized, mean ± SD = 35.1 ± 20.8 %; parasitized: 13.8 ± 17.8 %), whereas the opposite was true for color saturation ($F_{1,18}$ = 6.58, P = 0.019; non-parasitized: 54.5 ± 3.7 %; parasitized: 60.0 ± 5.8 %). Color saturation of males of the non-parasitized population increased along the breeding season, rising from 54.5 ± 3.7 % in early May to 61.4 ± 4.8 % in late June (two-way ANOVA with date and population as factors; date: $F_{1,39}$ = 9.15, P = 0.004; population: $F_{1,39}$ = 3.82, P = 0.058; interaction: $F_{1,39}$ = 4.17, P = 0.048). The surface of the colored area did not vary significantly between May and June in either population ($F_{1,39}$ = 0.34, P = 0.561).

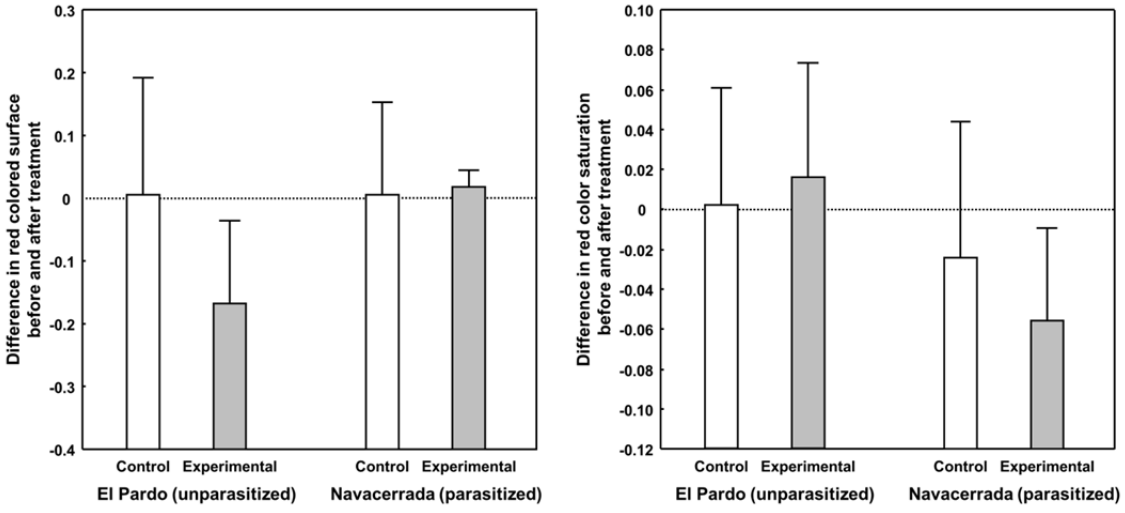


Figure 2. Differences between measurements of red-colored surface (left) and red saturation (right) taken before (measurement 1) and two weeks after (measurement 2) treatment for males from the parasitized (El Pardo) and non-parasitized (Navacerrada) populations: mean values of the difference measurement 2 – measurement 1 (± 95% confidence interval). Control males (white bars) were injected with saline solution whereas experimental males (gray bars) were injected with LPS.

Experimental males of the non-parasitized population responded to the intraperitoneal injection of LPS by reducing the red-colored surface on their heads, whereas experimental males of the parasitized population did not show such response (Fig.2, planned comparison for the interaction between the effects of treatment and population: $F_{1,25} = 5.77$, $P = 0.024$). On the other hand, experimental males of the parasitized population responded to LPS with a significant reduction of color saturation, whereas experimental males of the non-parasitized population did not show such response (Fig. 2, planned comparison: $F_{1,25} = 5.57$, $P = 0.026$).

Parasite-mediated trade-offs in high-elevation males

Our model selection approach with the 2015 data produced a logistic regression (Table 1) in which the ability of males to express sexual coloration was predicted by structural size (log-likelihood = -12.57, $\chi^2 = 25.1$, $df = 1$, $P < 0.001$) and parasite load (log-likelihood = -4.84, $\chi^2 = 9.7$, $df = 1$, $P = 0.002$). Consistently, males that expressed

Table 1. Model selection approach to predict the presence or absence of sexual coloration in males of the parasitized population (binomially distributed response variable) as a function of structural size, physical condition, and parasite load. AIC = Akaike Information Criterion.

Variables in model	*AIC*	*Δ AIC*	*Log-likelihood* χ^2
Structural size - Parasite load	6.27	--	24.65
Physical condition - Structural size - Parasite load	8.00	1.73	26.92
Structural size	13.83	7.56	17.09
Physical condition - Structural size	15.67	9.40	17.25
Physical condition	30.37	24.10	0.55
Parasite load	30.50	24.23	0.42
Physical condition - Parasite load	31.14	24.87	1.79

sexual coloration were larger than those that did not show head coloration (mean SVL ± SD = 73.4 ± 2.6 mm and 68.2 ± 3.1 mm, respectively; $F_{1,18} = 15.11$, $P = 0.001$). For a given structural size, colored males had less parasites (mean number of ticks adjusted

for body size ± SD: 1.6 ± 2.5 ticks per individual) than non-colored ones (12.1± 7.7 ticks per individual; ANCOVA: $F_{1,17}$ = 5.85, P = 0.027), suggesting that parasite load negatively affected the ability of lizards to display sexual coloration.

The PLS regression analyzing colored surface and color saturation as predictors of inflammation produced a single factor that explained 40.4% of the variance in swelling ($F_{1,18}$= 12.22, P = 0.003; Fig. 3). Predictor weights showed that physical condition and color saturation, but not structural size, parasite load or colored surface, explained a significant amount of variation in the inflammatory response induced by the injection of LPS (Fig. 3). Thus, inflammatory response was higher for males with a better body condition (simple correlation: r = 0.535, df = 18, P = 0.015). Controlling for body condition, the inflammatory response was negatively correlated with color saturation (partial correlation: r = -0.565, df = 17, P = 0.012), but not with colored surface (partial correlation: r = -0.128, df = 17, P = 0.601).

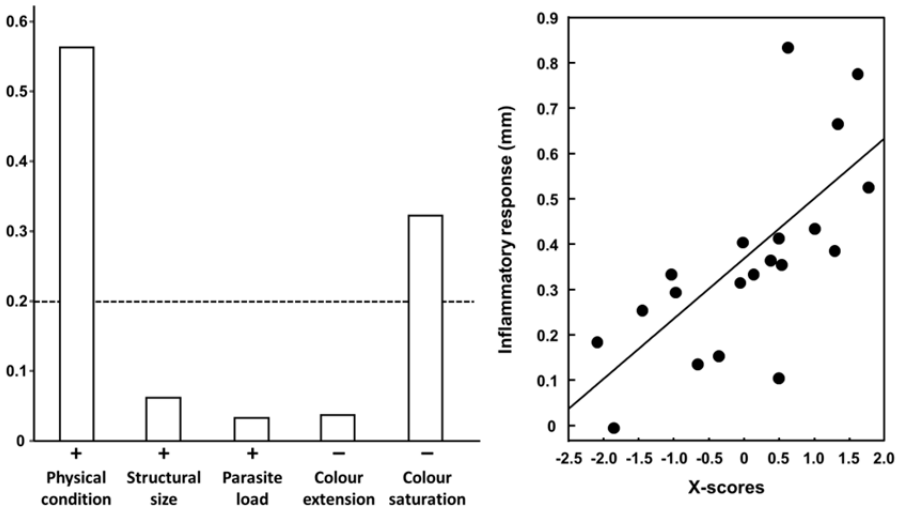


Figure 3. Results of a Partial Least Squares analysis (PLS) used to test for the relationship between the phenotypic variables considered (physical condition, structural size, parasite load, extension of colored surface, and color saturation) and the inflammatory response of male lizards to the subcutaneous injection of LPS. Left: relative contribution of the phenotypic variables to the multivariate factor (X) predicting inflammatory response; variables with squared weights > 0.20 (dashed line) are significant. Right: scatterplot depicting the regression of inflammatory response on X-scores.

Discussion

Our 2005 data showed that parasite load was higher for males than for females, as it has been reported in other lizards (Brace et al. 2015, Reedy et al. 2016). Males of *P. algirus* have larger home ranges than females, and larger and more colored males display greater activity to improve their reproductive success (Díaz 1993). As a consequence, their probability of becoming infested with tick nymphs should also be higher, what would lead to an increased parasite load by the end of the breeding season. Higher parasite loads in males than in females could be explained by the effects of testosterone, which increases activity and compromises immunocompetence (Folstad and Karter 1992, Salvador et al. 1996, Belliure et al. 2004). Circulating testosterone levels remain high until the end of the breeding season, as spermatogenesis proceeds and sexual ornaments enlarge (Díaz et al. 1994). All these processes reach a peak between April and mid-June, when the area of sexual coloration is positively correlated with circulating testosterone levels in lizards from the non-parasitized population (Díaz et al. 1994). Also, tick nymph load reaches a maximum in males of the parasitized population at the end of the breeding season, especially if their testosterone levels are experimentally increased (Salvador et al. 1996). Thus, the concomitant phenology of tick nymph infestation and color enlargement supports the trade-off between sexual ornamentation and parasite load that has been described in these lizards (Salvador et al. 1996) and other taxa (Malo et al. 2008, Pollock et al 2012).

Within this context of ecoimmunological differences, our experimental approach suggests that the color trait that is traded-off against immune system activation varies between non-parasitized and parasitized populations. Our results support this interpretation in two ways. Firstly, and foremost, in the non-parasitized population experimental males responded to LPS injection by reducing the extent of the red-

colored surface, whereas in the parasitized population they reduced color saturation. Secondly, these results were consistent with inter-population differences at the beginning of the breeding season, when the effects of sexual signals on mating success should be maximal. Thus, the features that were traded-off against the immunological activation caused by LPS could be those employed as sexual signals in each population (Hamilton and Zuk 1982, Jacot et al. 2005).

How do parasites influence these different responses? The sexual ornament was more conspicuous in low-elevation males, with no parasites, than in parasitized, high-elevation ones, which suggests that parasites limit the amount of resources available for expressing a color signal of a certain size (Kekalainen et al. 2014). This would explain why lizards of the non-parasitized population employ colored surface as an honest signal of male quality. The signal is effective because, in the absence of tick nymphs, it has enough inter-individual variation to remain informative, a necessary condition for the evolution of costly ornaments (Delhey and Peters 2008). And it is honest because it is traded-off with immune response, as demonstrated by its consistent decrease in experimental males, whichindicates that males may have trouble maintaining a large colored area and simultaneously mounting an efficient immune response. The males that succeeded to produce a red-colored sexual signal were larger (i.e. they were older and/or had grown faster) than those that did not display it; and, for a given body size, colored males had less parasites than uncolored ones. This is what would be expected if parasites led males to allocate to the immune system part of the resources they would otherwise invest in sexual coloration. This could also explain why color saturation, rather than colored surface, could be used to signal male quality (as suggested by the decrease in saturation found in LPS-injected males but not in control ones).

Our reasoning is based on the assumption that, in the face of a parasitic

infection, it is more demanding to grow a larger colored surface than to increase the saturation of a relatively smaller one. This assumption is consistent with the difference in the average extension of colored surface between males of both populations (equivalent to a 2.5-fold increase), that was larger than the same difference in saturation (equivalent to a 1.1-fold increase). Also, the effect of the LPS treatment was higher for the extent of colored area in the non-parasitized population, which experienced a 1.9-fold decrease, than for color saturation in the parasitized one, which experienced a 1.1-fold decrease. Thus, we suggest that males express their sexual coloration at the beginning of the breeding season as much as they can, and subsequently increase the saturation of the colored surface. When released from parasite load, they use colored surface as a signal, and they eventually increase color saturation with the remaining resources, perhaps in the face of new matings (second clutches are common at low-elevation; Díaz et al. 2007). However, when lizards are challenged by parasites and fail to grow a large colored surface, as in the tick-infested population, they use saturation as the main signaling trait from the beginning of the breeding season. Thus, after controlling for the effects of body condition, inflammatory response was negatively correlated with color saturation, suggesting that parasitized males had to trade-off their already compromised resources between a more attractive sexual ornament and a more effective immune response.

Finally, it should be stressed that, because our study compares only two sites that differ in altitude, inferences drawn from our results are, in a strict sense, restricted to the two sites used. However, our two-sites comparison may be representative of altitudinal effects (tick-parasitized lizards are widespread at montane oak forests, and very scarce at lowland holm oak forests; authors, personal observation), and our results may suggest general patterns that could be tested with future work. Also, other ecological differences

between sites, such as habitat visibility (Endler and Houde 1995), may have contributed to shape sexual ornaments. For example, mean distance between shrubs is larger, and cover of open areas is higher, at lower elevation (Iraeta et al. 2010), and this may have favored color surface, which can be perceived from a longer distance, as a sexual signal. At high-elevation, however, differences in red saturation would be effective only from a short distance.

In summary, we suggest that parasites may constrain the expression of sexual ornaments to the extent of regulating what trait characteristics are most informative about individual quality. Thus, parasites like tick nymphs, independently of their effects on individual survival (but see Salvador et al. 1996), may have a significant impact on individual fitness by influencing the expression of key sexual ornaments. Although our work suggests that the color features measured are reliable signals of male quality, experimental approaches would be needed to clarify the link between parasites and sexual selection (female preferences and competition among males). Furthermore, we encourage future studies on other organisms in which ecoimmunological differences might be affecting signaling systems. Because parasite-host coevolution is widespread in life, this approach could shed more light on the role of parasites as modulators of the phenotypic effects of sexual selection, including qualitative changes in the traits involved in signaling systems.

Supplementary material

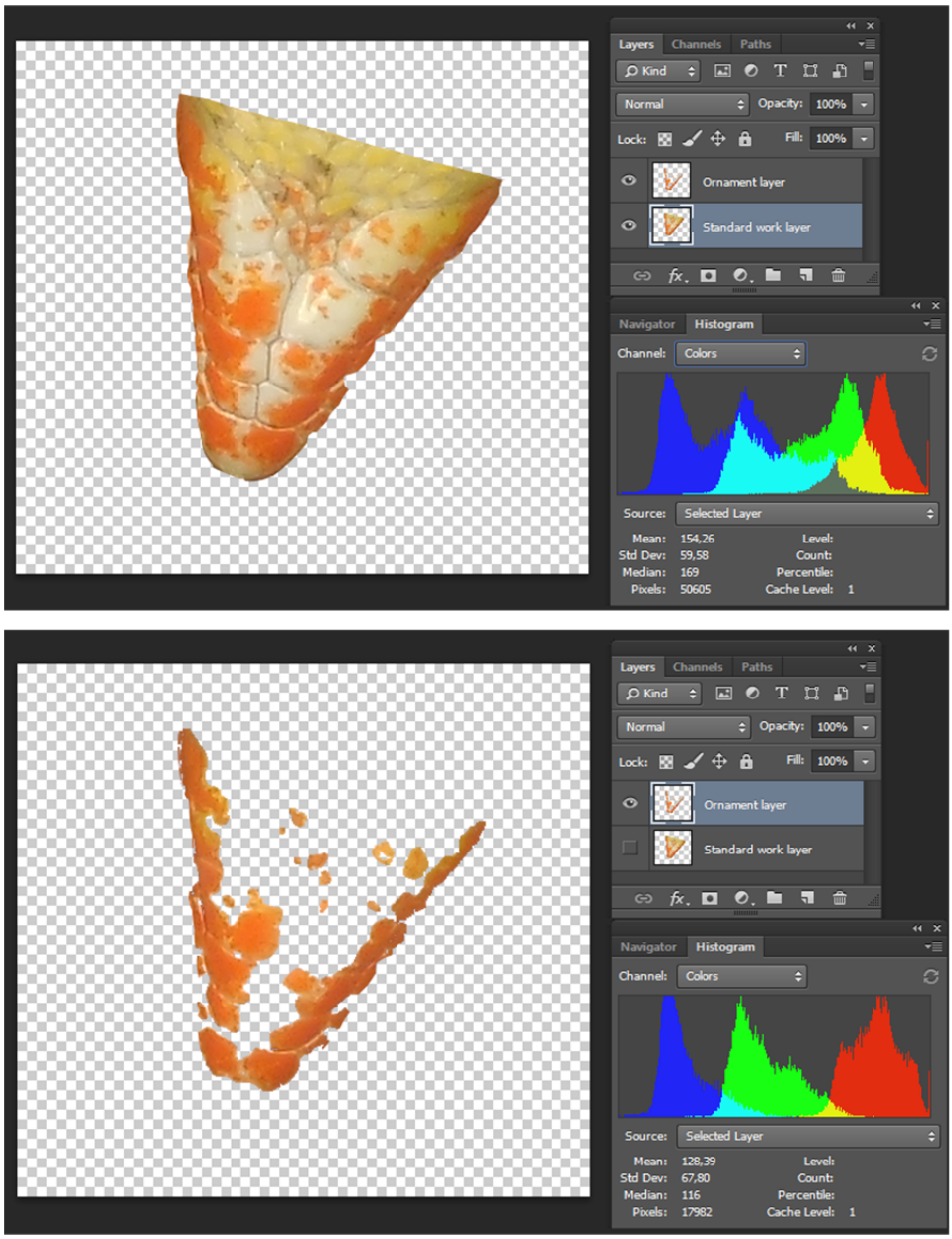


Figure S1. Key steps in Photoshop CS6 image processing. Above: standardized area for the analysis of ventral coloration of the head, obtained using the magnetic lasso tool to select the surface comprised within the triangle defined by the snout and the posterior edges of the infralabial scutes. Below: extent of red colored surface, measured using the magic wand tool. We used histograms to count the numbers of pixels in the appropriate layers, which allowed us to determine the overall amount of skin surface involved in the production of the sexual ornament, and to obtain the average value of the R, G, and B channels of the graphics card to calculate color saturation as R / (R+G+B).

CAPÍTULO 2: Baja divergencia genómica entre dos poblaciones de lagartija colilarga con marcada diferenciación fenotípica adaptativa.[2]

Resumen

La base genética de los procesos de adaptación local que surgen como consecuencia de la heterogeneidad ambiental a lo largo del área de distribución de una especie puede explorarse mediante escaneos genómicos en busca de variantes genéticas sometidas a selección. En general, se espera que el patrón de diferenciación genética sea congruente con el patrón de diferenciación fenotípica vinculado a los procesos de adaptación local, pero puede también ocurrir que la diferenciación genética de fondo sea baja, porque el flujo genético es intenso y/o porque la divergencia es muy reciente. En este capítulo, se estudia la diferenciación a escala genómica entre dos poblaciones de *Psammodromus algirus* que muestran múltiples y sólidas evidencias de estar adaptadas localmente a un gradiente altitudinal, aunque su divergencia genética no ha podido ser detectada utilizando datos de mtDNA y 7 loci de microsatélites. Nos valimos de herramientas de detección de SNPs con frecuencias alélicas indicativas de selección divergente para buscar una explicación genética al patrón de divergencia fenotípica observado, tratando de encontrar loci cuyo grado de divergencia fuera significativamente mayor que la diferenciación genética de fondo calculada a lo largo de todo el genoma (i.e., diferenciación genética neutral). Los resultados de dicho análisis no encontraron ningún locus putativamente subyacente a la divergencia adaptativa entre ambas poblaciones, confirmando su reducida diferenciación genética con muchísimo mayor potencia de análisis (gracias al estudio de más de 70.000 SNPs). La falta de resultados concluyentes en este trabajo pone de manifiesto las limitaciones del método de secuenciación masiva empleado (GBS) para la detección de sitios sometidos a

[2] Basado en el siguiente artículo: Llanos-Garrido, A., Pérez-Tris, J. y Díaz, J. A. Low genomic divergence between two lizard populations with high adaptive phenotypic differentiation. En preparación.

selección divergente entre poblaciones filogenéticamente muy próximas. Así, el genotipado mediante GBS en especies no-modelo dificulta mucho localizar en el genoma la variación genética encontrada, impidiendo el estudio de las divergencias genéticas a escala local. Además, este método de secuenciación tiene una potencia de muestreo menor de la que posiblemente necesitaríamos para encontrar los pocos SNPs divergentes supuestamente responsables de la diferenciación adaptativa entre estas dos poblaciones, por lo demás genéticamente indiferenciadas. Se discuten perspectivas de análisis más prometedoras para este tipo de estudios en general, y en particular para avanzar en el conocimiento de las bases genéticas de la diferenciación fenótipica de la lagartija colilarga.

Abstract

The genetic basis of local adaptation to environmental variation across a species' distribution range can be addressed through genome-wide methods of outlier detection (where outliers are genetic variants whose frequencies reveal the effects of natural selection). Generally, we should expect that phenotypic differentiation in adaptive traits should be paralleled by genetic divergence between populations. However, genetic divergence could be lower than expected due to intense gene flow and/or to remarkably recent differentiation. In this chapter we explore genome-wide traces of divergence between two populations separated by a 600 m elevational gradient; these populations seem to be differentially adapted to their environments despite low levels of genetic differentiation (according to mtDNA and microsatellite data). We performed a search for outliers (i.e. loci subject to selection) trying to identify specific loci with FST statistics significantly higher than those expected on the basis of overall, genome-wide estimates of genetic divergence (i.e. background genome-wide neutral differentiation). Our analyses failed to reveal any locus that could underlie the observed patterns of

adaptive divergence; instead, they confirmed the lack of differentiation on the basis of more than 70,000 SNPs. Perhaps the lack of positive results can be related to the drawbacks of our method of massive sequencing (GBS), because genotyping by sequencing in non-model species complicates the location of genetic variants in the genome and hampers the identification of genetic divergence on a local scale. Furthermore, this sequencing method probably lacks the power required to find the few outliers that could underlie adaptive differences between two populations whose overall level of differentiation is otherwise very low.

Introduction

Environmental heterogeneity throughout the range of a species can promote, through natural selection, the emergence of locally adapted phenotypes that can vary among different populations (Shaner et al. 2015). In general, the genotypes and phenotypes of these populations are expected to be spatially congruent, i.e. the genetic differentiation between populations should follow the same pattern as the adaptive phenotypic differentiation (Tigano and Friesen 2016). However, when there are no barriers to gene flow or when divergence between populations is very recent, discordances may appear within that pattern, whereby two populations may be genetically undifferentiated and, at the same time, show remarked phenotypic differences (Moody et al. 2015, Palmer and Kronforst 2015, Shaner et al. 2015). In these cases, it is theoretically easy to find the genetic basis of the phenotypes that distinguish them through large-scale genomic scans (Poelstra et al 2014). These methods offer the opportunity to identify fractions of the genome whose divergence is significant with respect to an undifferentiated genomic background (Safran et al 2016). These divergent regions can be attributed to the action of natural selection on the basis of their allelic frequencies, typically deviated from the homogeneous background which is characteristic of a genome that has not yet been

differentiated (Campagna. et al. 2015)

The hybridization of grey and black crow along central Europe is a classic example of genomic characterization of phenotypic divergence based on these principles of genome-wide differentiation (Poelstra et al. 2014). In this work, it was possible to identify genomic-scale introgressions that extended far beyond the morphological hybrid zone (the region in which crows show a phenotypic mixture between grey and black phenotypes). However, there were clear-cut differences in the transcriptomic profile of divergent phenotypes that were concentrated in pigmentation genes, even within the same individual (black feather follicles vs. grey feather follicles). In this paradigmatic study, only a small number of 'genomic islands' were found to be resistant to gene flow between the two phenotypes. This work showed how localized selection in very specific regions of the genome can promote phenotypic divergence even under scenarios of continuous, uninterrupted gene flow (Moody et al. 2015, Palmer and Kronforst 2015, Shaner et al. 2015).

When the genomic background of the study populations is actually differentiated, it is important to take geographical isolation into account in order to locate the divergence apparently underlying local adaptation (Campagna et al. 2015). For that purpose, genomic studies on adaptive divergence rely on the fact that geographic isolation generates an approximately homogeneous pattern of divergence along the genome, while selection gives rise to differentiation peaks that stand out in genomic scanning and that are congruent with the expected pattern of local adaptation (Poelstra et al. 2014). Moreover, when local adaptation is favored by barriers that restrict gene flow, such barriers generate a similar genomic pattern of divergence peaks that can be detected with the same approach (Ravinet et al. 2017). However, all these patterns of genomic heterogeneity can be confounded by several factors that make

adaptive divergence studies challenging: genomic background selection, selective waves, recombination, variation in mutation rate or variation in gene density (Campagna et al. 2015, Ellegren et al. 2012). Separating the effects of divergent selection from all these confounding factors represents one of the most important challenges for studies of local adaptation on a genomic scale (Campagna et al. 2015).

For such studies to be successful, it is desirable to sequence and assemble relatively long fragments of the genome (about 1 Gb, depending on the size of each species' genome; Poelstra et al. 2014, Lawson and Petren 2017). In addition, the resulting assembly must be physically mapped on the different chromosomes of the species using linkage maps so that samples can then be 're-sequenced' within the regions of interest (Lawson and Petren 2017). This type of approach results in a 'genomic landscape' where we can explore how genetic differentiation is distributed among species, populations or phenotypes (Poelstra et al 2014). Regions with high levels of differentiation are usually characterized by low levels of nucleotide diversity, biased allelic frequencies, high levels of linkage disequilibrium, and/or high levels of divergence in homologous polymorphisms between locally adapted groups (Campbell et al. 2018). When these conditions are met, it is possible not only to determine that these different groups are subject to divergent selection, but also to locate the genetic basis of divergence within the genome (Lawson and Petren, 2017).

The aim of this study is to elucidate the patterns of genetic differentiation that underlie phenotypic differences between two populations of the Large Psammodromus (*Psammodromus algirus*) separated by an altitudinal gradient. Although these populations show very little genetic differentiation (Verdú-Ricoy et al. 2010, Díaz et al. 2017), they differ in a wide variety of adaptive phenotypic traits, including several life-history characteristics (Iraeta et al. 2006, 2010, 2011, 2013, Capítulo 1: Llanos-Garrido

et al. 2017; Table 1).

In addition, previous studies have shown through reciprocal transplant and common garden experiments that these adaptive differences are not supported solely by environmental effects (Iraeta et al. 2006, 2013). Therefore, there must be a genetic basis to determine such phenotypic differences between these apparently undifferentiated populations. To define the degree of genetic differentiation between both populations, we used a genomic scan based on 73,291 SNPs that allowed us to analyze the genetic structure and genetic distance between them. In addition, we used a Bayesian method of outlier detection, i.e. detection of SNPs with between-population genetic distances greater than expected given the degree of overall background genomic differentiation. By doing this, we tried to identify polymorphisms that could provide the genetic basis for divergent selection undepinning observed adaptive differences (Bonhomme et al 2010). These approach has the same theoretical basis as the ones based on large assemblies explained above. However, the analytic power in detecting peaks of genetic differentiation is relatively lower, because the physical position of each variant is unknown and it is not possible to find islands of selection with respect to adjacent positions. Instead, overall genomic differentiation is assumed to represent the background (or basal) level of genetic differentiation, with which we can compare SNPs with a greater degree of divergence and whose allelic frequencies deviate from what could be expected under neutrality (Lewontin and Krakauer 1975). The reason why comparing differentiation peaks with respect to adjacent regions facilitates this task is that the degree of differentiation is heterogeneously distributed throughout the genome (Campagna et al 2015). Therefore, detecting a divergence peak in a specially conserved region is not possible for the approaches that use delocalized genetic variation; such degree of divergence may be well below the overall genomic level of background

Table 1. Phenotypic differences found between the populations of El Pardo and Navacerrada in previous studies.

Phenotypic trait	**Reference**
Shorter incubation time in the mountain population	Iraeta et al. 2006
Largest youth hatching in the plains population	Iraeta et al. 2006
Faster growth in lowland juveniles (reciprocal transplant experiment)	Iraeta et al. 2006
Faster growth in lowland juveniles controlling food abundance (*common garden* experiment)	Iraeta et al. 2013
More plastic activity levels in response to food availability in the plain population	Iraeta et al. 2008
Larger clutches of smaller eggs in the plain population	Iraeta et al. 2008 Iraeta et al. 2013
Longer flight distance for pregnant females in the mountain population	Iraeta et al. 2010
Relatively longer hind limbs in the plain population	Iraeta et al. 2011
More relatively larger femoral pores in males in the plain population	Iraeta et al. 2011
Greater extent of sexual coloration of the head in males in the plain population	Iraeta et al. 2011 Llanos-Garrido et al. 2017
Larger adult females in the mountain population	Iraeta et al. 2013
Counteragradient variation in body size: the genotypes that presumably control the key adaptations of the plain population (eggs and large post-natal fast growing juveniles) occur in a poorly productive environment in which juveniles grow more slowly and reach a smaller adult size.	Iraeta et al. 2006 Iraeta et al. 2013
Increased saturation of the sexual coloration of the head in males in the mountain population	Llanos-Garrido et al. 2017
Males respond to the activation of the immune system by reducing the extent of sexual coloration of the head in the plain population and reducing the saturation of sexual coloration of the head in the mountain population.	Llanos-Garrido et al. 2017
High rates of infestation by tick nymphs in the mountain population but not in the plain population (without ticks)	Llanos-Garrido et al. 2017

differentiation, but it may actually represent a clear-cut peak of adaptive divergence within its own genomic location (Lawson and Petren, 2017). This scenario is especially common in coding regions, where the genetic basis of phenotypic diversity is located. As a consquenece, approaches based on delocalized SNPs do not usually allow to identify the genetic basis of a phenotypic divergence, but they rather describe the general patterns of genetic differentiation that lie behind the process of local adaptation (for an example, see Tigano et al. 2017).

Material and methods

Study system

Psammodromus algirus (Linneus, 1758) is the most abundant lacertid lizard in Mediterranean scrubland and forest habitats of the Iberian Peninsula (excluding the northern Eurosiberian belt). It is medium sized species (snout-vent length: 60-90 mm; body mass: 6-16 g) in which variation in body size of adult individuals is an important indicator of fitness, as it depends directly on growth rate and age, and determines survivorship and reproductive success (Díaz et al. 2005, Martín and Forsman 1997, Iraeta et al. 2013). *P. algirus* males present a reddish nuptial coloration during the months of highest activity (April-June) that can range from small sublabial marks in subdominant males to occupy practically the entire head in the largest males. This visual signal is complemented by the secretions of femoral pores that are used for chemical communication. The number of femoral pores is higher in males than in females, and their activity increases as the breeding season advances in response to increased blood androgen levels (Chiu and Maderson 1975, Cole 1966, Diaz et al. 1994).

The populations in this study are separated by an altitudinal gradient of 600-700

m, and they differ in mean annual temperature, precipitation, habitat structure, and productivity. Previous studies on these populations revealed very little genetic differentiation based on mitochondrial DNA data (Verdú-Ricoy et al. 2010), despite the fact that they differ in a wide variety of adaptive phenotypic traits (Table 1).

The montane population is found at Navacerrada (Cerro de la Golondrina, Sierra de Guadarrama: 40°44′N, 4°00′W; 1300 m a.s.l.), and its habitat is composed by oak forests (*Quercus pirenaica*), scrub patches dominated by *Cistus laurifolius* and, to a lesser extent, rock outcrops and grasslands. Average annual temperature is 6.2 ºC and average precipitation is 1,170 mm. The lowland population is found at El Pardo (Madrid: 40°31′N, 03°47′W; 650 m a.s.l.), and it is a perennial forest of holm oaks (*Quercus ilex*) with shrubs of *Cistus ladanifer*. It is separated from the montane population by a linear distance of 32 km, and it has an average annual temperature of 12.5 ºC and an average annual rainfall of 438 mm. *P. algirus* is the most abundant lacertid lizard species at both sites, but it reaches higher densities at the montane location (Díaz 1997).

Field sampling and phenotypic analysis

The field sampling consisted in the capture by hand or with a loop of 40 individuals equally distributed between sampling locations and sexes (10 males and 10 females from each population). The captures were made between April and May 2015 and the number of ticks carried by each individual was counted in the field. The lizards were moved to the laboratory, where all phenotypic traits (snout-vent length, head width and length, hind leg length, number of femoral pores, and body mass) were measured. In addition, we photographed each lizard's head in ventral position (i.e. throat and infralabial scutes), to obtain data about sexual coloration. For this last purpose, we processed the photographs with Adobe Photoshop CS6 as explained in Chapter 1

(Llanos-Garrido et al. 2017). We standardized the area of analysis using the 'magnetic loop' tool and we used the 'magic wand' tool to measure the red coloured surface (30% tolerance) after having selected a random red colored point. Then, we used the 'similar' option of the magic wand tool (with the same 30% tolerance), to select all pixels with a similar coloration. The colored surface was measured as the percentage of colored pixels in the analyzed area. To calculate red saturation, we used the proportion of the red channel within the RGB channel, that is: R/(R+G+B), where R, G and B are the red, green and blue channels of the graphics card, respectively. All these measures were taken blindly with respect to the population of origin. After all this work was done, the lizards were released at their site of capture.

All phenotypic analyses were performed with Statistica software (Statsoft). In order to analyze between-population differences, we used two-way AN(C)OVAs with sex and population as the categorical predictors and the different morphological measures as dependent variables; when it was necessary to control for the effects of the body size, snout-vent length was used as a covariate. In order to evaluate the relationship between tick load and different predictors (sex, size, or number of femoral pores), the analysis was restricted to the Navacerrada population (no ticks were found at El Pardo, see below). Structural body size of Navacerrada males was estimated by the scores of individual lizards on the first axis of a Principal Component Analysis (PCA) (with snout-vent length, head length, head width, mean leg length, and body mass) that retained 84.2% of the total variance in the data set.

DNA extraction and sequencing

We obtained tissue samples by clipping 2 cm of the tail tip of lizards, which were afterwards released at their site of capture. We purified DNA for library preparation using the Speedtools Tissue DNA Extraction kit (Biotools). We used the restriction

enzyme Pst1 for GBS library preparation. Sequencing was done in an Illumina HiSeq2500 sequencer. To recover SNPs we used the pipeline UNEAK, implemented in TASSEL v.3.0 (Bradbury et al. 2007), which is specifically designed for samples with no reference genome. We aligned sequence tags to each other to form 'networks' of tags, where each node is a single tag sequence, and each edge represents a single base pair difference between two tags. We pruned the networks to remove putative sequencing errors (low frequency alleles) using the error rate threshold parameter. We also discarded loci with minor allele frequencies < 0.01 or that could be successfully sequenced in less than 10% of individuals. The resulting dataset had 73,291 biallelic SNPs (Single Nucleotide Polymorphism), a site depth of 6.60 ± 6.75 and a site missingness of 0.42 ± 0.31.

Characterization of neutral variation and SNPs under selection

Prior to performing clustering analyses, we used the software PLINK v.1.9. (Purcell et al. 2007) to prune the SNP database for linkage disequilibrium (LD), according to observed sample correlation coefficients. This was necessary because our clustering model (described below) did not take into account LD, and therefore linked SNPs could bias the grouping. We performed clustering analyses to find genetic structure among populations and lineages (eastern or western), and to obtain the assignment probabilities of each individual to each one of the resulting clusters. We used the program ADMIXTURE v.1.3 (Alexander et al. 2009) for maximum likelihood estimation of individual ancestries. We used cross-validation errors to identify the number of clusters (K-value) for which the model had highest predictive accuracy. However, we also explored higher K values to detect patterns of genetic structure among the five populations.

To minimize false positives in outlier analyses, we discarded loci that could not

be successfully sequenced from at least 75% of individuals in each population, and loci with minor allele frequencies < 0.05 in each population, thus excluding all private alleles from the dataset. The resulting SNP dataset included 6,421 loci. We used a Bayesian approach to perform an outlier analysis as implemented in Bayescan v.2.1 (Foll and Gaggiotti 2008). Bayescan uses a logistic regression model to partition F_{ST} coefficients into a population-specific term (β) and a locus-specific term (α). We selected loci with $\alpha > 0$ as suggesting positive selection, and a false discovery rate (corrected for multiple testing) $q < 0.05$. To obtain these parameters, we ran the MCMC algorithm implemented in the program with a prior odd value of 10, and using 20 pilot runs of 5,000 iterations each, followed by 100,000 iterations with a burn-in of 50,000 interactions.

We subsequently tried to annotate the SNPs with highest values of α by running a BLAST analysis against all records in the NCBI nucleotide database. In addition, we used χ^2-tests to look for significant deviations from Hardy Weinberg conditions, and we applied a sequential Bonferroni correction to correct for multiple testing. This was done to detect significant departures from Hardy-Weinberg in only one of the two populations considered, because such departures might be indicative of the action of natural selection in that population but not in the other one.

Results

Phenotypic variability

We found significant differences between both populations that basically confirmed the results of previous studies. Thus, lizards from the lowland population (El Pardo) were smaller, had longer legs, and tended to have more femoral pores, than those from the montane population (Navacerrada). In addition, no lizard from El Pardo was parasitized

by ticks, whereas levels of infestation were high in Navacerrada (Table 2). Males' sexual coloration data were especially scarce in our sample (very few colored males with weak sexual colorations), so that differences between populations could not be analyzed.

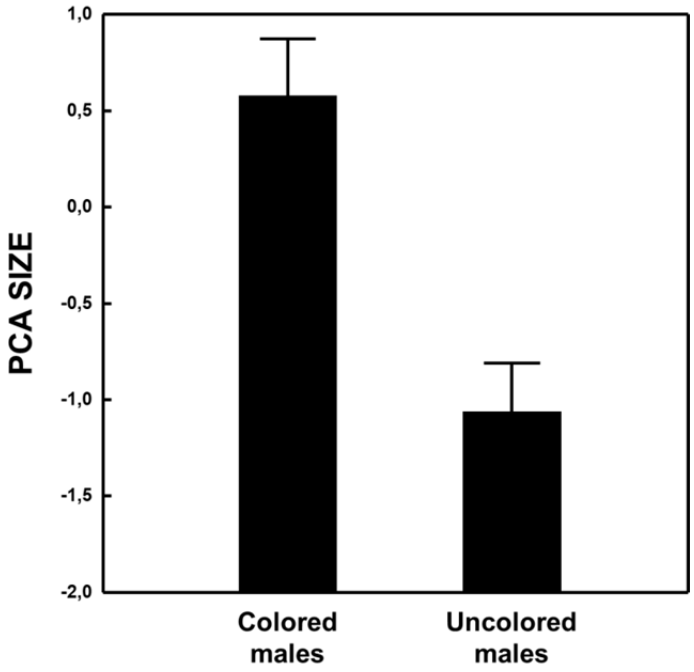


Figure 1. Size differences between males with sexual ornamentation and those without it in Navacerrada (mean ± IC95).

When we only considered lizards from Navacerrada, we found that males had more ticks than females ($F_{1,18} = 10.85$, $p = 0.004$), and that males that succeeded to develop a sexual coloration were larger than those that did not suceed ($F_{1,8} = 42.17$, $p < 0.001$; Fig. 1). In addition, we detected that after controlling for the effects of the number of femoral pores (an indicator of investment in sexual signalling, in the absence of a well developed sexual coloration), uncolored males tended to have more ticks (ANCOVA; number of pores: $F_{1,7} = 7.54$, $p = 0.029$; colour [categorical]: $F_{1,7} = 3.93$, $p = 0.088$).

Table 2. Phenotypic differences between the populations studied in this chapter: population means, with results of two-way ANOVAs with sex and population as factors. The last two variables (marked with an asterisk) show adjusted means and results of two-way ANCOVAs after controlling for the effects of body size (SVL).

Feature	El Pardo (mean ± sd)	Navacerrada (mean ± sd)	Effect of population
Body length	68.5 ± 2.5	71.9 ± 5.1	$F_{1,38} = 6.73$; $p = 0.01$
Head length	10.4 ± 1.2	11.2 ± 1.5	$F_{1,38} = 3.50$; $p = 0.07$
Ticks	0.0 ± 0.0	4.4 ± 4.1	$F_{1,38} = 66.77$; $p < 0.001$
Long. legs*	22.2 ± 0.4	20.6 ± 0.4	$F_{1,35} = 19.54$; $p < 0.001$
No. of pores*	17.8 ± 0.4	17.1 ± 0.2	$F_{1,35} = 2.56$; $p = 0.12$

Neutral genetic variability and SNPs under selection

The analysis of genetic structure with Admixture yielded a single more plausible model that grouped all samples into a single cluster (K = 1, cv-error = 0.382). Thus, no genetic structure is apparent between both populations. Forcing the model towards higher values of K (with lower levels of likelihood) does not help to find genetic structure between populations (K = 2, cv-error = 0.405; Figure 2).

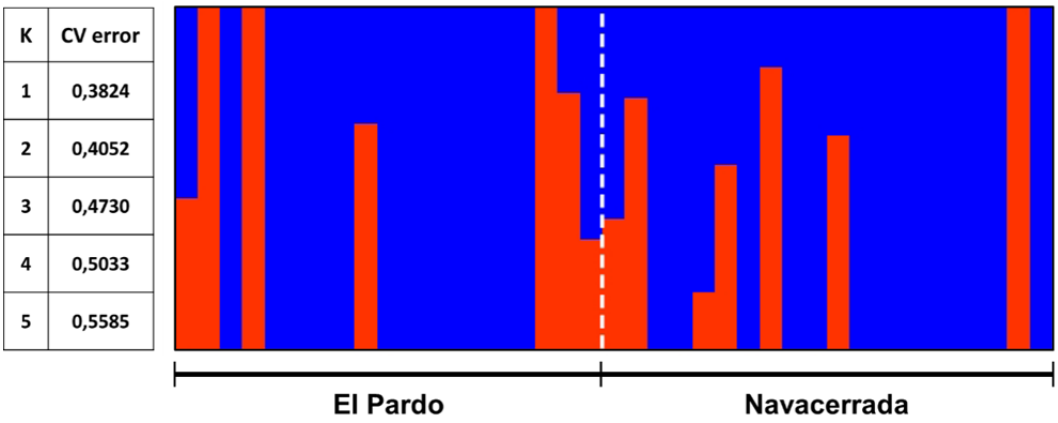


Figure 2. Probability of assignment to groups generated by Admixture for the model with two groups (K=2). On the left, CV-errors for each of the models analyzed; the most credible model is the one that includes only one group (K=1).

Regarding our outlier analysis, Bayescan did not detect any loci subject to selection (range of q-values: 0.883 - 0.910). The largest increase in α (which is the magnitude of the locus-specific effect of selection) concentrated on four SNPs (Table 3), whose α values were still very low. Of these SNPs, only two could be annotated. The first one had a marginally non-significant E-value of 0.018 (conventional significance value = 0.01), and it corresponded to the coding region of a repeated transmembrane ankirin involved in the cellular response to salicylic acid (inflammatory response) and the regulation of the immune system. The second sequence was annotated with a much higher E-value (0.6) as the coding region of a nuclear photoreceptor protein that regulates gene expression based on the reception of steroids. None of these SNPs had allelic frequencies that deviated significantly from Hardy-Weinberg after applying the sequential Bonferroni correction (Table 3).

Table 3. SNPs with greater divergence between populations detected by Bayescan. The Bayescan result (α and q) and the Chi-square analysis are presented to verify that the allelic frequencies of these SNPs do not differ significantly from what was expected under Hardy Weinberg's equilibrium.

SNP_ID	α (q)	H-W (El Pardo)	H-W (Navacerrada)
14537818	0.067 (0.883)	$\chi^2 = 0.436$; p = 0.804	$\chi^2 = 0.900$; p = 0.638
17216934	0.059 (0.887)	$\chi^2 = 5{,}445$; p = 0,066	$\chi^2 = 0.183$: p = 0.913
17911336	0.050 (0.891)	$\chi^2 = 7{,}871$; p = 0,020	$\chi^2 = 1.389$; p = 0.499
18623045	0.042 (0.885)	$\chi^2 = 0{,}720$; p = 0,698	$\chi^2 = 0.106$; p = 0.949

Discussion

Our results point to a lack of agreement between the absence of genomic divergence and the consistent phenotypic differentiation in traits whose adaptive value seems obvious that has been found in different previous studies (Iraeta et al. 2006, 2010, 2011, 2013). The fact that at Navacerrada the load of ectoparasites in males was again related to secondary sexual characteristics (see chapter 1), and that once again there were no

infestations at El Pardo, shows that the presence of ticks implies a different selective pressure in both populations (Chapter 1, Llanos-Garrido et al. 2017). Apart from their negative effect on the survival of individuals (Salvador et al. 1996), ticks appear to have other significant impacts on the fitness of individuals, influencing the expression of key sexual ornaments for reproduction (Chapter 1, Llanos-Garrido et al. 2017). In addition, the altitudinal gradient must involve many other selective pressures that, although not as evident as the one imposed by ticks, should also cause an undeniable impact on fitness components of as significant as growth rate or female clutch size (Iraeta et al. 2006, 2010, 2011, 2013). Moreover, the adaptive phenotypic differences found in this and earlier studies span throughout the entire life histories of lizards (affecting hatchling size, growth rate, sexual ornamentation, and the offspring size vs. offspring number trade-off). This phenotypic differentiation should involve a large number of divergent selective pressures underlying the observed differences (Roff 2007). Also, the fact that many of such differences persist in common garden or reciprocal transplant experiments, strongly suggests a genetic (or at least inheritable) basis for their divergence (De Kort et al. 2014).

However, the background genetic differentiation between these two populations is very low (Díaz et al. 2017, Verdú-Ricoy et al. 2010), and it remained undetectable when analytical power was greatly increased using genome-wide scans (results of this study). This should not hinder the genetic differentiation of the specific loci on which adaptive divergent traits are based. There are other systems in which the background genetic differentiation between phenotypically well-differentiated groups is low, and yet islands of selection with outstanding levels of local divergence can be found within relevant fractions of the genome (Aguillon et al. 2018). Such islands of divergence may remain differentiated even in scenarios of continuous gene flow (Moody et al. 2015,

Poelstra et al. 2014, Shaner et al. 2015). However, in our study system we failed to evaluate the amount of migration between putative genetic groups (populations), because we were not able to identify the genetic structure that makes such estimation meaningful (Phillips et al. 2008)

It seems that in our case the divergence between populations is so recent that it has not translated into a genomic footprint to a scale large enough to be detected with the techniques we have used (Aguillon et al al. 2018). However, there is compelling evidence that at least a small fraction of the genome has truly diverged, giving rise to the phenotypic differences observed in our system. One possible explanation is that the only divergent regions are those shaped by natural selection and expressed as divergent adaptive phenotypes (Safran et al. 2016). These regions would be located in very specific areas of the genome and they could be really scarce (Campagna et al. 2015). In this case, the main problem to detect these regions with our methods could well be the genotyping power of GBS; if so, the limiting factor would be the amount of genetic variability uncovered by our sequencing approach that contains enough meaningful information to reveal the effects of natural selection.

Another possibility is that the SNPs that give raise to the observed adaptive differences could be located in genomic regions with low genetic diversity. Such regions should include a peak of divergence at a local genomic scale, but they would still escape the *outlier* detection methods used in this study (Campbell et al 2018). This might occur because the degree of genetic variability is distributed heterogeneously throughout the genome (Feulner et al 2015, Nosil et al. 2009), but our outlier detection approach only captures peaks of divergence among SNPs that stand out from background (i.e. average) genomic divergence. As a consequence, longer assemblies would be needed to detect genomic divergence at a local genomic scale, by comparing

regions putatively under selection with contiguous sequences (Poelstra et al. 2014). However, heterogeneity in genome-wide levels of genetic divergence is unlikely to be confounding our analysis, because the degree of background genomic divergence between both populations is actually very low.

A final possible explanation is that the 'adaptive' phenotypic differences reported in this and previous studies lack a genetic basis and, as a consequence, have no correlates on genomic scans. It is true that some of those studies point to a genetic basis, especially those that report reciprocal transplant (Iraeta et al. 2006) or *common garden* experiments (Iraeta et al. 2013). However, it could also be the case that the differences observed were the result of maternal effects or inheritable epigenetic responses to the environment (Groot et al. 2018, Trerotola et al. 2015). In those experiments, the juveniles being compared came from gravid females that laid their clutches in the lab but whose eggs were formed under the environmental conditions of their respective habitats (Iraeta et al. 2006, 2013). Epigenetic marks can be inherited, but their duration is restricted to a low number of generations; afterwards, the signal is blurred (Trerotola et al. 2015). Thus, the demonstration that phenotypic differences have a genetic basis would require to prolong the experiment a few more generations, so that we could check whether differences have been blurred or not (Groot et al 2018). In addition, since adaptive epigenetic marks act as modifiers of gene expression (i.e. they control what genes are transcribed and which are not), a study of transcriptomic differentiation could be very useful, eventually guiding the search for variants subject to selection in differentiated regions of the genome (McGirr and Martin 2018).

In summary, our analysis of GBS data has not allowed us to detect a genetic basis on which the phenotypic differences observed in this and other studies can be based. If such a genetic basis exists, our results suggest a scenario in which a small

number of changes have a large effect on the phenotype. In these circumstances, even a large-scale genomic scan like the one carried out in this study would have enormous difficulties to identify the changes responsible for phenotypic divergence, because in such undifferentiated populations the divergence peaks would be located in very specific regions of the genome, whose inclusion and subsequent detection in sampled sequences would be extremely unlikely (Joo et al. 2016). Another possibility is that the discordance between genetic and phenotypic differentiation reflects the presence of inheritable epigenetic marks. In short, this study highlights the need to work with larger assemblies, and even to perform transcriptomic analyses, in order to provide a reliable picture of the genetic basis of adaptive phenotypic differences (Harris et al. 2018).

CAPÍTULO 3. El uso combinado de dos métodos de detección de *outliers* revela las dinámicas de adaptación local a escala genómica.[3]

Resumen

La adaptación local es un proceso dinámico mediante el cual diferentes combinaciones alélicas son seleccionadas en diferentes poblaciones en distintos momentos de su historia evolutiva, y cuya señal genética puede ser inferida mediante análisis de *outliers* a escala genómica. En este capítulo se combinan estimas de flujo genético con dos métodos de detección de *outliers*, uno de ellos independiende del patrón de coancestría entre poblaciones (CIOA) y otro que no discrimina los efectos de la coancestría (ROA), para identificar las variantes genéticas que se ven favorecidas cuando la ecología promueve convergencia fenotípica. Analizamos cinco poblaciones de *P. algirus* con un área de distribución amplia y heterogénea que además ha ido cambiando desde la separación de los dos linajes (oriental y occidental) que componen la especie desde hace aproximadamente 3 ma. En general, las lagartijas del oeste habitan ambientes forestales y carecen de patrón dorsal rayado, mientras que las del este habitan ambientes más abiertos y son rayadas. Sin embargo, una población (Lerma) presenta un fenotipo sin raya dorsal a pesar de pertenecer al linaje oriental. El análisis de 73,291 SNPs confirmó la división este-oeste e identificó 12 (ROA) y 9 (CIOA) *outliers* no solapantes. ROA reveló divergencia adaptativa para las poblaciones del este y del oeste localizadas en los extremos del gradiente ambiental, y reveló variación genética ancestral como fuente de alelos localmente seleccionados en Lerma; CIOA detectó variantes innovadoras en Lerma. En general, la separación adaptativa de los linajes parece haber sido compensada por selección equilibradora y patrones asimétricos de flujo genético, lo que podría estar

[3] Basado en el siguiente artículo: Llanos-Garrido, A., Pérez-Tris, J. y Díaz, J. A. The combined use of raw and phylogenetically independent methods of outlier detection uncovers genome-wide dynamics of local adaptation in a lizard. Enviado a Evolution (en primera revisión).

dificultando la especiación a pesar de la fuerte estructura genética este-oeste y de la baja capacidad de dispersión de las lagartijas.

Abstract

Local adaptation is a dynamic process by which different allele combinations are selected in different populations at different times, and whose genetic signature can be inferred by genome-wide outlier analyses. We combined gene flow estimates with two methods of outlier detection, one of them independent of population coancestry (CIOA) and the other one not (ROA), to identify genetic variants favored when ecology promotes phenotypic convergence. We analysed five populations of a lizard distributed over a wide, heterogeneous range that has been changing since the split of eastern and western lineages ca. 3 mya. Overall, western lizards inhabit forest habitat and are unstriped, whereas eastern ones inhabit shrublands and are striped. However, one population (Lerma) has unstriped phenotype despite eastern ancestry. The analysis of 73,291 SNPs confirmed the east-west division and identified 12 (ROA) plus 9 (CIOA) non-overlapping outliers. ROA revealed adaptive divergence for eastern and western populations at the extremes of the environmental gradient, and uncovered standing variation as a source of alleles locally selected in Lerma; CIOA uncovered innovative variants selected in Lerma. Overall, adaptive lineage splitting was seemingly counterbalanced by balancing selection and asymmetrical patterns of gene flow, which could hamper speciation despite strong genetic structure.

Introduction

Local adaptation is at the core of the origin and maintenance of genetic variation and its geographical patterning across species' distribution ranges (Endler 1977, Schluter 2001,

Hereford 2009, Savolainen et al. 2013). Local adaptive challenges contribute to transform adaptive landscapes during the evolution of species' ranges, a process which leads to different gene combinations being selected in different populations at different times during their evolutionary history (Wright 1932, Kokko et al. 2017, Svensson and Clasbeek 2017). It is thus relevant to investigate the genetic footprint of past and present contingencies that may shape adaptation as an incomplete and dynamical process, not only by studying how local adaptation influences current genetic dynamics, but also by inferring how it has been favored, constrained and/or blurred by the evolutionary history of a system (Rundle and Nosil 2005, Rosenblum et al. 2007, Laurent et al. 2016, McEntee et al. 2018). Genomic approaches to infer local adaptation have become very popular, because they offer a temporal perspective that other approaches (such as common garden or transplant experiments) cannot provide (Tiffin and Ross-Ibarra 2014).

It is possible to screen genome-wide patterns of DNA polymorphism to detect locus-specific signals of divergent selection (Luikart et al. 2003, Schlötterer 2003). This method uses theoretical predictions about the effects of positive natural selection on allele frequencies to identify regions in the genome that harbor adaptive mutations (Kim and Stephan 2002, Przeworski 2002). By screening genome-wide patterns of polymorphism, it is possible to identify 'islands' of locally elevated genetic differentiation between groups of individuals putatively subject to divergent selection (Wu 2001, Han et al. 2017). However, the genomic landscape of divergence varies depending on the evolutionary histories of the lineages involved, making the detection of selection islands challenging in many occasions.

One evolutionary scenario that may obscure the contribution of adaptation to population genetic structuring is the existence of deep population differentiation

accumulated during long periods of lineage divergence. In these cases, adaptive mutations can be identified by screening genetic variants randomly distributed along the genome, and identifying those that follow patterns of allele frequencies deviated from the general pattern of divergence (outliers), and that are therefore compatible with adaptive divergence of the genomic regions that contain them (Rhode et al. 2017, Schield et al. 2017, Tigano et al. 2017).

The most immediate approach to outlier analysis consists of uncovering which loci present traces of selection, and whether this selection is divergent or not (Tiffin and Ross-Ibarra 2014). To this end, Bayesian approaches that use logistic regression to partition F_{ST} coefficients into a population-specific term and a locus-specific term (e.g. Foll and Gaggiotti 2008), have proved convenient. It is also possible to get traces of selection independent of the phylogenetic configuration, as revealed by outliers whose allele frequencies deviate from the expected under genetic drift alone, given the population tree (e.g. Bonhomme et al. 2010). For simplicity, we refer to these two approaches as raw outlier analysis (ROA) and coancestry-independent outlier analysis (CIOA), respectively. If the two sets of outliers detected by these alternative approaches do not overlap, ROA and CIOA, combined with isolation-migration analyses, can shed light on the relative roles of standing genetic variation, recent admixture, and genetic innovations, as the major drivers of local adaptation dynamics (Barrett and Schluter 2008).

We use genotyping by sequencing (GBS; Elshire et al. 2011) and combined ROA and CIOA to explore how standing genetic variation and genetic innovation underlie local adaptation in the most abundant and widespread lizard in the Mediterranean region of the Iberian Peninsula: the Large Psammodromus *Psammodromus algirus* (Díaz et al. 2017). This species shows striking genetic and phenotypic divergence associated with

environmental heterogeneity across its distribution range. The north-west part of its range is more humid and temperate than the southeast, and this is mirrored by broad changes in the vegetation patterns, with forests dominating in the west and more open spaces in the east (Díaz et al. 2017). The genetic diversity of *P. algirus* is structured in two mtDNA lineages, eastern and western, separated ca. 3-3.5 Ma (Carranza et al. 2006). Each mtDNA lineage typically shows a distinct phenotype for a heritable trait that could be adaptively linked to crypsis: lizards that inhabit the eastern, more open regions tend to display a striped pattern absent among western lizards, which inhabit more vegetated habitats (Díaz et al. 2017).

Interestingly, striped and unstriped phenotypes do not perfectly match the eastern-western lineage dichotomy. Moreover, the population in Lerma, a favorite study site of the authors in northern Spain (see Díaz et al. 2005, Santos et al. 2008, 2009, Telleria et al. 2011), inhabits a forest archipelago that resembles the typical habitat of western lizards. At first glance the phenotype of these lizards is unstriped, fitting the expected from local adaptation to the forests they live in. However, the analysis of their mtDNA places Lerma within the E2 clade of the eastern lineage, which is almost exclusively composed of striped lizards (Díaz et al. 2017). Therefore, the Lerma population provides an excellent opportunity to investigate the source of genetic variants that are locally selected in populations of the large Psammodromus.

One possibility is that genetic variation standing within lineages since they diverged allowed populations to keep locally adapted in the face of environmental heterogeneity. In Lerma, this source of standing genetic variation should be detected by ROA in the form of outliers with higher heterozygosity than the average for the genome. However, these outliers should not be detected by CIOA, as they would not deviate from the expected from the phylogenetic structure of populations. In addition, local adaptation of

populations faced with novel environments could have been attained through genetic innovation. In Lerma, genetic innovation might be provided by immigrants bringing genetic variants from populations already adapted to similar environment. If genetic innovation is the result of recent admixture, we also expect that ROA detects outliers with high heterozygosity, but not higher than the average heterozygosity of the entire genome. Finally, genetic innovation may also evolve from natural selection of locally arisen variants. In Lerma, such innovations can be uncovered by CIOA in the form of outliers with higher levels of homozygosity than the average for the genome (ROA would have more difficulty to detect such outliers if the phylogenetic inertia of adaptation is large). These scenarios are not mutually exclusive, and their contribution to lizards' local adaptation can be inferred with the outlier analyses described above. Our goal is not to identify the genetic basis of existing adaptations. Rather, we aim to unravel how evolution supplies the genetic variation that ecology demands for local adaptation in populations scattered over a heterogeneous species' range.

Materials and Methods

Field sampling

We sampled lizards in the focal population (Lerma) and in two pairs of populations that replicated either its eastern genetic lineage (Aranjuez and Brihuega of the E2 clade; Verdú-Ricoy et al. 2014; Díaz et al. 2017), or its forested habitat type within the distribution of the western lineage (El Pardo and Navacerrada of the W2 clade; Díaz et al. 2017). All these populations are located near the putative contact zone between the eastern and western lineages. Lerma (42.058 ºN, -3.611 ºE; 900 m asl) is a fragmented forest of evergreen and deciduous trees interspersed with grassland patches. Aranjuez

(40.016 ºN, -3.586 ºE; 594 m asl) is a hot, dry site with a high cover of herbs and shrubs, and no trees, where all lizards are striped. Brihuega (40.778 ºN, -2.911 ºE; 1,009 m asl) is a deciduous open forest with a mosaic of grassland and woodland patches, where 67% of lizards are striped. El Pardo (40.511 ºN, -3.755 ºE; 658 m asl) is a xeric, lowland evergreen forest where 20% of lizards are striped, whereas Navacerrada (40.726 ºN, -4.023 ºE; 1,230 m asl) is a more productive montane location covered by deciduous forest with a very similar species composition to that of Lerma, where 17% of lizards are striped. A quantitative assessment of the degree of environmental variation between our study sites is provided in Supplementary Appendix S1 (for further details about the sites and the genetic and phenotypic characterization of their populations except Lerma, see Díaz et al. 2017).

Phenotypic variation

During the breeding season of 2012, we took tissue samples from 18 adult lizards noosed or captured by hand in Lerma. During the breeding season of 2015, we captured and sampled with the same methods another 77 lizards (20, 18, 19, and 20 from Aranjuez, Brihuega, El Pardo, and Navacerrada, respectively). These 95 lizards were used for genomic analyses. In 2016, we captured 18 lizards in Lerma that were used for the analysis of phenotypes along with the 77 lizards captured in the other populations during 2015. Therefore, different individuals from Lerma were used for genetic and phenotypic analyses. In order to characterize lizards' phenotype, we brought them to the lab to record their snout-vent length (SVL) and body mass and to measure their dorsal coloration from images. For this later purpose, we held the individuals stretched under an anti-reflective glass and immobilized against a soft sponge to avoid damaging them. We took pictures of the lizards' back in a dark room, using a fix setup with two white light sources placed on either side of the lizard at 25 cm from the subject, with an angle

of 45º. The camera was set at 35 cm from the subject to obtain an overhead image, with automatic exposure compensation and shutter speed of 0.5 ms. We used Adobe Photoshop CS6 (Adobe Systems, 2002) for image processing (Capítulo 1: Llanos-Garrido et al. 2017): we standardized the analyzed area using the '*magnetic lasso'* tool to delimit a 5-cm long surface from the shoulder joint (which set the width of the analyzed surface) towards the posterior end of the animal. We measured the size of the striped surface with the '*magic wand'* tool (at 10% tolerance) after clicking at the middle of the stripe. We subsequently used the '*similar*' option of the '*magic wand'* tool (at 10% tolerance) to select pigmented scales, and used the percentage of colored pixels in the analyzed surface as a measure of the size of the stripe. We measured the darkness of the stripe using the inverse of brightness of the colored layer. Note that we did not use a color standard to reduce the error in the measurement of brightness, but the fact that all lizards were photographed under the same conditions of illumination and exposure made our images directly comparable.

DNA extraction, sequencing and variant calling

We obtained tissue samples by clipping 2 cm of the tail tip of lizards, which were released unharmed at their site of capture. We kept the samples in absolute ethanol at 4 ºC until DNA extraction. We purified DNA for library preparation using the Speedtools Tissue DNA Extraction kit (Biotools) with a cell lysis step of 24 hours and resuspension in DNase-free water at 60 ºC.

We used the restriction enzyme *Pst1* for GBS library preparation. Sequencing was done in an Illumina HiSeq2500 sequencer. The pipeline used for SNPs discovery was UNEAK, implemented in TASSEL v.3.0 (Bradbury et al. 2007) and specifically designed for samples with no reference genome. Sequence tags were aligned to each other to form 'networks' of tags, where each node is a single tag sequence, and each

edge represents a single base pair difference between two tags. The networks were pruned to remove putative sequencing errors (low frequency alleles) using the error rate threshold parameter. The resulting dataset had 83,648 SNPs with a site depth of 5.96 ± 6.56 (mean ± sd) and a site missingness of 0.49 ± 0.33. We discarded loci with minor allele frequencies < 0.01 or that could be successfully sequenced in less than 10% of individuals. The resulting dataset had 73,291 loci, with only one SNP per locus, and with a site depth of 6.60 ± 6.75 and a site missingness of 0.42 ± 0.31.

Admixture and gene flow

Prior to performing clustering analyses, we used the software PLINK v.1.9. (Purcell et al. 2007) to prune the SNP database for linkage disequilibrium (LD), according to observed sample correlation coefficients. This was necessary because our clustering model (described below) did not take into account LD, and therefore linked SNPs could bias the grouping. We performed clustering analyses to find genetic structure among populations and lineages (eastern or western), and to obtain the assignment probabilities of each individual to each one of the resulting clusters. We used the program ADMIXTURE v.1.3 (Alexander et al. 2009) for maximum likelihood estimation of individual ancestries. We used cross-validation errors to identify the number of clusters (K-value) for which the model had highest predictive accuracy. However, we also explored higher K values to detect patterns of genetic structure among the five populations. To complement such clustering analysis, we performed a Multi-Dimensional Scaling analysis (MDS) on TASSEL 5.2. (Bradbury et al. 2007).

In order to improve the interpretation of our outlier analyses, we estimated gene flow among populations to test if there is less opposition to immigration of genetically western lizards in Lerma than in other populations of the eastern lineage. To obtain estimates of gene flow among populations, we used a maximum likelihood method

based on coalescence implemented in MIGRATE 3.6 (Beerli and Felsenstein 2001). MIGRATE uses an equilibrium model that estimates migration rates averaged across the coalescent history. Including undifferentiated populations in the analysis of migration rates would hamper the estimation of migration rates among less connected populations (Pfenninger and Posada 2002). Therefore, we pooled individuals from populations that remained undifferentiated in an admixture analysis with K = 5 (equal to the number of sampled populations). Although grouping will make the estimates of local population size not interpretable for pooled populations, this strategy does provide robust estimates of gene flow (Beerli and Felsenstein 2001). We used MIGRATE to estimate the effective population size scaled by mutation rate $\Theta = N_e\mu$, together with the effective number of migrants N_em, where N_e is the effective population size, μ is the mutation rate per generation, and m is the migration rate per generation. Migration rates were estimated as $M = m/\mu$ and presented with 95% credibility intervals. Starting values for all parameter estimates were initially obtained using F_{ST} (Beerli and Felsenstein 1999); all other searching parameters were set to default values. We performed a Markov Chain Monte Carlo (MCMC) search running ten heated chains with 100,000 recorded genealogies each, and three cold chains with 50,000 recorded genealogies, sampling each chain every 100 steps with a burn-in period of 10,000 steps.

Outlier analyses

After describing overall genetic structure and migration, we applied another filtering step to our SNP dataset to minimize false positives in the subsequent outlier analyses. To this end, we discarded loci with minor allele frequencies < 0.05 in each population (thereby excluding all private alleles from the dataset), or loci that could not be successfully sequenced from at least 75% of individuals in each population. The resulting SNP dataset included 6,421 loci.

We used a Bayesian approach to perform a ROA as implemented in Bayescan v.2.1, a conservative approach which is hardly prone to false positives and very useful when the number of populations is low (Foll and Gaggiotti 2008). Bayescan uses a logistic regression model to partition F_{ST} coefficients into a population-specific term (β) and a locus-specific term (α). We selected loci with $\alpha > 0$ suggesting positive selection, and false discovery rate (corrected for multiple testing) $q \leq 0.05$. To obtain these parameters, we ran the reversible-jump MCMC algorithm implemented in the program with a prior odd value of 10, and using 20 pilot runs of 5,000 iterations each, followed by 100,000 iterations with a burn-in of 50,000. In order to search for outliers taking into account the variation in the degree of relatedness among populations, we performed a CIOA using the Bonhomme et al. (2010) extension of the Lewontin-Krakauer test, using R code following Bonhomme et al. (2010). This approach infers the historical branching of populations from the matrix of pairwise distances among populations, and identifies outliers based on F_{ST} deviations from the expected if fixation rates were proportional to the branch lengths of the population tree. We selected loci based on the statistical significance of the FLK statistic, with a restrictive significance threshold of $p \leq 0.001$ to account for multiple testing. Finally, we used outliers inferred from ROA and CIOA to run new admixture analyses and compare them with the overall patterns of genetic admixture and gene flow.

We analyzed variation among populations in the heterozygosity of outlier SNPs detected by ROA and CIOA, using ANCOVA with population as factor, heterozygosity of the relevant set of outliers as the dependent variable, and overall heterozygosity as the covariate.

Results

Phenotypic variation

Phenotypic differences among populations were significant for all variables (ANOVA's for SVL, body mass, surface of the mid-dorsal stripe, and darkness of the mid-dorsal stripe: all $p < 10^{-6}$; Fig. 1a-c). Post-hoc tests (Tukey HSD for unequal N) revealed similar patterns for all phenotypes, with Aranjuez at one end of the phenotypic gradient and Lerma grouping with the western populations at the opposite end. These patterns were roughly consistent with habitat differences among populations (Table S1 and Fig. S1).

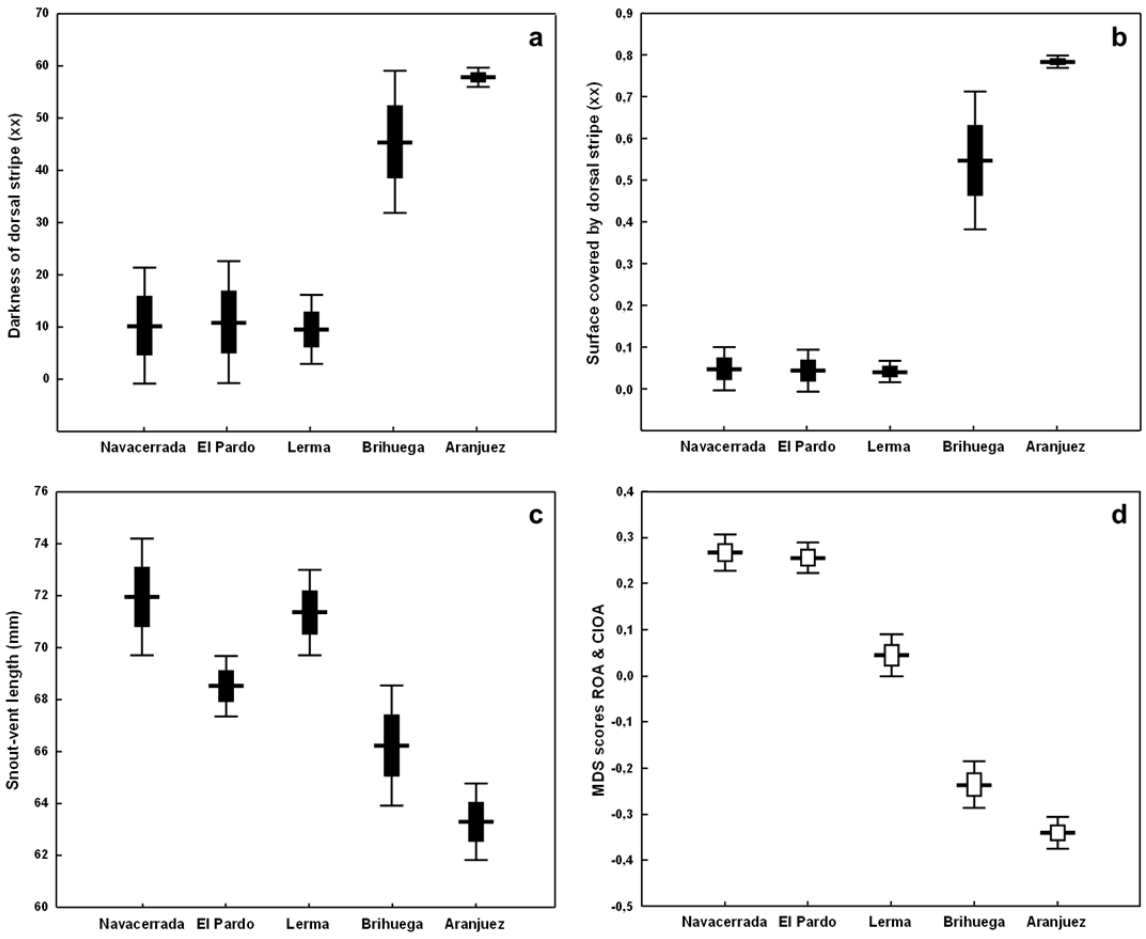


Figure 1. Phenotypic measures among populations (a-c) and MDS scores for outlier SNPs (d) (mean ± sd and 95% confidence interval).

Admixture and gene flow

Using 73,291 SNPs, the best model produced by ADMIXTURE had K = 2, recovering two clusters corresponding to the two mtDNA lineages (cv-error = 0.586; Fig. 2a). When we forced the model towards higher K values (with higher values of cv-error), Aranjuez was the first population to separate from the rest of its lineage (K = 3, cv-error = 0.665), followed by Brihuega and Lerma (K = 4, cv-error = 0.744). However, when

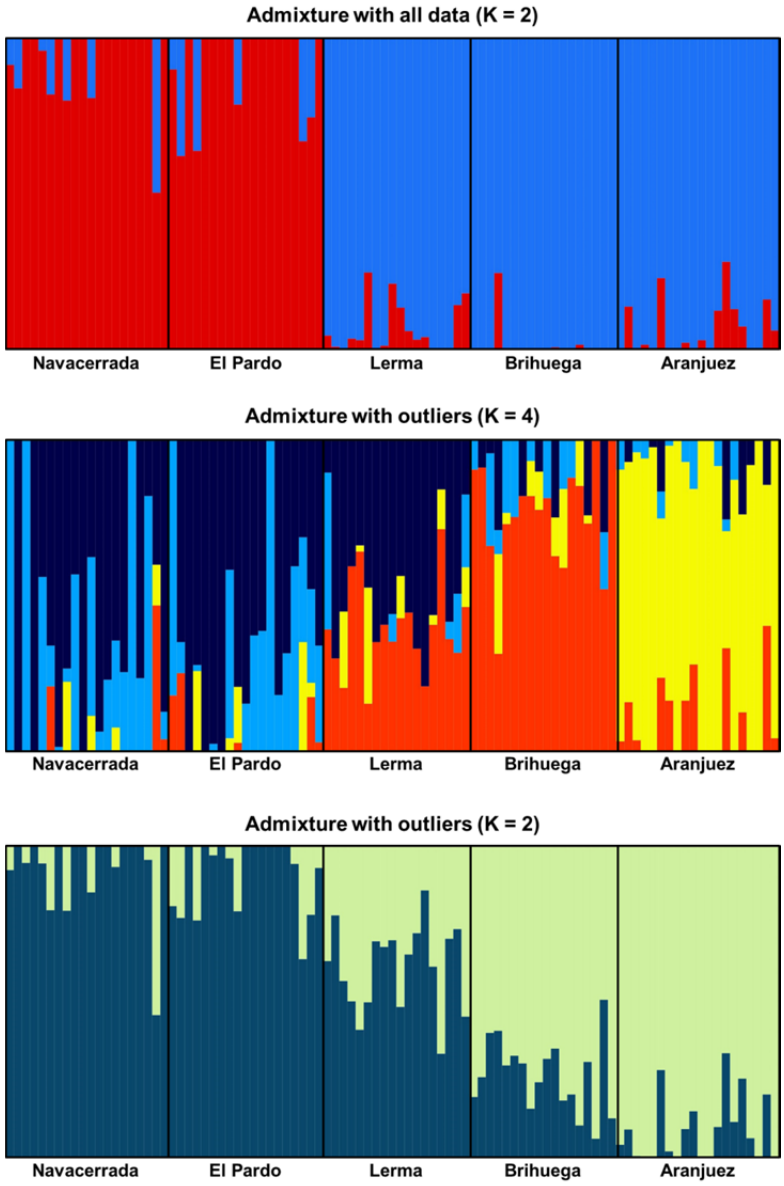


Figure 2. Genetic structure according to 73,291 SNPs (a) or 21 outlier SNPs (b, K=4; c, K=2). Bar colors represent posterior probabilities of cluster membership.

we set K equal to the number of populations in our sample, we did not detect any genetic structure between the two western populations (K = 5, cv-error = 0.810). The MDS using the same dataset produced an axis with eigenvalue = 0.463, which defined an east-west gradient that discriminated between both mitochondrial lineages with p = 1.0 ($F_{1,93}$ = 1369.19, R^2 = 0.936, $p < 10^{-6}$; mean ± sd scores of eastern and western lizards = -0.056 ± 0.017 and 0.081 ± 0.019, respectively). Our maximum likelihood estimates of gene flow revealed highly asymmetrical patterns of migration among lizard populations (ANOVA for among-group variation in the estimates of immigration: $F_{3,8}$ = 105.10, P < 0.001; Fig. 3). Among western lizards, MIGRATE recovered high levels of immigration from populations of the eastern lineage. Aranjuez was by far the most isolated population, while the other two eastern populations (Brihuega and Lerma) received intermediate levels of gene flow. Gene flow from western populations was higher in Lerma (Fig. 3: M = 65.6 ± 0.7) than in Brihuega (49.1 ± 0.6) or Aranjuez (6.25 ± 0.07).

Outlier analyses

Our ROA with Bayescan detected 12 outlier loci with $\alpha > 0$ ($0.97 < \alpha < 1.35$) and $q \leq 0.05$. The CIOA identified nine additional outlier loci with $p < 0.001$, none of which was detected by Bayescan. The MDS analysis performed with all 21 outliers yielded a single axis (eigenvalue = 6.78) representing a divergence gradient with western populations on one end and Aranjuez on the opposite end. This analysis placed Brihuega near Aranjuez, and Lerma close to the western populations, thus recovering the same pattern observed for lizard phenotypes (ANOVA with MDS scores: $F_{4,90}$ = 180.98, $p < 0.001$; Fig. 1d).

The best clustering model inferred by ADMIXURE using all the SNPs putatively

under selection had K = 4 (cv-error = 0.641), placing all individuals from Aranjuez and Brihuega into two clusters, and dividing individuals of Navacerrada and El Pardo into two

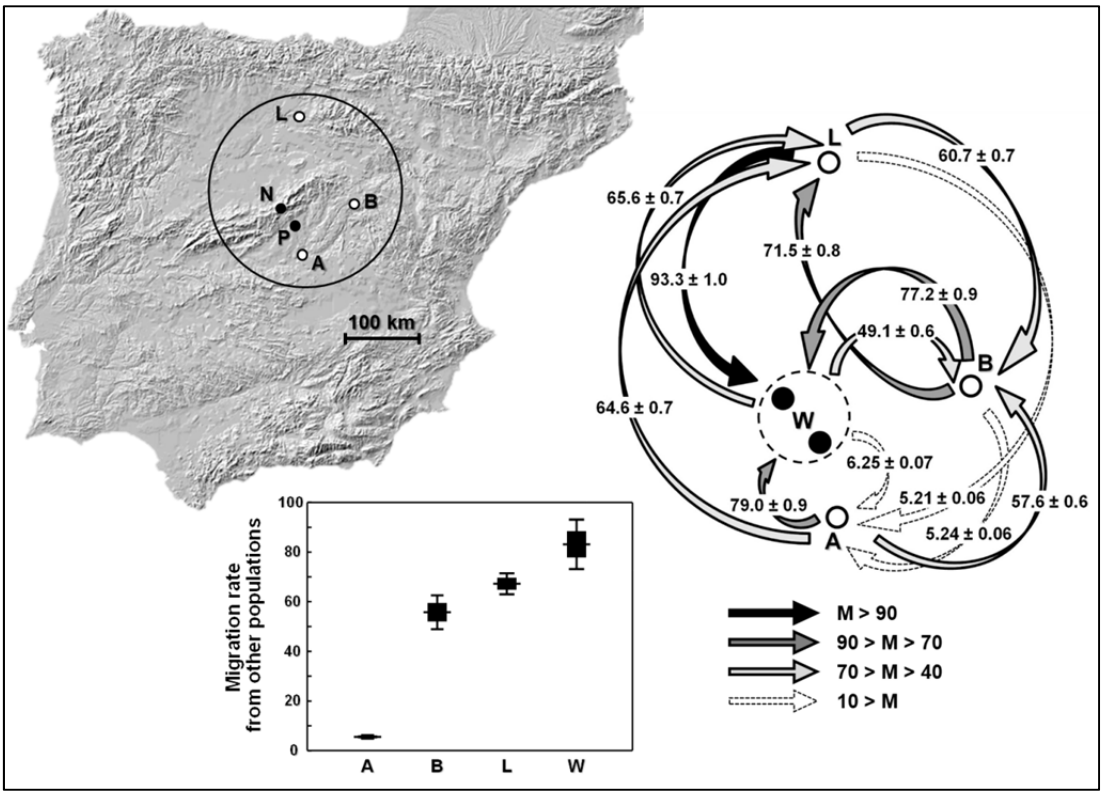


Figure 3. Maximum likelihood estimates of gene flow revealing asymmetrical patterns of migration among lizard populations and genetic isolation in Aranjuez.

clusters that did not match the western populations (Fig. 2b). The population of Lerma was composed of individuals assigned either to one of the eastern clusters (the one including Brihuega, with a mean ± sd probability of assignment of 0.393 ± 0.147) or to one of the two western groups (probability of assignment of 0.547 ± 0.155). When we forced the model towards higher cv-errors, the next result was for K = 2 (cv-error = 0.686; Fig. 2c), with all lizards being correctly assigned to their mitochondrial lineage except for those from Lerma, that were assigned to the western cluster with mean ± sd probability of 0.613 ± 0.140.

The populations located at the extremes of the genetic gradient identified by

MDS and admixture analyses, Aranjuez and the two western populations, scored high homozygosity (with different alleles approaching fixation) for the outliers detected by ROA. Lerma, and to a lesser extent also Brihuega, scored significantly higher heterozygosity of ROA outliers than expected from overall genome heterozygosity (ANCOVA with population as the factor, H_{ROA} as the dependent variable and H_{ALL} as the covariate: $F_{4,89} = 10.121$, $p < 0.001$; Table 1). However, for the outliers detected by CIOA, Lerma scored significantly higher homozygosity than the other populations (H_{CIOA} as the dependent variable: $F_{4,89} = 5.812$, $p < 0.001$; Table 1). To further support the hypothesis that genetic variants subject to selection have been fixed along the evolutionary history of highly divergent populations (Aranjuez and western), we examined the probabilities of correct assignment of each individual to its lineage using SNP data with either whole genome data (PG, Fig. 2a) or outliers (PO, Fig. 2c). In these populations located at the extremes of the genetic gradient, we found a strong correlation between PO and PG (Aranjuez: $r = 0.823$, $p < 0.001$; western populations: $r = 0.924$, $p < 0.001$). However, the correlation was not significant in Brihuega ($r = 0.294$, $p = 0.236$) or in Lerma ($r = 0.154$, $p = 0.541$).

Discussion

The large Psammodromus provides a paradigmatic example of local adaptation, in which the phenotype of individuals fits better to the expected from habitat characteristics than from population ancestry. Our combined analysis of phenotypic variation, population genetic structure and genetic polymorphisms subject to natural selection, allowed us to explore the sources of genetic variants favored in different ecological contexts in this system.

The analysis of > 73,000 SNPs recovered the east-west division of genetic lineages described before based on the analysis of mitochondrial DNA sequences, thereby confirming the ancestral division of this species on a genome-wide basis (Carranza et al. 2006). Compared to previous studies, our approach boosted the resolution with which the genetic background of lizard populations (and individuals) could be characterized. We found different degrees of population structure within genetic lineages, with hierarchical levels of between-population differentiation within the eastern lineage (in which Aranjuez arose as an evolutionarily distinct population), and absence of differentiation between the two western populations studied (cf. Díaz et al. 2017). Within this genetic framework, our analysis of traits with recognized adaptive value in this species (Díaz et al. 2017) showed a phenotypic gradient with Aranjuez at one end, and the western populations at the other. The lizards from Lerma are phenotypically very similar to western lizards, a pattern that contradicts their unmistakably eastern genetic background, but is consistent with their western-like habitat (Santos et al. 2008 and Supplementary Material). This link between habitat and phenotype was in harmony with the pattern of genetic structuring of the F_{ST} outliers identified in the genome, an observation which substantiates the adaptive value of these genetic variants (Rosenblum et al. 2010, Rhode et al. 2017, Schield et al. 2017). Highly divergent populations (Aranjuez and western) have genetic variants subject to selection that seem to have been fixed along their evolutionary history. The high correlation between PG and PO in these populations further indicates that the position of individuals along the genetic gradient defined by selected alleles (outliers) corresponds to their location along the axis of genomic backgrounds opposing eastern and western lineages. However, the same correlation was not significant in Brihuega or in Lerma, suggesting that in these populations selected alleles do not fit the overall pattern of

divergence subject to selection, either for the presence of new selected alleles (CIOA results in Lerma) or for high heterozygosity in alleles selected during lineage differentiation (ROA results in Brihuega and Lerma). Therefore, our results provide compelling evidence of standing genetic variation contributing to lizards' adaptability (Barrett and Schluter 2008).

The process of local adaptation in Lerma seems to be favored not only by a genetic background consolidated through a long history of local adaptation to changing conditions, but also by the incorporation of genetic novelty (locally or imported from other populations) (Barton and Etheridge 2018). Gene flow among lizard populations was apparently influenced by ecological processes and not just by geographic proximity (Rosenblum 2006; Orsini et al. 2013). For example, gene flow into ecologically extreme populations was asymmetric: Aranjuez was almost completely isolated while western populations scored the greatest levels of immigration. Aranjuez seems thus to be the place where incoming genotypes are more strongly purged out, probably because lizards occupy singular habitat of reduced quality. In forested habitat typical of western populations, the large Psammodromus prefers open forests were sunlit patches are readily available (Santos et al. 2008). In this habitat, genetically eastern lizards would not be selected against as strongly as the western ones may be purged out in shrubby habitat typical of eastern populations. In addition, eastern lizards may find opportunities in western habitat, where higher vegetation cover results into lower detectability by visually-oriented predators (even with their striped phenotype; Díaz et al. 2017), and they can benefit from higher food abundance in forests (Consuegra et al. 2005, Tenan et al. 2017). The opposite is not true: western lizards find poor habitat in open scrubland (such as Aranjuez in our study; Díaz et al. 2017). Also, eastern (striped) lizards could be better adapted to dispersal through unforested, grassy habitat than western (unstriped)

ones (Duckworth 2008, LaRue et al. 2018). Whatever the causes of the asymmetry in migration rates found in our study, unrestricted gene flow from western populations, along with standing variation, would allow eastern populations (especially Lerma) to approach the western lizard phenotype in forested habitat (Phillips 1996, Alleaume-Benharira et al. 2006).

Eastern and western lineages of the large Psammodromus have most likely remained locally adapted during the glacial cycles, maintaining their profound divergence but also great adaptive potential favored by balancing selection in the face of habitat heterogeneity (Barrett and Schluter 2008, Svardal et al. 2015, Gallet et al. 2018). Our outlier analysis (ROA) supports this idea, by documenting both the divergence of populations at the extremes of the environmental gradient, and the maintenance of high heterozygosity in eastern populations whose habitat has seemingly acquired western-like conditions. In addition, admixture may also have been important for maintaining populations adapted to changing local conditions. Thus, gene flow, with variable permeability to immigration conditioned by a dynamic landscape configuration, may have prevented speciation while maintaining strong genetic structure (Räsänen and Hendry 2008, Whitney et al. 2018). This is remarkable for a terrestrial vertebrate with limited dispersal ability after 3 my of lineage divergence. At the same time, the combination of divergence subject to selection, balancing selection, and permeability to gene flow, may have rendered the large Psammodromus a species able to adapt to a wide variety of habitats, a trait which may underlie its success as the most abundant and widespread Iberian lizard.

However, an obvious question is why new genetic variants (CIOA) have evolved in Lerma when this population already had standing genetic variation (at ROA loci highly homozygous in the west) that could be selected to warrant local adaptation to

western-like habitat. This paradox may hide a more complex scenario in which environmental gradients are multidimensional, with other factors (such as predators, parasites or competitors) affecting the dynamics of local adaptation besides the forest-to-open gradient driving the divergence between the two lineages (Levinst 1962, Aguirre-Liguori et al. 2017). After all, if only one dimension of environmental variation is considered, moving away from one end of the gradient inevitably leads to approaching the other end (Lahti et al. 2009).

All in all, the combined use of ROA and CIOA provided compelling evidence of different sources of genetic variation contributing to local adaptation dynamics in a lizard species. We are aware that this was feasible thanks to various singularities of our study system, with populations distributed over a wide, heterogeneous range that has been changing since the split of eastern and western lineages in the late Pliocene. In this biogeographical scenario, adaptive lineage splitting was counterbalanced by standing genetic variation and asymmetrical patterns of gene flow (likely promoted by historical or current environmental instability), which could hamper speciation despite strong genetic structure. Complex systems like this one may thus exemplify how to shed light on a variety of important questions at the interface between local adaptation and the first steps towards speciation (or the lack of it) (Schluter 2001, Rundle and Nosil 2005, Rosenblum 2006, Via 2009). To summarize, the combined use of ROA and CIOA may offer a powerful tool to elucidate how evolution provides the genetic variation that ecology demands throughout the history of a species.

Supplementary material

To quantitatively describe the structure of the habitat occupied by each population we used data from the CORINE Land Cover dataset. We merged the polygons delimited by the sampling points of each locality with the CORINE database layer using QGIS v2.18.16, and within each polygon we recorded the cover of three variables that are important for our sampling species (cover of broadleaved trees, grassland and shrubs; Díaz and Carrascal 1991). To quantify the habitat of each population we computed its score on a principal component analysis (PCA) that combined the three cover variables recorded within each polygon (Table S1). This PCA yielded a single axis that opposed areas with high cover of broadleaved woodlands interspersed with grasslands, to xeric areas dominated by shrub cover (Table S1). Scores on this axis defined a habitat gradient opposing Navacerrada ('forest end') to Aranjuez ('shrubland end'), with Lerma, Brihuega and El Pardo occupying intermediate positions but much closer to Navacerrada than to Aranjuez (Fig. S1).

Table S1. *Factor loadings (with eigenvalue and explained variance) in a principal component analysis* with cover of forests, grasslands and shrublands in the five populations studied

	Factor loadings
Cover of forests	-0,966
Cover of grasslands	-0,702
Cover of shrublands	0,964
Eigenvalue	2,355
Explained Variance	0,785

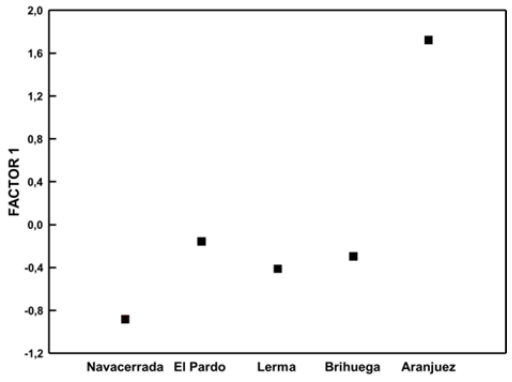


Fig S1. Scores of populations on a habitat principal component.

CAPÍTULO 4. Predicción ajustada del área de distribución de la lagartija colilarga mediante modelos de asociación ambiental con loci sometidos a selección divergente.[4]

Resumen

A lo largo del área de distribución de una especie, los individuos tienen que afrontar retos ecológicos diversos confiando en la variación genética que la evolución les ha proporcionado. Nuestro objetivo es definir el área total de distribución de una especie a partir de la distribución geográfica de unos pocos loci bajo selección (obtenidos mediante una búsqueda a lo largo de todo el genoma) y explorar algunas de las causas de las restricciones adaptativas que subyacen al establecimiento y a la forma de los límites del área de distribución. Muestreamos cinco poblaciones del lacértido mediterráneo *Psammodromus algirus* que, a pesar de encontrarse cerca unas de otras, se distribuyen a lo largo de un gradiente ambiental (definido por la temperatura y la precipitación) muy contrastado. Utilizamos 21 SNPs sujetos a selección y previamente caracterizados para correlacionar el genotipo de 95 individuos con la variación ambiental registrada entre sus poblaciones, utilizando celdas de 1x1 km como unidades de muestreo. Extrapolando el modelo resultante a todas las posibles combinaciones alélicas, inferimos todas las localizaciones geográficas potencialmente aptas para el establecimiento de la especie. El área de distribución inferida coincidió en gran medida el área de distribución descrita para esta especie, permitiendo además una predicción precisa de los huecos internos y, especialmente, de los límites de su distribución. Nuestros resultados sugieren un umbral adaptativo determinado por la cantidad de variación genética disponible que se requeriría para garantizar la adaptación más allá de

[4] Basado en el siguiente artículo: Llanos-Garrido, A., Briega-Álvarez, A., Pérez-Tris, J. y Díaz, J.A. Environmental association modelling with loci under divergent selection accurately predicts the distribution range of a lizard. En estado manuscrito.

un límite concreto de variación ambiental. Como consecuencia, la expansión del área en nuestro sistema estaría ligada al surgimiento de nuevas variantes genéticas susceptibles de ser seleccionadas.

Abstract

All across the distribution range of a species, individuals have to confront contrasting ecological challenges relying on the genetic variation supplied by evolution. Our aim is to define entire species ranges from the geographical distribution of loci under selection (obtained after a genome-wide search for SNP's) and to explore some of the causes of the adaptive boundaries that underlie the formation and shape of range edges. We sampled five populations of the western Mediterranean lizard *Psammodromus algirus* that, despite being close to each other, inhabit a noticeable environmental gradient of temperature-precipitation. We used 21 previously characterized SNPs under selection to correlate the genotypes of 95 individuals with environmental variation among their populations, using 1x1 km^2 grid cells as sampling units. By extrapolating the resulting model to all possible combinations of alleles, we inferred the locations that were suitable for the species. The inferred distribution range overlapped to a large extent with the realized range of the lizard, allowing an accurate prediction of internal gaps and specially range borders. Moreover, the inference of how many genotypes were adapted to each grid cell showed that the borders were suitable for only a few combinations of alleles. Thus, our results suggest an adaptability threshold determined by the amount of genetic variation available that would be required to warrant adaptation beyond a certain limit of environmental variation. As a consequence, range expansion in our system would be ultimately linked to the arising of new variants under selection.

Introduction

Species' borders are commonly assumed to be the result of the constraints imposed by their ecological niches and/or the gradual environmental change to suboptimal conditions towards the edge of their ranges (Hutchinson 1957, Brown 1984). All in all, the factors that shape distribution borders are ultimately linked to local adaptation dynamics; simply put, a species does not occur outside its distribution border because it is not adapted to the conditions beyond it (Kirkpatrick and Barton 1997, Bridle et al. 2007). However, the edge of a species' range is typically more abrupt than expected, given that environmental change towards suboptimal conditions or niche boundaries occurs usually in a gradual way (Sexton et al. 2009). Moreover, all across their ranges species may confront ecological challenges much more demanding than the ones that take place at their distribution borders (Kirkpatrick and Barton 1997). To understand these seemingly arbitrary boundaries to range expansion, Haldane (1956) proposed gene 'swamping' as a center-border effect by which gene flow from central to marginal habitats causes maladaptation at the edges of the range, reducing population density and preventing range expansion. This hypothesis has been subjected to continuous debate (Nosil et al. 2004, Sexton et al. 2011, Polechová 2018).

Evaluating these different models is crucial to understand what are the adaptive causes underlying the formation and shape of range edges (Sexton et al. 2009, Lee-Waw et al. 2018). When ecological niche is the main cause behind range limits, because a species has completely filled its ecological niche, a niche expansion must precede range enlargements (Hutchinson 1957). In such cases, since niche expansion implies adaptation to more extreme conditions along one or more environmental gradients, this process is limited by the magnitude of the additive genetic variance associated with adaptation to these gradients (Lande and Shannon 1996). Conversely, if habitat

suitability remains high at and beyond range boundaries, then dispersal constraints, gene swamping and/or marginal demographic effects could be defining the location and shape of distribution borders (Kirkpatrick et al. 1997, Bridle and Vines 2007, Charlesworth 2009, Peterson et al. 2011). Thus, if the edges of the range are tightly linked to the exhaustion of relevant genetic variance, gene swamping cannot properly explain those limits, because it cannot operate beyond the limits imposed by additive genetic variance at relevant adaptive loci.

Landscape genomics approaches have boosted our understanding of how environmental variables drive the genetic dynamics of local adaptation (Hoban et al. 2016, Ahrens et al. 2018). These methods can be applied to define range boundaries by looking at the shifts in allelic frequencies along environmental gradients (Ecker et al. 2008, Herrera and Bazaga 2008). Thus, it is possible to explore what loci govern the adaptability of a species, and to model the suitability of certain genotypes to different habitats all over a species' range (i. e, environmental association analyses; Lotterhos and Whitlock 2015, Rellstab et al. 2015). However, describing correlations between genotypes and environmental gradients is but one part of the challenge. It is also paramount to identify the loci underpinning local adaptation, since they make the fraction of genetic variation that is relevant to explain an individual's ability to disperse to and thrive in new habitats (Dudaniec et al. 2018). Identifying these combinations of loci under selection is crucial to understand the origin and maintenance of new populations and, therefore, the genetic dynamics that shape range boundaries (Hargreaves et al. 2013). Yet, new populations of a species can be established either by 1) the arrival of individuals carrying genetic adaptations to that new site, or 2) the arrival of genetic variants that can recombine in situ to generate new locally adapted genotypes (Barton and Etheridge 2018). Discerning between these two possibilities is a

hard challenge. In particular, the last scenario is controversial because it assumes that an individual is able to reproduce in a location to which it is not adapted. However, the potential to produce new genetic combinations increases with dispersal rate, by rising the probability that different (suboptimal) genotypes eventually co-occur at the same new habitats (Barton and Etheridge 2018; LaRue et al 2018). Thus, it is important to consider dispersal ability in landscape genomic studies, and to compare systems with different dispersal rates. In particular, study organisms with low dispersal rates should reduce the confounding effects of dispersion, allowing us to focus on local adaptation dynamics (Lee-Yaw et al. 2018).

In this study, we address these questions using data from a lacertid lizard species, the Large Psammodromus *Psammodromus algirus*, whose phylogeographical and ecological differentiation is well characterized (Díaz et al. 2017). This lizard is widespread across the Western Mediterranean region, and its range encompasses contrasting environmental conditions, extending from northern Africa in the south to southwest France in the north, and from Portugal in the west to Tunisia in the east (Fig. 1). We used 21 loci putatively under selection (hereafter outliers; Llanos-Garrido et al. under review) to model distribution boundaries on the basis of five nearby populations that cover a representative fraction of the environmental variation faced by *P. algirus* across its entire distribution range. To this end, we run an environmental association analysis (Rellstab et al. 2015) with allelic variants at loci under selection as predictors, and we extrapolated, for all possible allelic combinations at those loci, the geographical locations with suitable environmental conditions. By doing so, we were able to infer not only an ecological niche model of the whole distribution range of the species, but also the genotypes that would potentially be adapted to each geographical grid cell within it. We assumed a simple model without center-border biases, and in which every genotype

is able to reach every geographic cell. Also, we only used the (adaptive) genetic variation likely to be associated with environmental conditions, in such way that we could know whether actual range limits are linked to adaptability thresholds determined by the amount of additive genetic variance available for selection. This approach allowed us to test whether a species' distribution range can be explained by the genetic dynamics that shape local adaptation at the individual level, without invoking demographic processes such as gene swamping or increased homozigosity near the edge of the range (Herrera and Bazaga 2008, Polechova et al. 2009).

Figure 1. Known distribution range of *Psammodromus algirus* (based on Bons and Geniez 1996 [north-west Africa] and Pleguezuelos 1997 [Iberian Peninsula]). Orange circles mark the location of the sampling populations. The gaps within the range that correspond to croplands or cities are not considered. A question mark is placed where the species is known to inhabit but there is no accurate information about its distribution (22 presences scattered all over North Africa). The discontinuous line defines the assumed southern edge of the distribution range according to IUCN.

Specifically, we aimed to answer the following questions: 1) Is it possible to infer an entire distribution range on the basis of a limited number or outliers? 2) How important are limitations to dispersal in defining a species' range limits? 3) Are there

fewer genotypes adapted to marginal conditions than to core conditions (which would difficult adaptation to range edges)? And 4) is there an adaptability threshold, determined by the availability of genetic variance under selection, that constrains the expansion of the range beyond its actual boundaries?

Materials and Methods

Study system

Psammodromus algirus is a ground-dwelling, heliothermic lizard from the Western Mediterranean region whose distribution range encompasses a wide variety of habitats, from arid shrublands to temperate forests (Díaz and Carrascal 1991). In the Iberian Peninsula, where *P. algirus* is the most abundant and widespread lizard species, climatic heterogeneity is mirrored by broad changes in vegetation patterns: forests dominate in the west of its range, whereas shrublands prevail in the east. The genetic diversity of this species is broadly structured in two mtDNA lineages, eastern and western, which diverged ca. 3-3.5 mya (Carranza et al. 2006, Verdú-Ricoy et al. 2010). These lineages show some degree of ecologically-driven divergence, because eastern lizards typically display a striped dorsal pattern absent among western ones; striped and unstriped phenotypes seem to be adaptively linked to crypsis in the predominant habitat where lizards live (Díaz et al. 2017).

We sampled 95 lizards in five populations along a significant environmental gradient in the center of the Iberian Peninsula, covering both mtDNA lineages (Fig.1). Three sampling sites were of eastern mtDNA adscription: 1) Lerma (42.058 ºN, -3.611 ºE; 900 m asl), a fragmented mixed forest interspersed with grassland patches, 2) Aranjuez (40.016 ºN, -3.586 ºE; 594 m asl), a hot, dry site with a high cover of herbs

and shrubs and no trees, and 3) Brihuega (40.778 ºN, -2.911 ºE; 1,009 m asl), a deciduous open forest with a mosaic of grassland and woodland patches. Two additional sampling sites belonged to the western lineage: 4) El Pardo (40.511 ºN, -3.755 ºE; 658 m asl), a xeric, lowland evergreen forest, and 5) Navacerrada (40.726 ºN, -4.023 ºE; 1,230 m asl), a montane location covered by deciduous forest. Several particularities of these populations make them representative of a wide range of selective pressures gathered around the core of this species' range: 1) lizards from Lerma inhabit a very fragmented forest archipelago that resembles the typical habitat of western lizards (although they belong to the eastern lineage) (see Díaz et al. 2005; Santos et al. 2008, 2009; Tellería et al. 2011 for further information about habitat fragmentation effects in this system); 2) Aranjuez lizards inhabit the typical hot and dry habitat of eastern lizards, and although this locality is very close to the western populations included in this study (El Pardo and Navacerrada), it receives very little gene flow from them (Díaz et al. 2017), so that its isolated condition promotes the accumulation of genetic divergence subject to selection (Llanos-Garrido et al. under review); and 3) the two western populations are separated by a significant altitudinal gradient, and although lizards from both populations show little genetic differentiation (Díaz et al. 2017), they differ in important phenotypic traits such as escape tactics, sexual dimorphism, sexual ornaments, ectoparasite loads and other life history traits (Iraeta et al. 2006, 2010, 2011, and 2013; Capítulo 1: Llanos-Garrido et al. 2017).

DNA extraction, sequencing and variant calling, and outlier analyses

The 21 loci under selection used in this study were detected by outlier search analyses conducted in a previous study (Llanos-Garrido et al. under review). Shortly, we obtained tissue samples by clipping 2 cm of the tail tip of lizards, which were afterwards released at their site of capture. We purified DNA for library preparation

using the Speedtools Tissue DNA Extraction kit (Biotools).

We used the restriction enzyme Pst1 for GBS library preparation. Sequencing was done in an Illumina HiSeq2500 sequencer. To recover SNPs we used the pipeline UNEAK, implemented in TASSEL v.3.0 (Bradbury et al. 2007), which is specifically designed for samples with no reference genome. We aligned sequence tags to each other to form 'networks' of tags, where each node is a single tag sequence, and each edge represents a single base pair difference between two tags. We pruned the networks to remove putative sequencing errors (low frequency alleles) using the error rate threshold parameter. We also discarded loci with minor allele frequencies < 0.01 or that could be successfully sequenced in less than 10% of individuals. The resulting dataset had 73,291 biallelic SNPs (Single Nucleotide Polymorphism), a site depth of 6.60 ± 6.75 and a site missingness of 0.42 ± 0.31.

To minimize false positives in outlier analyses, we discarded loci that could not be successfully sequenced from at least 75% of individuals in each population, and loci with minor allele frequencies < 0.05 in each population, thus excluding all private alleles from the dataset. The resulting SNP dataset included 6,421 loci. We used a Bayesian approach to perform an outlier analysis as implemented in Bayescan v.2.1 (Foll and Gaggiotti 2008). Bayescan uses a logistic regression model to partition F_{ST} coefficients into a population-specific term (β) and a locus-specific term (α). We selected loci with $\alpha > 0$ as suggesting positive selection, and a false discovery rate (corrected for multiple testing) $q < 0.05$. To obtain these parameters, we ran the MCMC algorithm implemented in the program with a prior odd value of 10, and using 20 pilot runs of 5,000 iterations each, followed by 100,000 iterations with a burn-in of 50,000 interactions. In order to search for outliers while accounting for coancestry effects, we performed a second outlier analysis using the Bonhomme et al. (2010) extension of the

Lewontin-Krakauer test. We selected loci based on the statistical significance of the FLK statistic, with a restrictive significance threshold of $p < 0.001$ to account for multiple testing.

The outlier analysis performed with Bayescan detected 12 outlier loci with $\alpha > 0$ ($0.97 < \alpha < 1.35$) and $q < 0.05$, while the FLK analysis identified nine additional loci with $p < 0.001$, none of which was previously detected by Bayescan (Llanos-Garrido et al., under review).

Quantification of environmental variation

To quantify environmental variation all over the potential range of the species, we selected an area that included its actual distribution range plus a 450-850 km wide perimeter belt around it. Within this area, we used data from the Bioclim 2.0 dataset (cell resolution = 1x1 km; Fick and Hijmans 2017). We computed the score of each cell on a principal component analysis that combined all Bioclim dataset variables using R core. This PCA yielded a principal axis that opposed hot areas with low precipitation to temperate ones with high precipitation (results not shown).

Environmental Association Analysis (EAA)

Environmental values (dependent variable for EAA models) were assigned to 1x1 km grid cells (sampling units, N = 93) with QGIS v2.18.16 (QGIS Development Team 2018) using a layer of PCA-scores within polygons defined by sampling locations. The genotypes for each loci (independent variables for EAA models) were recoded as 0, 1 or 2 depending on whether they were homozygous for the reference allele, heterozygous or homozygous for the alternative allele, respectively (with reference and alternative alleles arbitrarily defined by the order of appearance in the process of variant calling).

Our EAA was constrained by the fact that the 95 individuals genotyped

belonged to only five different populations. While this ensures a proper characterization of genetic variation within populations, it leads to unavoidable pseudoreplication of environmental data (and, depending on the extent of genetic differentiation and aggregation, also of genetic data). To circumvent this problem, we used a triple randomization approach. Firstly, we based our EAA on the random assignment of genotypes within sampled populations to 1x1 km grid cells. Given that environmental variation is several orders of magnitude larger among populations than within them (97.6 % of the variance in PCA scores explained by population adscription), we chose to prioritize the among-populations component of the models by assuming that all genotypes could occupy every grid cell within the geographical boundaries of their own population. This is more realistic than assuming a large component of genotype-environment covariation at a local, within-population scale. Moreover, the genetic component of such covariation could not be detected by our methods of outlier detection, which were specifically designed to search for genetic divergence among populations. Our first set of 1,000 intra-population randomizations should therefore capture the extant association between genotypic and environmental variation at the scale of the sampled gradient.

Secondly, we randomized 1,000 times the geographical grid cell assigned to each genotype, but without taking into account its population. This allowed us to produce a null hypothesis of no association between genotypic and environmental variation, but which takes into account the fact that environmental values are pseudoreplicated. By comparing the distribution of predictive powers of the models obtained in both randomizations, we could check the extent to which predictive power was higher in the first [intra-population] set than in the second [inter-population randomization]) one.

Thirdly, we performed a test of genetic randomization to control for the potential effects of genetic structure among populations and genetic aggregation within them. For that purpose, we constructed 1,000 new sets by randomly selecting 21 loci from each genotype (i.e. the number of detected outliers) but without taking into account whether they were outliers or not. This was done to account for the fact that neutral genetic variation (random SNPs) is expected to have the same degree of aggregation than variation under selection (outliers). However, while we should expect that at least a fraction of the genetic variation subject to selection should be correlated with environmental variation, the opposite is true for the neutral differentiation of populations.

We performed a backward stepwise multiple regression analyses for each randomized data set (N = 3,000 EAAs), with SNPs as predictors and environmental scores as the dependent variable using *lm* function in R core. By doing this, we obtained a distribution of adjusted R^2 levels and P-values for each set of randomizations. Final model building was achieved by considering the mean p-values of partial correlations (N = 1,000) calculated for all the datasets obtained with the intra-population randomization strategy. In each step, we removed all SNPs with a mean p-value > 0.5, and we recalculated all partial correlations with the remaining SNPs. In the last step, when all remaining markers had mean p-values < 0.5, we removed all SNPs with mean p-values > 0.05. Our final model (genotype-environment association model, or GEAM) was built with the mean intercept and mean beta values of the remaining SNPs.

Range inference

To infer the distribution range of the species, we followed a two-steps procedure. Firstly, we used R coding to include all the geographical cells that presented the same environmental scores than the sampled populations (predicted range #1). This first

approach provides a baseline prediction with no genetic information that can be used to quantify the improvement in predictive ability supplied by GEAM. Secondly, we considered all possible combinations of alleles for the outlier loci selected by GEAM (i.e., all possible genotypes under selection) to predict all environmental values suitable for at least one genotype according to GEAM. By fulfilling all grid cells with those environmental values, we could extrapolate our prediction to the overall distribution range of the species. Finally, we removed from the inferred range a few disconnected patches (in France, coastal Italy, and the Mediterranean islands) that were too far from the main distribution range of *P. algirus*, whose low dispersal rate (Santos et al. 2009) is supported by the fact that genetic differentiation can be detected even among forest fragments separated by 350 m of unsuitable arable land (Pérez-tris et al. submitted). This produced our second (and final) inferred distribution range (predicted range #2).

The extent of overlap between real and predicted distribution ranges was estimated using QGIS v2.18.16 (QGIS Development Team 2018).

Results

Environmental Association Analysis

The PCA with all Bioclim environmental variables (N = 19 variables) yielded a single principal component (eigenvalue = 0.679) that retained four variables (annual mean temperature [BIO1], max temperature of the warmest month [BIO5], mean temperature of the warmest quarter [BIO10] and annual precipitation [BIO12]) defining a bioclimatic gradient with a hot and dry extreme in the area occupied by the Sahara desert (highest values) and a temperate and wet extreme in northwestern Spain (lowest values; Fig. 2).

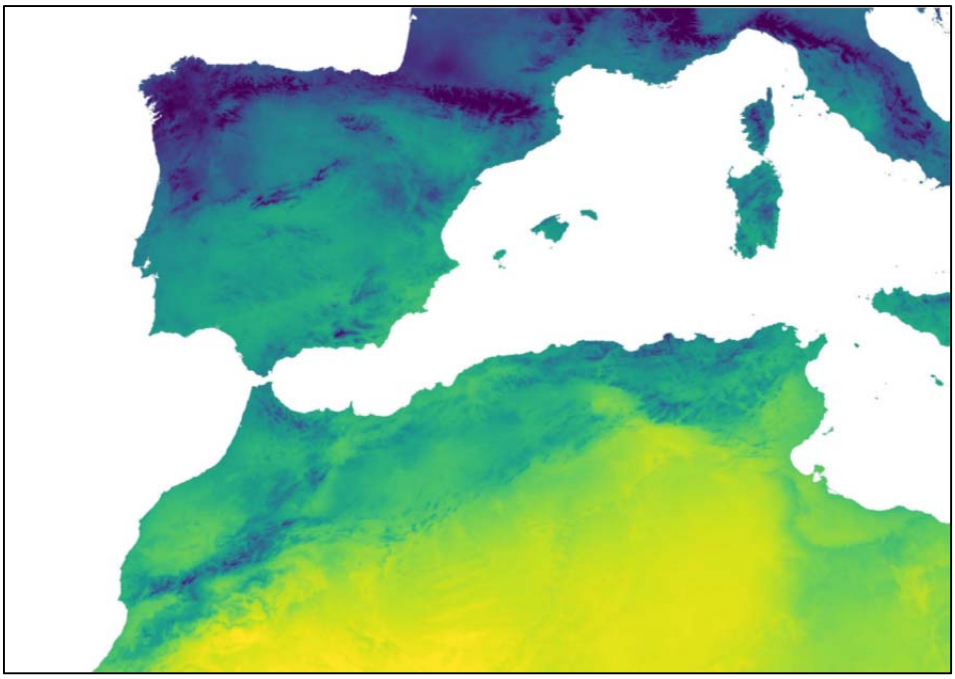


Figure 2. Bioclimatic gradient defined by the environmental PCA for all the western Mediterranean region (potential distribution range of *P. algirus*).

A vast majority of inter-population and genetic randomizations resulted in non-significant models (Fig. 3). Mean adjusted R^2 for the random assignment of genotypes to populations (inter-population randomization) was 0.002 (SD = 0.038, range = 0 - 0.199), with a mean p-value of 0.487, (SD = 0.292, range = 0.0004 - 0.999); only 6% of the 1,000 randomized datasets yielded significant results. Similarly, mean adjusted R^2 for genetic randomization (i.e. random within-population selection of SNPs, either outliers or not) was 0.002 (SD = 0.105, range = 0 - 0.324), with a mean p-value of 0.494; SD = 0.296, range = 0.001- 1); only 5.6 % of the datasets yielded significant models. Conversely, all the datasets built by intra-population randomization produced highly significant models (mean adjusted R^2 = 0.646 ± 0.007, range = 0.623 - 0.670; mean p-value ± SD =$3.5 \times 10^{-17} \pm 2.8 \times 10^{-17}$, range = $2.02 \times 10^{-18} – 2.93 \times 10^{-16}$; Fig. 3). Thus, the environmental association models including the SNPs under selection were 323 times more explicative than those built with the other two randomization strategies.

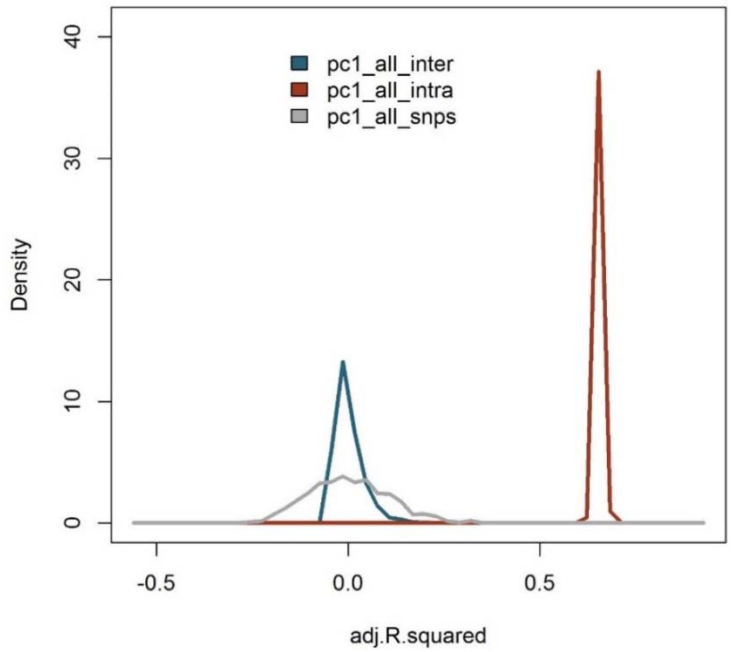


Figure 3. Distributions of adjusted R^2 values for the regression models obtained with three different randomized datasets (see text for details).

The final individual-based genotype-environment association model (GEAM) included four SNPs with significant partial correlations (Table 1). Of these, one (SNP3) was inferred by Bayescan (standard outlier detection), while the rest of them were inferred by FLK (co-ancestry independent outlier detection).

Table 1. Parameter estimates for the regression coefficients of the SNPs under selection that entered the final model.

	Parameter estimate ± SD (min, max)	P-value ± SD (min, max)
Intercept	- 3.630 ± 0.033 (- 3.739, - 3.539)	$2.42 \times 10^{-22} \pm 2.03 \times 10^{-22}$ (1.08×10^{-23}, 1.92×10^{-21})
SNP1	+ 0.605 ± 0.012 (0.567, 0.642)	$3.31 \times 10^{-7} \pm 1.85 \times 10^{-7}$ (6.62×10^{-8}, 1.46×10^{-6})
SNP2	+ 0.488 ± 0.019 (0.430, 0.551)	$5.55 \times 10^{-4} \pm 2.66 \times 10^{-4}$ (1.01×10^{-4}, 1.71×10^{-3})
SNP3	+ 0.347 ± 0.011 (0.315, 0.385)	$3.87 \times 10^{-4} \pm 1.78 \times 10^{-4}$ (7.31×10^{-5}, 1.49×10^{-3})
SNP4	+ 0.366 ± 0.010 (0.328, 0.396)	$2.81 \times 10^{-4} \pm 1.03 \times 10^{-4}$ (7.41×10^{-5}, 8.68×10^{-4})

Range inference

Of the overall number of cells of the real species' range, 27.83% had the same environmental scores than the sampled populations (or, in other words, predicted range

#1 allowed to forecast 27.83% of the species' range); 25.56% of predicted range #1 fell outside real range limits. Predicted range #2 (the range inferred by extrapolating GEAM to include all grid cells suitable for any possible combination of alleles at the four loci in the final model) was similar to the real species' range. All grid cells in predicted range #1 were included in predicted range #2, accounting for 36.64 % of its total amount. In turn, predicted range #2 captured 75.09 % of the actual distribution of *P. algirus*, with 13.41 % of inferred presences beyond real range limits (Fig. 4). Of the 24.91% fraction of the real distribution range unpredicted by GEAM, a large proportion corresponded to the northwest corner of the range, as well as to a large number of small predicted gaps within it (Fig. 4). Nevertheless, predicted range #2 accurately reflected not only the northern and southern edges of the real distribution range, but also many of the gaps within it, both in northwest Africa (i.e. far from the sampled populations) and in many Iberian mountain ranges (Fig. 4).

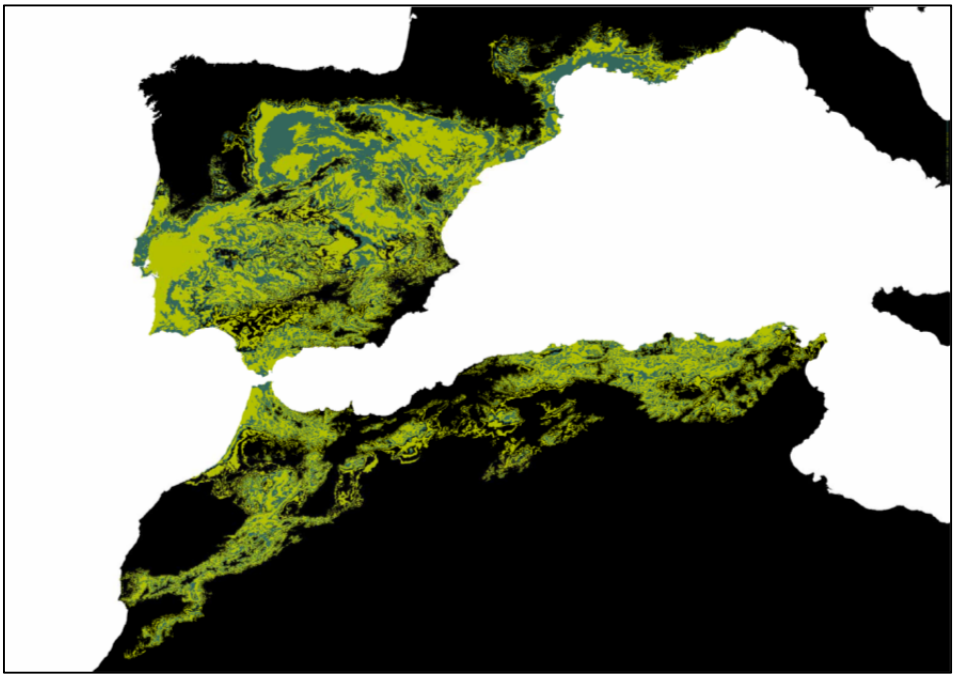


Figure 4. Inferred distribution range. In dark green predicted range #1 (grid cells with the same environmental scores than the sampled populations) and in clear green predicted range #2 (inferred by extrapolating GEAM to any possible combination of alleles at the loci included in the final model).

Discussion

Extrapolating genotype-environment association analyses to all possible combinations of alleles at a few loci subject to selection allowed us to explore how much environmental variation could be exploited by a given amount of genetic variability. This, in turn, revealed an adaptability threshold that ultimately defined the distribution boundaries of *P. algirus*. Thus, our successful inference of a species' range from the geographical distribution of a few adaptive loci uncovered the role of individual genotypic uniqueness in shaping a species' distribution range, which provides compelling evidence of how the genetic dynamics of local adaptation underlie distribution patterns.

Environmental Association Analysis

Our intra-population randomization approach showed that the predictive power of GEAM was more than 300 times larger than expected by chance. Such high predictive power, based on the genetic diversity found in only five populations, depended critically on the selection of an appropriate sample of populations. Thus, our sampled populations included both the eastern and western lineages into which *P. algirus* is divided (Carranza et al. 2006, Díaz et al. 2017), with among-populations geographical distances correlated with neither genetical nor environmental distances. Yet, a small number of inter-population and genetic randomizations also yielded significant models, suggesting a certain degree of environmental pseudoreplication and genetic aggregation in our data. However, the rate of significance was close to 5%, i.e. the conventional level of type I error rate for significance in statistical tests. As a consequence, and given the low standard error of parameter estimates (Table 1), our final genotype-environment association model should be regarded as very robust.

Range inference accuracy

Our approach combines the advantages of correlational and mechanistic distribution models (Kearney and Porter 2009), because it relies on genotype-environment correlations based on the geographical distribution of field-captured individuals, but deals with SNPs under selection that should ultimately underlie functional differences in morphology, physiology, and/or behavior. This should allow us to enlarge the scale of our analysis to the species level (Buckley 2010), because the extrapolation of our results to all possible allele combinations at loci under selection covers a much wider range of adaptive phenotypic variation than the one revealed by the physiological measurements of a restricted sample of individuals or populations. Moreover, as long as local adaptation leads to a heterogeneous distribution of the genotypes adapted to different parts of the range, our GEAM should be more realistic than mechanistic models. This is because mechanistic models are based on physiological measurements of individuals that may lack the specific adaptations required to thrive in specific habitats different from their own, such as the ones that determine range boundaries (Svardal et al. 2015).

Genotype-based range inference was especially accurate at the southern edge of the species' range, including a precise delimitation of range gaps in Morocco, where detailed corological information is available (Bons and Geniez, 1996). However, we could not test the accuracy of our model for the rest of North Africa due to the lack of detailed distribution maps of *Psammodromus algirus* in this area. The only information about these locations was obtained from the IUCN Red List database (which does not provide data about within-range gaps) and the GBIF database (which has only 22 records in this area, all of which were predicted by our model). Nevertheless, the distribution borders suggested by these databases were accurately predicted by our range inference.

Regarding northern boundaries in the Iberian Peninsula, where detailed corological data are also available (Pleguezuelos 1997), we did not recover the presence of the species in a relatively large NW area where lizard populations do occur, inhabiting suitable habitat patches near the cool, humid end of the tested environmental gradient. This is probably because our outlier analyses did not capture all the genetic variation under selection that is associated with such gradient (see below). Across southern France, the real distribution range of the species does not exceed the Rhône River delta, a geographical barrier which could not be predicted by our method of range inference. However, in the Iberian Peninsula our model successfully recovered the central and eastern parts of the northern range boundary, as well as several within-range gaps associated with mountain ranges (around central plateaus and river valleys) and arid regions in the SE.

The role of niche boundaries and dispersal limitations in shaping range limits

The accuracy of our prediction of range limits reflects that, overall, these were revealing ecological niche boundaries. If other environmental factors (e.g. prey, competitors, predators, parasites, etc.) had been constraining range expansions, our inferred range would have extended beyond real range boundaries, and realized range edges would be explained by the existence of limitations to expansion before fulfilling all cells within the spatial projection of the species' niche (Holt 2003). In fact, this happened only in the northeastern border of the range, where GEAM predicted the presence of *P. algirus* beyond the Rhône River. As steted before, the Rhône delta creates a geographical barrier to expansion, whereas niche boundaries would actually allow the species to reach Italy.

Also, our results lead us to dismiss the existence of demographical center-border

effects such as gene swamping or increased homozigosity near the edge of the range (Herrera and Bazaga 2008, Polechova et al. 2009, Pironon et al. 2017). This is because if these processes were acting, they would be limiting the persistence of marginal populations near range boundaries, and we would systematically infer false positives beyond range limits (Case and Taper 2000, Bridle and Vines 2007, Lee-Yaw et al. 2018). However, we did wrongly infer a relatively large inland area of false positives at the northern side of the Pyrenees (Figs. 4 and 5). Interestingly, this area was suitable for a small number (≤ 3) of genotypes, which provides a reasonable explanation for these false positives, because it seems unlikely that the few genotypes that could be adapted to these unoccupied areas were available in nearby marginal populations (Pujol et al. 2009, Dawson et al. 2010, Barton, and Etheridge 2018).

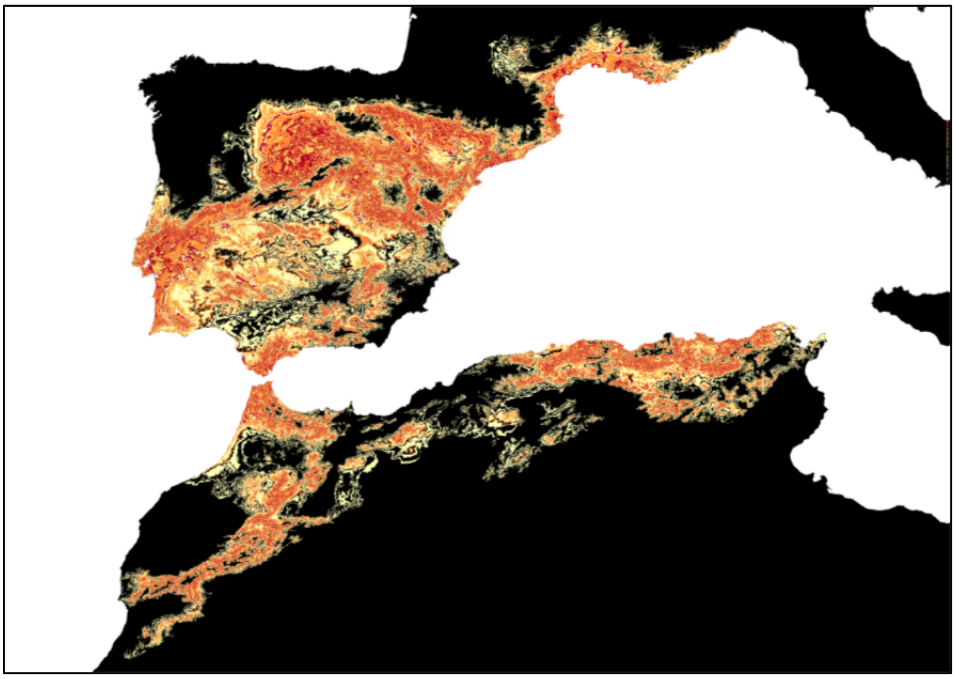


Figure 5. Temperature map representing the number of adapted genotypes per grid cell according to our GEAM.

The dispersal ability of the genotypes arising at range margins plays an important role in the colonization of new areas beyond range limits (Simmons and

Thomas 2004, Hardie and Hutchings 2010). In our system, for example, the range of *P. algirus* would extend ca. 13% beyond its eastern European border if lizards were able to disperse across the Rhone River. Furthermore, genetic diversity would be fostered by greater dispersion abilities (Duckworth 2008), which would facilitate the co-occurrence of adaptive allele combinations at range margins. For instance, the low dispersal rate of these small terrestrial ectotherms would complicate their expansion towards suitable but unoccupied areas north of the Pyrenees (false positives in our model), given the synergistic effects of low dispersal rates and a low probability of finding adapted genotypes at nearby marginal populations. To clarify the role of dispersal ability in the colonization of new locations beyond range limits, our genotype-based modelling approach should be applied to species showing different dispersal abilities (Sanford et al. 2006, Dawson et al. 2010), or we could perform transplant experiments beyond range limits (i.e. manipulating species' dispersal ability; Hargreaves et al. 2018). By doing this, we should be able to discern between the role of dispersal ability *per se* and the genetic contribution of pre-adapted genotypes arising at marginal populations (or the lack of them) (Bridle and Vines 2007, Sexton et al. 2009, Phillipsen et al. 2015).

Number of adapted genotypes per cell across the species' range

Besides the general pattern by which less genotypes were adapted to range boundaries than to core areas, we found two distinct scenarios within the Iberian Peninsula. On the one hand, in the northern half of the peninsula there were many areas suitable for a high number of genotypes. In such context, there is a high probability of finding the few genotypes that are adapted to the challenging environmental conditions characteristic of northern boundaries, which should facilitate the establishment of the marginal populations that shape the corresponding edge (Kawecki 2008, Hardie and Hutchings 2010, Halbritter et al. 2015).

Conversely, most of the southern half of the peninsula seemed to be suitable only for a small number of genotypes, despite the fact that *P. algirus* is abundant in this area. However, several small areas locally suitable for many genotypes were interspersed all across the region. Such areas could therefore play an important role as sources of specific genetic diversity adapted to the demanding, singular environments that surround them (Holt and Keitt 2005; Sagarin et al. 2006). The genotyping of populations that inhabit demanding environments, suitable for a small number of allele combinations, would be crucial to sustain this assertion (Eckert et al. 2008, Gallet et al. 2018). Similarly, a model simulating the intensity of selection in both sources and sinks of genetic diversity should be useful to test whether the genetic variants adapted to demanding environments arise with higher probability in source populations with more relaxed selection regimes (Alleaume-Benharira et al. 2006).

Adaptability thresholds constrain range limits

Our results suggest that species' ranges are determined by the maximum possible span of the environmental values to which a certain number of loci subject to selection are adapted. As a consequence, range expansions should be constrained by adaptability thresholds. In our system, it seems that the environmental range to which *P. algirus* is adapted is ultimately linked to the amount of genetic variance under selection associated to a specific precipitation/temperature bioclimatic gradient. If positive, a range expansion promoted by adaptations towards more extreme environments should entail the selection of new genetic variants. Moreover, such range expansion would require that the effect of the new adaptive mutations is additive with respect to the ones that define the adaptability threshold (Polechova et al. 2015, Polechova 2018). Whilst our results support this line of reasoning, further theoretical exploration is needed to uncover the hypothesized positive relationship between the magnitude of the increment

in environmentally correlated additive genetic variance, and the extent of range expansion that can be achieved (Angert et al. 2008, Polechova et al. 2009).

Overall, we have shown that inferring species' ranges from the geographical distribution of SNPs under selection can be not only very accurate, but also truly informative about the genetic dynamics that underlie local adaptation all over a species' range. Our results point out towards the amount of genetic variability subject to selection as the ultimate determinant of the location and shape of range boundaries in our system. This conclusion sheds light on the basic knowledge required to understand the arrangement and evolution of species' distribution ranges and range boundaries (Connallon and Sgrò 2018).

CAPÍTULO 5. Baja aptitud asociada a la erosión genética en una población fragmentada de lagartijas.[5]

Resumen

El aislamiento poblacional en fragmentos de hábitat disminuye la aptitud a través de la pérdida de diversidad genética y de la depresión por endogamia. Así por ejemplo, se espera que los individuos homocigotos en poblaciones erosionadas como efecto de la fragmentación de hábitat alcancen niveles bajos de aptitud. En una población fragmentada de lagartija colilarga (*Psammodromus algirus*) en el norte de España, encontramos evidencias de estructura genética a una escala geográfica pequeña (140-12,800 m), posiblemente relacionada con la reducción del intercambio genético entre poblaciones acantonadas en los fragmentos forestales y separadas por una matriz de cultivos cerealistas que constituyen un hábitat inhóspito para la especie. Las lagartijas del mismo fragmento de hábitat estaban más relacionadas genéticamente unas con otras de lo que se esperaría por azar, y la homocigosidad individual fue mayor en los fragmentos pequeños que en los grandes. Dentro de los fragmentos, la homocigosidad individual se correlacionó negativamente con el tamaño corporal de los adultos y con la fecundidad de las hembras (peso relativo de la puesta), revelando una conexión entre la reducción del flujo genético, el incremento de la homocigosidad, y la disminución de la aptitud; esta conexión permitiría explicar de forma determinista la reducción de la viabilidad poblacional que tiene lugar en los hábitats fragmentados, reforzando las evidencias disponibles acerca del impacto de la pérdida de diversidad genética en poblaciones naturales fragmentadas.

[5] Basado en el siguiente artículo: Pérez-Tris, J., Llanos-Garrido, A., Bloor, P., Carbonell, R., Tellería, J., Santos, T. y Díaz, J.A. Low fitness associated to genetic erosion in a fragmented lizard population. Enviado a Ecology and Evolution (primera revisión).

Abstract

Population isolation in habitat fragments is believed to lower fitness through the loss of genetic diversity and inbreeding depression. For example, homozygous individuals are expected to achieve low fitness in populations that have become genetically eroded as an effect of anthropogenic habitat fragmentation. In a fragmented population of the large psammodromus lizard *Psammodromus algirus* in northern Spain, we found evidence of genetic structuring at a very small geographic scale (distance between habitat fragments ranged 140-12,800 m), which was compatible with restricted gene flow among populations disconnected in a matrix of inhospitable habitat. Lizards from the same habitat fragment were genetically more related to one another than expected by chance, and individual homozygosity was greater in small than in large fragments. Within fragments, individual homozygosity was negatively associated with adult body size and fecundity, revealing a link among reduced gene flow, increased homozygosity and lowered fitness that may deterministically reduce population viability in fragmented habitat. Such an observation adds up to mounting evidence of the impact of the loss of genetic diversity on fragmented wild populations.

Introduction

Habitat fragmentation, the anthropogenic decrease in size and spatial connectivity of habitat fragments, is a major driver of the ongoing biodiversity crisis (Fahrig, 2003; Foley, 2005; Pimm et al., 2014). Fragmentation not only involves the effective loss of habitat, but also hinders the movement of individuals among habitat fragments, which may compromise species persistence if reduced dispersal prevents the restoration of local extinction events or reduces gene flow among disconnected populations (Hanski

1998, Nathan et al. 2017). Isolated populations may lose genetic diversity owing to genetic drift after population bottlenecks, an effect which can be aggravated if immigration from other habitat fragments is infrequent and this circumstance forces individuals to breed with close relatives (Young et al. 1996). In turn, genetic erosion may jeopardize population viability due to inbreeding depression, or low fitness as a result of breeding of related individuals (Westemeier et al. 1998; Crnokrak & Roff, 1999; Reed & Frankham, 2003), an effect which has long been believed to exacerbate extinction risks in fragmented habitat (Frankham, 2005; Wootton & Pfister, 2013). Inbreeding depression occurs when fixation rates of deleterious alleles are higher than rates of genetic rescue by migration or genetic purging by selection; the larger this difference, which is enhanced by the loss of connectivity between fragments, the larger the magnitude of the inbreeding depression caused by fragmentation (Hedrick & Garcia-Dorado, 2016).

Inbreeding depression has been found in a variety of wild populations (Kardos et al. 2016), yet its effects on population dynamics have been historically debated because they may be masked or effectively overridden by environmental and demographic processes (Lande, 1988; Caro & Laurenson, 1994; Frankham, 2010; but see Wootton & Pfister, 2013). Ecologists and conservation biologists are therefore urged to determine the true magnitude and fitness effects of inbreeding effectively caused by loss of genetic diversity in wild populations (Keller & Waller, 2002; Frankham, 2010), which is especially true when populations are threatened by habitat fragmentation. For example, restricted gene flow among disconnected population fragments is a well-documented phenomenon (Dixo et al. 2009), and reduced demographic growth of small populations inhabiting fragmented habitat has been attributed to negative genetic effects (Hanski and Saccheri, 2006), including a decrease in the ability of fragmented populations to

adapt to environmental change (Yates and Fraser, 2014).

In fragmented populations, inbreeding may be common if subpopulations are small and dispersal among habitat fragments is rare. This circumstance would oblige individuals to mate within their habitat patch, which would increase genetic relatedness of individuals inhabiting the same patch and increase average homozygosity as a consequence. Increased individual homozygosity may contribute to inbreeding depression if homozygous individuals have problems to cope with life challenges such as parasite resistance (Coltman et al. 1999) or competition for mates (Hoffman et al. 2007), or show poor lifetime performance in terms of fecundity, longevity or offspring quality (Huisman et al. 2016). Mounting evidence of negative homozygosity-fitness correlations (Hansson and Westerberg, 2002; Coltman and Slate, 2003; Reed and Frankham, 2003; Chapman et al. 2009; Huisman et al. 2016) supports the idea that such correlations can be used to estimate the impact of inbreeding in nature, which may be particularly true in fragmented populations where inbreeding is expected to be stronger than usual (Hansson and Westerberg, 2002; Szulkin et al. 2010).

Populations of the large psammodromus *Psammodromus algirus* persist in fragmented habitat close to the edge of the species distribution range in northern Spain (Santos et al. 2008). In that region, this forest lizard is absent from many small habitat fragments that are otherwise suitable for the sustenance of populations, according to known habitat preferences of the species (Díaz et al. 2000). In addition, lizards show lower breeding success in small than in large habitat fragments (Díaz et al. 2005). However, the question remains as to why habitat fragmentation may hamper breeding performance of these lizards. We analysed if low fitness of lizards in fragmented habitat is associated with loss of genetic diversity due to isolation in small habitat fragments. We explored four lines of evidence of the negative genetic effects of forest

fragmentation, expecting to find (1) population structure among population fragments owing to reduced gene flow, (2) reduced out-crossing opportunities within habitat fragments, (3) increased individual homozygosity in small population fragments, and (4) negative correlations between individual homozygosity and fitness.

Methods

During May-June in 2001 and 2002, we studied lizards in a fragmented forest located near Lerma, in northern Spain (42º06'N, 03º40'W), where suitable habitat fragments are separated by inhospitable cultivated land (Fig. 1). The fragmentation of continuous habitat to open cultivation fields in the past, which was especially intense after the Spanish Civil War (1936-1939), led to the formation of an archipelago of deciduous or evergreen woodlots. We captured lizards from three large habitat fragments (> 200 ha; CW, CE and Q) and from 14 small fragments (< 10 ha), which represent over 70% of the known species range in this locality (we did not obtain lizard samples from seven sites with known presence of the species). To obtain adequate sample sizes, these habitat fragments were grouped into five different geographically defined habitat sectors (Fig. 1). These five groupings were defined on the basis of geographical proximity measured using the geometric centre of sites (between-site connectivity is inversely proportional to the distance between sites). To make the groupings objective, we conducted a cluster analysis of the matrix of among-site geographic distances, using Euclidean distances and complete linkage. Habitat fragments within the resulting clusters were separated by a maximum of 350 m of arable land.

Overall, we captured 131 lizards (mean sample size ± 1 SE = 20.6 ± 4.8 individuals per habitat sector, range = 14 - 42). The individuals analysed had been

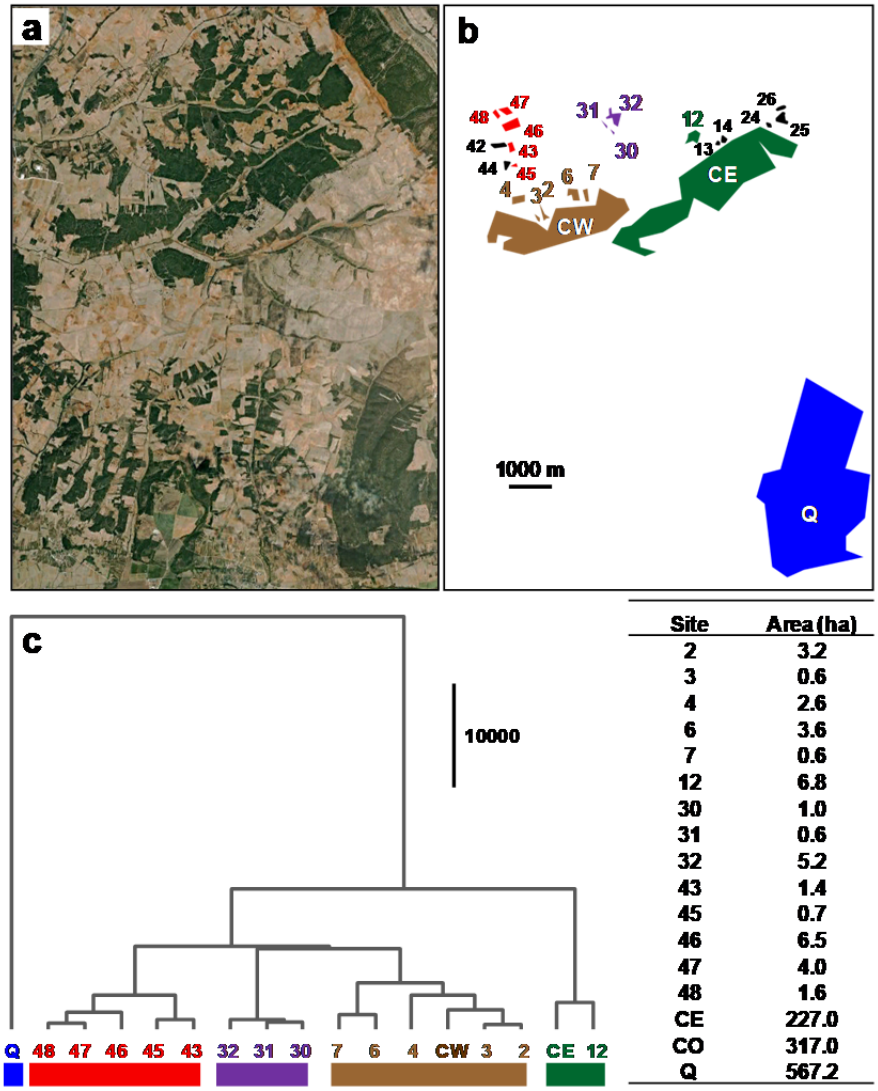


Figure 1. The fragmented forest of Lerma (a) and the 17 habitat fragments where lizards were captured (b), which represent 71% of sites where the species is known to occur in the area (seven not sampled sites are also shown in black). The cluster tree (c) shows Euclidean linkage distances (the vertical bar shows the scale) used to define five habitat sectors (coded with colours) based on proximity between sampling sites. The name according to published nomenclature (Santos et al. 2008) and the size (area in ha) of each sampling site is also shown.

included in an earlier study of the breeding performance of lizards (Díaz et al. 2005), which allowed us to measure various traits associated with individual fitness (see below), thereby offering a unique oportunity to assess the correlations between fitness traits and individual homozygosity. The 14 small forest fragments are separated by variable distances from the two large forest fragments CW and CE. The most distant

site (Q, which is the largest habitat fragment) was also sampled because it had been considered to be a potential source of colonizers of the fragmented landscape in previous analyses of the regional distribution of lizards in this area (Díaz et al. 2000; Santos et al. 2008), and it also contributed data to the analysis of breeding investment of lizards (Díaz et al. 2005).

In the field, we recorded snout-to-vent length as a measure of lizard body size. We also collected tissue for DNA analyses by removing approximately 5 mm of lizards' tail tip. Tissue was stored in ethanol until DNA analysis. Four immature individuals (i.e. individuals not reaching the species' minimum body size at maturation) that were sampled for DNA analysis were excluded from the analysis of variation in body size. Captured gravid females (n = 48) were kept captive until they laid their clutches in the laboratory. All individuals were returned to site of capture after the study. A detailed description of field and husbandry methods can be found elsewhere (Díaz et al. 2005). Sampling work, measuring methods and tissue collection were performed in accordance with relevant guidelines and regulations approved by the Junta de Castilla y León (Consejería de Medio Ambiente) under license E.P. 117/2001.

We used DNA extracts to genotype all individuals for eight microsatellite loci specifically developed for *P. algirus* as described elsewhere (Bloor and Dávila, 2008), only six of which were polymorphic in lizards from Lerma. We are aware that our analysis of population genetic structure and individual homozygosity based on microsatellite markers is less powerful than the current high-throughput sequencing methods for genome-wide typing (Hoffman et al. 2014; Huisman et al. 2016). However, we typed our lizards in 2006, before these methods became widely accessible, and when we later tried to recover next-generation sequencing data from the same samples, the analysis was unfeasible because the samples had lost quality (not enough DNA of

sufficient quality remaining). Therefore, we were forced to explore the relationships between genotype and fitness traits based on only six microsatellite loci. Nevertheless, the utility of microsatellites has been validated with large SNP datasets both for fitness-heterozygosity correlations and for population genetic differentiation and structure (Putman and Carbone, 2014). In addition, even a small number of microsatellite loci can provide much insight into the genetic features of individuals and populations (Forstmeier et al. 2012), as shown for example in an experiment with *Anolis* lizards that analysed population genetic structure and individual homozygosity using six microsatellite loci (Kolbe et al. 2012). However, the inference of genome-wide properties based on limited microsatellite data remains challenging due to uncertainty about the mutational processes generating genetic diversity at different loci, dependence on genomic context (e.g. if mutation rates vary widely among genomic regions), or lack of power if the number of markers is too low as it happens to be our case (Putman and Carbone, 2014). Given the impossibility to gain next-generation sequencing information from our samples, but also considering the importance of genotyping individuals that had been scored for different fitness components in the laboratory, we decided to validate our measures of genetic diversity using a completely new sample of lizards for which we could obtain both microsatellite data and estimates of genome-wide diversity based on SNP calling. For that purpose, in 2015 we genotyped 17 lizards captured in 2008-2009 in the same area, using the same set of six microsatellites, and also SNPs sequenced on an Illumina HiSeq2500 sequencer after GBS library preparation. Filtering and SNP calling were performed using the UNEAK pipeline (Lu et al. 2013), specifically designed for samples with no reference genome and implemented in TASSEL v.3.0 (Bradbury et al. 2007). Sequence tags were aligned to each other to form 'networks' of tags, where each node is a single tag sequence, and each edge represents a

single base pair difference between two tags. The networks were pruned to remove putative sequencing errors (low frequency alleles) using the error rate threshold parameter. We tested for Linkage Disequilibrium (LD) as a measure of correlation between loci (r^2) using the pruning criteria --*indep-pairwise* of plink software (Purcell et al. 2007). We obtained 73,291 loci, with a site depth of 6.6 ± 6.75 (mean ± sd) and a site missingness of 0.424 ± 0.313.

An analysis with GENEPOP (Rousset, 2008) of the six microsatellite markers did not detect significant deviations from Hardy-Weinberg equilibrium or pairwise linkage disequilibrium within habitat patches. Population differentiation and isolation-by-distance effects were estimated using Arlequin 3.5 (Excoffier and Lischer, 2010), using R_{ST}-like genetic distances (Slatkin, 1995), although using standard F_{ST}-like genetic distances produced the same results. We computed maximum likelihood estimates of relatedness between pairs of individuals using the ML-relate software (Kalinowski et al. 2006). We used one-way ANOVA to test for variation in pairwise relatedness between individuals sampled in the same site or in different sites, but the significance of the analysis was computed as the probability of obtaining as large or greater F values than the observed one in 10,000 random matrices in which the site of origin of each individual was randomly assigned.

Homozygosity of individual lizards was measured using uncorrected homozygosity, internal relatedness, and homozygosity by loci, which were calculated using CERNICALIN (Aparicio et al. 2006). Uncorrected homozygosity (H_O) was calculated as the proportion of loci at which an individual is homozygous. Internal relatedness (*IR*), which weighs on the basis of the frequency of the alleles, was calculated as $IR = (2H - \Sigma f_i)/(2N - \Sigma f_i)$, where H is the number of loci that are homozygous, N is the number of loci and f_i is the frequency of the ith allele contained in

the genotype. Finally, homozygosity by loci (*HL*), was estimated to weigh the contribution of each locus to the homozygosity value depending on its allelic variability (Aparicio et al. 2006). Homozygosity by loci is calculated as $HL = (\sum E_h)/ (\sum E_h + \sum E_j)$, where E_h and E_j are the expected heterozygosities of the loci that an individual bears in homozygosis (h) and in heterozygosis (j), respectively. These three estimates of homozygosity were highly correlated in our sample ($r > 0.97$, $n = 131$, $P < 0.0001$), and therefore we chose one of them (*HL*) for the analyses (using the other two estimates of homozygosity produced the same results).

We used a hierarchically nested variance components design to test for within-sector effects of patch size (small versus large habitat fragments) on lizard pairwise relatedness and homozygosity. In these analyses, sector was a random factor, and type of habitat fragment (small versus large) was a fix factor. We used Satterthwaite's method of denominator synthesis to find the linear combinations of sources of random variation that serve as appropriate error terms for each effect. Thus, the degrees of freedom for the denominator mean square can be fractional rather than integer values, meaning that fractions of sources of variation were used in synthesizing error terms for significance testing (StaSoft Inc, 2014). We used the same modelling approach to analyse homozygosity-fitness correlations using individual body size and clutch mass relative to female size as estimates of fitness. Body size indicates fitness in this population because larger lizards survive better and larger females lay more eggs (Díaz et al. 2005). Clutch mass controlling for the effect of female body size combines the contributions of egg number (or fecundity) and egg size (or offspring quality) to total reproductive output, given that these two variables are positively correlated in this population, while removing the effect of female body size on realised breeding investment (Díaz et al. 2005).

Results

The correlation between microsatellite and genome-wide measures of homozigosity was highly significant ($r = 0.730$, $p < 0.001$; Supplementary Fig. S1), meaning that heterozygosity estimates obtained with data of six microsatellite loci reliably measure genome-wide heterozygosity as scored using > 70k SNPs. However, with only 6 microsatellite loci there could be a chance for a single locus to be driving the effect. To exclude this possibility, we repeated the correlation six times, each time excluding one locus. Results remained significant in all cases (all $r > 0.627$ and all $p < 0.007$). It should also be noted that our conclusion is valid independently of the distribution of microsatellites along the genome (i.e. whether each locus is more or less linked to low-recombination regions of the genome).

Based on microsatellite analyses, we found population structure among 17 forest fragments separated from one another by 140-12,800 m in a matrix of inhospitable habitat. In a hierarchical AMOVA, variation among five habitat sectors (groups of forest fragments separated from other similar groups, Fig. 1) explained 5.6% of total genetic variance ($F_{CT} = 0.056$, $P < 0.0001$). However, populations from fragments within the same sector showed no sign of genetic isolation ($F_{SC} = 0$, $P = 0.830$). Isolation by distance could partly explain the pattern of among-site genetic differentiation within the study area, because the greatest forest fragment (Q) is distantly located from all other sites (Fig. 1) and shows the greatest average F_{ST} in pairwise population comparisons (comparisons with Q site $F_{ST} = 0.037$, mean ± SE of comparisons between other sites $F_{ST} = 0.013 \pm 0.003$). However, a Mantel test with 10,000 simulations of random distance matrices only found a weak and non-significant isolation-by-distance effect ($r = 0.178$, $P = 0.076$), which was clearly removed with the exclusion of fragment Q from the analysis ($r = 0.06$, $P = 0.29$), although population

genetic structure among habitat sectors still explained a significant 4.3% of variance (F_{CT} = 0.043, P = 0.013).

An analysis of maximum likelihood estimates of pairwise genetic relatedness revealed that lizards from the same forest fragment were genetically more related to one another than expected by chance (difference in pairwise relatedness between lizards from the same site compared to lizards from different sites: observed F = 76.55, mean ± SE estimated from 10,000 simulations of random distributions of individuals among sites: F = 1.02 ± 1.01; P < 0.0001; Fig. 2). We did not find variation in pairwise relatedness between lizards from small or large forest fragments ($F_{1,14.62}$ = 0.02, P = 0.888). However, lizards were more homozygous in smaller than in larger forest fragments ($F_{1,14.43}$ = 16.04, P = 0.0012; Fig. 3).

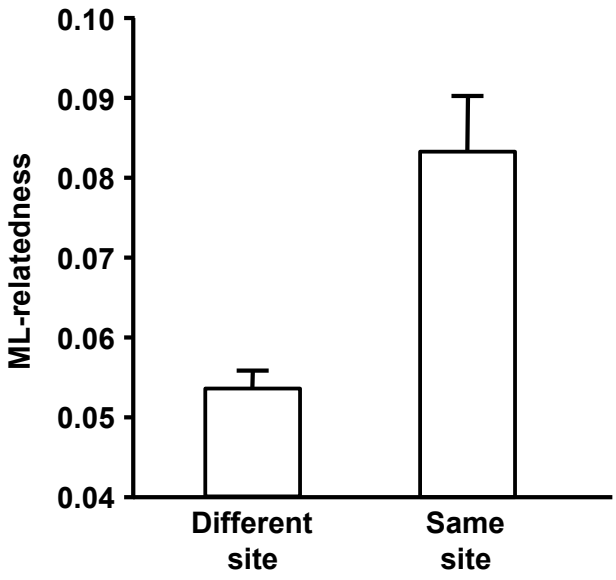


Figure 2. Average pairwise relatedness (maximum likelihood relatedness estimates based on allele sharing) between lizards captured in different or in the same habitat fragments (means ± 95% confidence intervals).

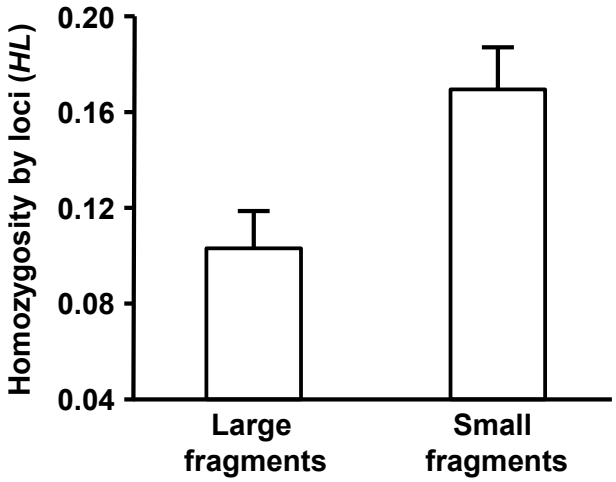


Figure 3. Average homozygosity values (means ± SE) among fully grown lizards captured in large or small forest fragments.

Controlling for the effects of sex and site (nested within forest type, because lizards grow faster in deciduous than in evergreen fragments (Iraeta et al. 2006), the former being more productive than the later (Santos et al. 2008)), homozygosity was negatively correlated with body size of adult lizards (Table 1). Moreover, female homozygosity was negatively correlated with clutch mass while controlling for the effects of site (nested within forest type) and female body size (Table 1).

Table 1. Homozygosity-fitness correlations based on estimates of homozygosity by loci (HL) of lizards. Effects of homozygosity on body size (measured as log snout-to-vent length) were estimated controlling for sex (females tended to have larger snout-to-vent length than males). Effects of homozygosity on clutch mass (log-transformed) were estimated controlling for body size (log snout-to-vent length). In both cases, site (nested within forest type, i.e. deciduous vs. evergreen forests) was included as a random factor to control for possible environmental differences among habitat fragments.

	Body size			**Clutch mass**		
	Beta	F (g.l.)	*P*	Beta	$F_{1,35}$	*P*
Homozygosity (HL)	-0.219	6.00 (1, 102)	0.016	-0.248	5.10 (1, 35)	0.030
Sex		3.92 (1, 102)	0.050	0.852	62.83 (1, 35)	< 0.001
Site (nested within forest type)		1.90 (16, 102)	0.034		1.79 (10, 35)	0.099

Discussion

Our results provide empirical evidence of the links among reduced gene flow, increased individual homozygosity, and reduced fitness of individuals in fragmented habitat. Other studies conducted at the individual level have reported reduced fitness of inbred individuals in fragmented habitat (Johansson et al. 2007; Reed et al. 2011). In our study, habitat fragments preserve what is left of a lizard population that was geographically unstructured a few decades ago, but has recently been fragmented by agricultural practices. In this context, genetic effects of habitat fragmentation (reduced gene flow and inbreeding) emerged with detrimental consequences for individuals. Thus, fragmentation-induced homozygosity was negatively correlated with two important components of lizards' fitness: adult body size and reproductive investment. The observed negative correlation between homozygosity and fitness may reflect a lifelong impact of individual genetic variability on viability and reproductive performance of lizards (Johansson et al. 2007; Huisman et al. 2016). Reduced reproductive value of homozygous lizards in fragmented habitat helps to directly link population genetic erosion to low individual fitness of lizards in small habitat fragments (Díaz et al. 2005). In a broader context, this kind of observation helps to interpret low population growth rates often observed in fragmented populations as a result of negative genetic effects (Hanski and Saccheri, 2006; Wootton and Pfister, 2013).

In our study, genetic structure could be viewed as the consequence of infrequent interchange of lizards among habitat fragments, an interpretation which was supported by genetic evidence that individuals had limited out-crossing opportunities. Habitat fragmentation theory predicts out-crossing opportunities to fall down as population size decreases with decreasing fragment size. Supporting this prediction, we found that lizards were more homozygous in smaller than in larger forest fragments. Therefore,

fragmentation may reduce genetic variability of lizards by frequent inbreeding between fragment-locked individuals, combined with stochastic effects of bottlenecks and genetic drift, which may exacerbate the loss of genetic variation in the smallest habitat fragments (Young et al. 1996). Evidence of inbreeding depression was provided by a negative correlation between individual homozygosity and two relevant components of lizards' fitness: adult body size, which directly depends on growth rate and age and determines survival, male attractiveness and female fecundity in this species (Martin and Forsman, 1999; Díaz et al. 2005), and female breeding investment measured as clutch mass relative to female size (Sinervo and Licht, 1991).

Our results provide compelling evidence of genetic deterioration as the cause of reduced individual fitness in this fragmented lizard population (Díaz et al. 2005). We lack data of population size in our study fragments because this species is extremely hard to census (Díaz et al. 2000; Santos et al. 2008). Therefore, we cannot provide direct evidence of negative population trends associated with the genetic effects detected in our study. However, circumstantial evidence supports the view that the association between population fragmentation, individual homozygosity and low fitness is already having a negative demographic impact on this lizard population. In our study area fragmentation increases westwards, and the westernmost part of the forest archipelago is separated from the eastern part by a motorway (Santos et al. 2008). Earlier studies of lizards in the area failed to find lizards in large (> 10 ha) fragments of favourable habitat located west of the motorway, although the size and structural characteristics of these forest fragments predicted the presence of the species (Díaz et al. 2000; Santos et al. 2008). Given that the motorway was built only 25 years ago, when the highway A–1 (E–50 according to European nomenclature) substituted an unfenced, single carriageway road (N–1), we interpret the apparent absence of lizards on the

western side of the road as a barrier effect of the motorway preventing restoration of locally extinct populations (Tellería et al. 2011), which reinforces the view that gene flow is important for the long-term sustenance of lizard populations in this area. Further supporting this idea, reintroduced lizards thrive in fragments where populations apparently went extinct in the recent past: we translocated genetically unrelated lab-born juveniles to a 4.1-ha fragment where lizards had never been seen before, and the population remained viable (as shown by the presence of adult and juvenile individuals) five years after the introduction (Santos et al. 2009).

Importantly, homozygosity was negatively correlated to lizard fitness in both large and small habitat fragments. Therefore, our results support the view that negative genetic effects may reduce population viability before population size becomes too small (Spielman et al. 2004), which may compromise the adaptability of populations in fragmented habitat (Yates and Fraser 2014, Cheptou et al. 2017). In our study area, populations are less likely found in evergreen than in deciduous forest fragments of the same size, arguably because deciduous habitat has higher thermal quality and more abundant food (Iraeta et al. 2006, Santos et al. 2008). Therefore, lizard may have reduced viability and fecundity in poor habitat (Díaz et al. 2005, Iraeta et al. 2006), where extinction risk may therefore increase for populations with lowered genetic variability. Understanding the links among reduced gene flow, individual genetic condition and fitness in a broad range of taxa may be useful if we are to rescue populations on the verge of extinction in fragmented habitat (Hedrick and Garcia-Dorado 2016). Such knowledge may also help policymakers to decide when management actions aimed at favouring out-breeding, such as translocation of unrelated individuals (Madsen et al. 1999, Santos et al. 2009), will help to enhance population viability.

Supplementary material

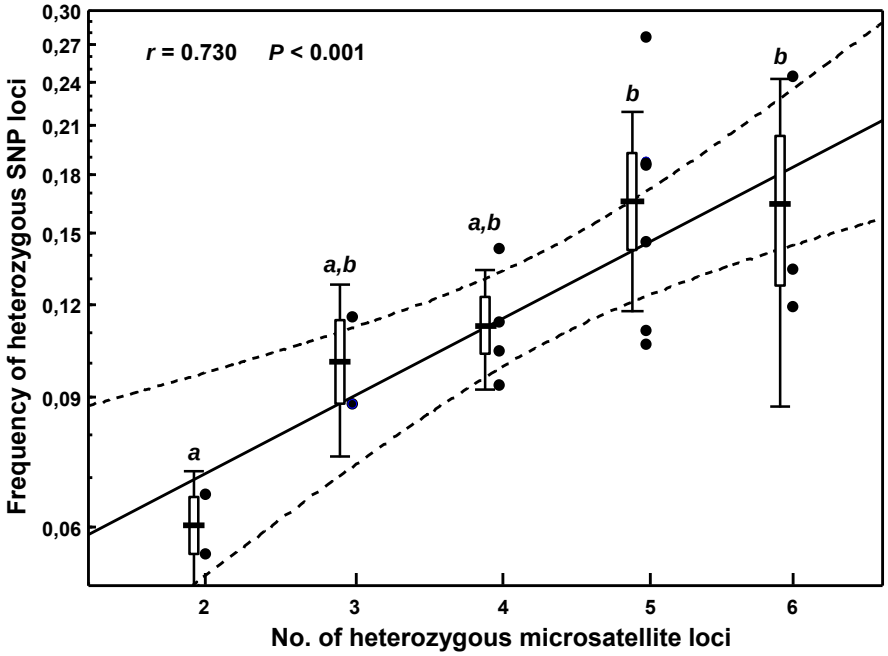


Figure S1. Regression of genome-wide (> 70k loci) against microsatellite (6 loci) measures of heterozigosity. Mean values (with SE and 95% CI) of SNP-heterozigosities corresponding to different numbers of heterozygous microsatellites are also shown. Superscripts indicate homogeneous groups according to Tukey's HSD test for unequal sample sizes (ANOVA: $F_{4, 13} = 4.86$, $P < 0.013$)

3. DISCUSIÓN GENERAL

Determinantes eco-inmunológicos potencialmente relacionados con la adaptación local

Como era de esperar, dados los resultados previos con otros saurios, los machos de lagartija colilarga de Navacerrada presentaron más carga parasitaria que las hembras (Brace et al. 2015, Reedy et al. 2016). Estas diferencias entre sexos pueden explicarse por el compromiso entre los niveles de testosterona en sangre (que promueven la actividad y la ornamentación sexual; Salvador et al. 1996) y la inmunocompetencia (Folstad y Karter 1992). Los niveles de testosterona en sangre presentan un ciclo estacional en esta especie de lagartija en el que la concentración de testosterona en sangre aumenta hasta el final de la época reproductora según se van desarrollando los ornamentos sexuales de los machos (Díaz et al. 1994). Este ciclo estacional es perfectamente congruente con la fenología de la infestación por garrapatas y el desarrollo de la coloración sexual, lo que apoya de manera sólida el compromiso entre este aspecto de la ornamentación sexual y la carga de parásitos en machos (Malo et al. 2008, Pollock et al. 2012). Además, a esta evidencia correlacional del compromiso entre sistema inmune y ornamentación sexual, se le suma nuestra demostración empírica de dicho compromiso en la que mostramos cómo la activación del sistema inmune desvía recursos que se utilizan para montar un ornamento sexual.

Nuestros experimentos ponen de manifiesto también que el aspecto ornamental que se encuentra comprometido con la activación del sistema inmune (y que por lo tanto es el que resulta informativo acerca de la calidad inmune del portador; Hamilton y Zuk 1982) varía entre la población parasitada y la que no lo está. Esto pone de manifiesto que la presencia de parásitos puede influir en esa respuesta inmune, de manera que si un individuo se encuentra parasitado, desvía recursos inmunológicos que de otra manera podría utilizar para aumentar la superficie coloreada de rojo y se ve forzado a utilizar

otro aspecto ornamental (saturación de rojo) que informe de su calidad individual en señales menos conspicuas (Kekalainen et al. 2014). Esta afirmación está apoyada en el hecho de que cada uno de esos aspectos se encuentra correlacionado positivamente con el tamaño de los machos en la población en la que aparentemente está siendo utilizados como señal, y asociado negativamente con la capacidad de respuesta inflamatoria de los mismos. Por lo tanto, nuestros resultados sugieren que la presencia de parásitos puede restringir la expresión de ornamentos sexuales hasta el punto de regular qué características son más informativas acerca de la calidad de sus portadores.

Este tipo de aproximaciones pueden ayudar a entender mejor cuál es el papel de los parásitos como moduladores de los efectos fenotípicos de la selección sexual. Las diferencias eco-inmunológicas entre El Pardo y Navacerrada podrían afectar de forma directa al aislamiento reproductivo (Endler y Houde 1995, Giery y Layman 2015), promoviendo así el proceso de adaptación local que parece que gobierna la divergencia fenotípica entre estas poblaciones. Esto es así porque cuando las presiones selectivas divergentes que subyacen a la adaptación local están relacionadas estrechamente con la selección sexual, como es nuestro caso, es posible que los ornamentos sexuales evolucionen de manera divergente en respuesta a dichas presiones (Endler y Houde 1995).

Búsqueda de patrones de diferenciación genética congruentes con procesos de adaptación local

Algunas de las diferencias fenotípicas adaptativas encontradas en otros estudios entre El Pardo y Navacerrada (Iraeta et al. 2006, 2010, 2011 y 2013) se confirmaron en el segundo capítulo de esta tesis, apoyando así la evidencia creciente de procesos adaptativos divergentes entre estas dos poblaciones. En estos estudios anteriores, se encontraron diferencias en rasgos distribuidos a lo largo de todo el ciclo de vida de las

lagartijas y además se pudieron controlar los efectos ambientales en situaciones de ambiente común (*common garden*) y trasplante recíproco, lo que pone de manifiesto la existencia de una base genética divergente entre esas poblaciones que pudiera estar detrás de este aparente patrón de adaptación local (De Kort et al. 2014). Por otra parte, pudimos confirmar la inexistencia de estructura genética neutral entre estas dos poblaciones de lagartijas (Verdú-Ricoy et al. 2010, Díaz et al. 2017), lo que apunta a un proceso de divergencia muy reciente o a un flujo genético intenso entre ellas. En este caso, no podemos evaluar la intensidad de flujo genético entre estas poblaciones porque el hecho de no poder identificar estructura genética entre ellas hace que este cálculo no tenga sentido (Phillips et al. 2008). Estos resultados contradicen aparentemente a los obtenidos en el capítulo 1 de la tesis, pero cabe destacar que la divergencia reciente puede todavía desembocar en procesos de aislamiento reproductivo mediados por las diferencias inmunológicas descritas en ese capítulo (Endler y Houde 1995).

Esta indiferenciación genética neutral no debería suponer un problema a la hora de identificar loci sujetos a selección natural, ya que los análisis que tratan de encontrar estas variantes se basan en encontrar islas o picos de selección cuya divergencia local resalta por encima del resto del genoma (Aguillon et al. 2018). Sin embargo, parece que la divergencia entre las poblaciones es tan exigua que no existe aún ninguna huella divergente en el genoma que permita la detección de esos picos de divergencia, al menos con las técnicas empleadas en este estudio (Aguillon et al. 2018). Por lo tanto, parece que los procesos adaptativos locales que puedan diferenciar estas dos poblaciones (descritos en el capítulo 1 y en estudios anteriores) no han dejado todavía una marca genética lo suficientemente generalizada en el genoma como para que puedan detectarse con la potencia de muestreo genético que ofrecen los métodos empleados en esta tesis (Joo et al. 2016).

Combinación de métodos de detección de outliers para descifrar las fuentes de diversidad genética que subyacen a la adaptación local

En el capítulo 3, la búsqueda de diferenciación genética sujeta a selección sí que se realizó entre poblaciones bien diferenciadas genéticamente, como muestran los resultados de los análisis de estructura genética neutral realizados en este capítulo. Estos análisis confirmaron la división entre linajes descrita previamente (Carranza et al. 2006) con una resolución mucho mayor (encontramos estructura genética también dentro del linaje del este). Además, en este caso sí se encontraron diferentes SNPs cuyo patrón de divergencia se ajustaba a lo esperado bajo selección natural y que fue congruente con el patrón de diferenciación fenotípica y ambiental entre las poblaciones muestreadas (Rosenblum et al. 2010, Rhode et al. 2017, Schield et al. 2017).

Gracias a la combinación de dos métodos distintos de detección de outliers, conseguimos discernir entre variantes genéticas sujetas a selección que se arrastran a lo largo de la historia evolutiva de los linajes y las que parecen evolucionar de manera independiente al patrón filogenético. Las poblaciones más divergentes presentaron SNPs sujetos a selección que parecían haberse fijado a lo largo de su historia evolutiva, pero las poblaciones que ocupaban posiciones ambientales y fenotípicas más intermedias no parecían seguir ese patrón. El proceso de adaptación local en Lerma, una población oriental con fenotipo y ambiente propios del linaje occidental, parece haberse visto favorecido tanto de mantener alelos alternativos seleccionados de forma divergente a lo largo de la evolución de los dos linajes como de innovaciones genéticas seleccionadas localmente o importadas a través del flujo genético con otras poblaciones (Barrett y Schluter 2008, Barton y Etheridge 2018); el resultado final es un proceso de adaptación convergente en el que se ven favorecidas variantes genéticas que proceden de distintos momentos de la historia evolutiva de la población.

Además, los resultados de los análisis de detección de outliers que no tienen en cuenta el patrón de co-ancestría, muestran cómo los dos linajes de lagartija colilarga no sólo son capaces de mantenerse localmente adaptados gracias a procesos de selección divergente, sino que además parece que el potencial adaptativo ante nuevos retos ecológicos se ve favorecido gracias a procesos de selección equilibradora (Barrett y Schluter 2008, Svardal et al. 2015, Gallet et al. 2018). Además de esto, el hecho de que el flujo genético mantenga una permeabilidad variable condicionada por la propia adaptación local de cada población, puede prevenir procesos de especiación a la vez que se mantiene una estructura genética patente (este-oeste) en este sistema (Räsänen y Hendry 2008; Whitney et al. 2018).

Modelos de asociación ambiental con loci sujetos a selección para inferir áreas de distribución de especies

Gracias a la congruencia espacial entre variabilidad ecológica y genotípica sujeta a selección descrita en nuestro sistema en el capítulo 3, fuimos capaces de realizar un modelo de asociación ambiental muy robusto que incluía parte de esa variabilidad genética sujeta a selección detectada en ese mismo capítulo (Rellstab et al. 2015). Este modelo resultó tan robusto porque las poblaciones que se incluían en él eran buenos representantes de las condiciones adaptativas de los dos linajes de lagartija, pero también porque las distancias geográficas no estaban correlacionadas con la distancia genética ni ambiental. Esto nos permitió realizar una inferencia del área de distribución de la especie bastante precisa (sobre todo para los huecos dentro del área en Marruecos y para la definición de los límites del área) que combinaba las ventajas de los modelos de distribución mecanísticos y correlacionales (Kearney y Porter 2009). Extrapolando el modelo de asociación genotipo-ambiente a todas las combinaciones posibles de alelos, conseguimos agrandar la escala de nuestro análisis al nivel de especie (Buckley 2010),

además de ser capaces de identificar qué combinaciones de alelos están adaptadas a qué celdas geográficas de acuerdo con nuestro modelo. Esto reveló posibles patrones de ocupación de ambientes solo aptos para unos pocos genotipos a partir de regiones que funcionarían como fuente de diversidad en contextos de selección natural más relajada (Holt y Keitt 2005, Sagarin et al. 2006).

Además, revelamos que el rango máximo de valores ambientales para los que un número concreto de loci sujetos a selección puede resultar adaptativo concordaba con los límites de distribución reales de nuestra especie modelo. Esto sugiere que es la propia cantidad de variabilidad genética sujeta a selección la que resulta realmente determinante a la hora de definir la forma y localización de los límites del área de distribución de una especie (Angert et al. 2008, Polechova et al. 2009). Si este fuera el caso, una expansión de área debería ir precedida por la selección de nuevas variantes genéticas cuyo efecto resultara aditivo con respecto a las que ya se encuentran definiendo dicho umbral adaptativo (Polechova et al. 2015, Polechova 2018). Por lo tanto, parece que inferir áreas de distribución de especies a partir de la distribución geográfica de SNPs bajo selección puede no sólo resultar muy preciso, sino además verdaderamente informativo acerca de las dinámicas genéticas que subyacen a los procesos de adaptación local a lo largo del área de distribución de una especie.

Relación entre disminución de aptitud, erosión genética y fragmentación del hábitat

En el último capítulo, mostramos evidencia empírica de la asociación entre la reducción de flujo genético, el incremento de la homocigosidad individual y la reducción de aptitud en ambientes fragmentados, lo que constituye una observación crítica para entender realmente los efectos genéticos de la transformación del hábitat antropogénica (Johansson et al. 2007, Reed et al. 2011). En este capítulo, los análisis basados en microsatélites muestran una estructura genética muy marcada, incluso entre fragmentos

de hábitat cercanos, lo que pone de manifiesto un infrecuente intercambio genético entre parches de hábitat que en última instancia desembocaría en la pérdida de diversidad genética por culpa de entrecruzamientos consanguíneos (Young et al. 1996). Este aumento de la consanguinidad a escala local se consigue correlacionar con una disminución en la aptitud de los individuos a través de dos importantes componentes de la misma: el tamaño del adulto (relacionado con la supervivencia, la tasa de crecimiento en edades tempranas, el éxito reproductivo y la fecundidad; Forsman 1999, Díaz et al. 2005) y la inversión reproductiva (Sinervo y Licht, 1991). Este último resultado, analizado con unos pocos marcadores clásicos (6 microsatélites), se ve muy apoyado por el hecho de que la homocigosidad medida con los mismos marcadores, se correlaciona muy significativamente con la medida a través de secuenciaciones al azar y a gran escala a lo largo de todo el genoma.

Por otra parte, aunque el estudio carece de medidas de tendencia demográficas, mostramos una importante línea de evidencia que hace pensar que estos efectos genéticos de la fragmentación del hábitat podrían afectar a la viabilidad de las poblaciones que conforman nuestro sistema (Díaz et al. 2000, Santos et al. 2008). La fragmentación del hábitat en el área de estudio se muestra más severa hacia el Oeste, donde se encuentra una autopista que constituye una importante barrera a la dispersión de las lagartijas (Santos et al. 2008, Tellería et al. 2011). Al oeste de esta autopista se encuentran varios fragmentos de bosque cuyas condiciones ambientales resultan perfectamente aptas para nuestra especie modelo (Díaz et al. 2000, Santos et al. 2008), pero que sin embargo, se encuentran desprovistas de lagartijas. La interpretación que le damos a este hecho es que el flujo genético resulta importante para el mantenimiento de poblaciones aisladas en esta área, y la ausencia de la especie al oeste de la carretera podría ser un buen ejemplo de cómo el impedimento a la restauración genética puede

desembocar en la extinción local de la especie (Tellería et al. 2011).

Limitaciones de esta tesis y líneas de investigación futuras

Para concretar el grado de aislamiento reproductivo que según los resultados del capítulo 1 podría estar promoviendo la presencia de parásitos serían necesarias varias líneas de estudio: 1) discernir entre divergencia reciente y flujo genético, de manera que si se encontrara una restricción al flujo genético se podría interpretar a la luz de los resultados obtenidos y continuar con los siguientes puntos; 2) para confirmar que la presencia de parásitos afecta a la selección sexual, harían falta experimentos de selección sexual en los que se ofreciera a las hembras de las diferentes poblaciones, machos con distintos aspectos ornamentales desarrollados; 3) aumentar la muestra a más poblaciones que mostraran diferencias en la carga de parásitos, incluyendo distintos tipos de parásitos (por ejemplo, endoparásitos sanguíneos; Álvarez-Ruiz et al. 2018, Carbayo et al. 2018) para poder aumentar la escala de extrapolación de las conclusiones del primer capítulo.

El capítulo 2 es íntegramente una prueba de las limitaciones de los estudios genómicos para dar con la base genética de la divergencia adaptativa. Para intentar definir esa base habría que, sobre todo, ampliar el muestreo a lo largo del genoma (con técnicas de secuenciación masiva más potentes) y realizar ensamblajes más largos para ser capaces de localizar la variabilidad genética encontrada y poder contraponer los picos de divergencia al grado de diferenciación de regiones contiguas.

En el capítulo 3, focalizamos nuestros esfuerzos en una población oriental de fenotipo y ambiente occidentales para intentar desenmarañar las dinámicas genéticas que subyacen a la convergencia adaptativa. Sin embargo, existe todo un clado dentro del linaje oriental que presenta estas mismas características (E3) y que además

aparentemente ha interrumpido el flujo genético con su propio linaje para "retomarlo" con el linaje occidental (Díaz et al. 2017). Sería interesante ampliar la muestra hacia estas poblaciones para investigar si son los mismos mecanismos genéticos los que subyacen a la convergencia adaptativa cuando se mantiene el flujo genético con las poblaciones de las que la muestra diverge y cuando no.

Además, encontramos patrones asimétricos de flujo genético en el capítulo 3 que requieren una explicación más concreta. No sabemos qué determina que la tasa de purga en según qué ambientes sea distinta para genotipos procedentes de distintos sitios. Parece que la divergencia adaptativa no es equivalente, es decir, estar más adaptado a un ambiente que los individuos foráneos hace más o menos adaptable a ambientes nuevos a los individuos, según a qué ambiente se hayan conseguido adaptar localmente. Postulamos una hipótesis en el capítulo 4 a este respecto, en la que describimos cómo los ambientes más singulares para los que menos genotipos están adaptados podrían ser ocupados por diversidad genética que surgiera en contextos de selección más relajada. Otra posibilidad es que los distintos fenotipos, adaptativos localmente, resultaran más o menos ventajosos a la hora de dispersarse. Por ejemplo, en el caso de las lagartijas rayadas (una adaptación para la cripsis) podría suceder que fueran menos detectable en las rutas dispersivas entre poblaciones, haciéndolas más propicias a funcionar como inmigrantes. El sistema de Lerma constituye una buena oportunidad para comprobar esta hipótesis, ya que allí conviven lagartijas rayadas y no rayadas en un ambiente fragmentado por tierras de cultivo (una matriz en la que se esperaría que las lagartijas rayadas fueran más crípticas). Por tanto, con el suficiente poder de resolución (análisis genómicos) se pueden realizar análisis de migración por coalescencia a escala local entre los distintos fragmentos discerniendo entre la contribución al flujo de genes que hace cada uno de los fenotipos.

La precisión de la inferencia del área de distribución del capítulo 4 está completamente sujeta a los datos previos acerca del área real de la especie que se quiere inferir. Se necesitan mejores datos de distribución de esta especie o implementar la metodología en otras especies para las que se conozca el área con mayor precisión. Además, las explicaciones probabilísticas que les damos a las pocas falsas presencias predichas más allá del límite de distribución, requieren una comprobación empírica: los bordes cercanos a esas zonas aptas pero desprovistas de lagartijas deberían estar poblados por individuos no adaptados a la expansión en esa dirección. En general, todo el modelo necesita una validación directa consistente en muestreos representativos de todo el área de distribución de la especie para comprobar que los genotipos que persisten en cada ambiente son aquellos que el modelo predice que están adaptados a ellos.

Por otra parte, todas las explicaciones a la limitación del área que tienen que ver con la capacidad de dispersión de la especie son fácilmente comprobables haciendo experimentos de trasplante más allá del borde o modelizando esta aproximación con diferentes especies con distintas capacidades dispersivas. Las rutas de ocupación de ambientes singulares que proponemos en el capítulo 4, en las que poblaciones que experimentan selección relajada funcionan como fuentes de la poca diversidad genética adaptada a dichos ambientes, necesitan una comprobación con modelos teóricos de dispersión entre sitios con distinta intensidad de selección. Esta sería la manera de concretar si las rutas dispersivas están supeditadas no solo a la existencia de la selección natural (lo cual parece evidente dados los datos de migración del capítulo 3, además de otros estudios al margen de esta tesis) sino también al hecho de que las mismas presiones selectivas presenten distintos grados de intensidad. La explicación más relevante de este capítulo a los límites de distribución de la especie es la que tiene que

ver con los umbrales adaptativos para una cantidad dada de varianza genética sujeta a selección. Esta afirmación necesita también más exploración teórica que revele la relación positiva entre la magnitud del incremento de varianza genética aditiva y la superficie del área a la que la especie es capaz de expandirse. Tiene que demostrarse una relación directa entre el surgimiento o extinción de una mutación aditiva que sume o reste de manera congruente una porción del área de distribución.

Por último, en el capítulo 5, sería deseable una validación de los análisis de estructura genética y patrones de flujo genético entre fragmentos con más cantidad de datos. En este mismo capítulo ya se han validado con una gran base de datos genómica las medidas de homocigosidad con microsatélites, pero hay otros muchos análisis que respaldan las conclusiones del estudio a los que convendría una validación genómica. Esto además hace que se pueda recuperar el uso de marcadores clásicos más accesibles para hacer estudios de genética de poblaciones en los que no se necesita tanta potencia de muestreo genómico para resultar relevante (como a todas luces es el caso de este capítulo). Además, con las grandes bases de datos que ofrecen los escaneos genómicos a gran escala, se pueden inferir con precisión las dinámicas demográficas que se mencionan en este capítulo cuando se invoca al riesgo de extinción como una posible consecuencia de la reducción de aptitud en contextos de poco intercambio genético. Por ejemplo, utilizando aproximaciones de coalescencia para inferir tamaños poblacionales pasados e inferir futuros, se podría intentar relacionar esa depresión por consanguinidad con una tasa de extinción inferida, lo que le daría mucha más consistencia a este trabajo y permitiría concluir una última consecuencia genética de la fragmentación del hábitat.

4. CONCLUSIONES

Capítulo 1

- Parasites seemed to limit the amount of resources available for expressing a color signal of a certain size. In the absence of ticks, the colored surface has enough inter-individual variation to remain informative, and it is honest because it is traded-off with immune response. Parasites led males to allocate to the immune system part of the resources they would otherwise invest in sexual coloration, so that color saturation, rather than colored surface (probably a more demanding trait), could be used to signal male quality when parasites are present.
    - *Los parásitos parecen limitar la cantidad de recursos disponibles para expresar una señal visual de un tamaño concreto. En ausencia de garrapatas, la superficie coloreada tiene la suficiente variación inter-individual como para resultar informativa, y es una señal honesta porque está comprometida con el sistema inmune. Los parásitos provocan que los machos tengan que desviar al sistema inmune parte de los recursos que de otra forma invertirían en ornamentación sexual, por lo que la saturación de color, en vez de la superficie coloreada (que parece más costosa de expresar), podría ser usada como señal de la calidad de los machos en presencia de parásitos.*
- Parasites may constrain the expression of sexual ornaments to the extent of being modulators of the phenotypic effects of sexual selection, including qualitative changes in the traits involved in signaling systems. Thus, parasites like tick nymphs, independently of their effects on individual survival, may have a significant impact on individual fitness by influencing the expression of key sexual ornaments.

- *Los parásitos pueden restringir la expresión de los ornamentos sexuales hasta el punto de resultar moduladores de los efectos fenotípicos de la selección sexual, incluyendo cambios cualitativos en los rasgos involucrados en los sistemas de señalización. Por tanto, los parásitos como las ninfas de garrapata, independientemente de sus efectos en la supervivencia de los individuos, pueden tener un impacto significativo en su aptitud influenciando la expresión de los ornamentos sexuales que modulan su éxito reproductivo.*

Capítulo 2

- We found contrasting patterns of differentiation between the lack of divergence between two populations separated by a 600 m elevational gradient in a genome-wide search for selected loci, and the disparity between them in a number of adaptive traits. Our results suggest that divergent adaptation has evolved so recently in these populations that selection has not had enough time to produce genome-wide marks of divergence that can be traced with our methods of outlier detection.
- *Encontramos una discordancia entre la ausencia de divergencia genómica entre dos poblaciones separadas por un gradiente altitudinal, y una diferenciación fenotípica en rasgos cuyo valor adaptativo parece evidente. Parece que la divergencia entre las poblaciones es tan reciente que no ha dejado todavía huella en el genoma a una escala lo suficientemente grande como para poder ser detectada con los análisis de identificación de* outliers *empleados en este estudio.*
- To be conclusive our results would need the following: 1) more sampling power along the genome to increase the genetic variability under study (if the adaptive divergence were specifically located in non-sampled loci) and 2) inferring longer

sequence assemblies to be able to measure genomic divergence at a local scale (if genetic divergence were heterogeneously distributed along the genome).

- *Para que estos resultados puedan resultar más concluyentes se necesitaría: 1) un aumento de la potencia del muestreo genómico para recuperar más variabilidad genética (si la divergencia adaptativa estuviera muy localizada) y 2) la construcción de ensamblajes más largos para medir el grado de divergencia a escala genómica local (si la divergencia se distribuyera de forma heterogénea a lo largo del genoma).*

Capítulo 3

- Our combined analysis of phenotypic variation, population genetic structure and genetic polymorphisms subject to natural selection, allowed us to explore the sources of genetic variation favored in different ecological contexts in our system. Moreover, these analysis revealed that the combination of divergence subject to selection, balancing selection and permeability to gene flow, may have rendered our model species able to adapt to a wide variety of habitats.
- *Los análisis combinados de variación fenotípica, estructura genética y polimorfismos genéticos sujetos a selección natural, nos han permitido explorar las distintas fuentes de variación genética que se ven favorecidas en diferentes contextos ecológicos. Además, nuestros análisis revelaron que la combinación de la divergencia sujeta a selección, la selección equilibradora, y la permeabilidad al flujo genético entre poblaciones localmente adaptadas, podrían haber conducido a que nuestra especie modelo sea capaz de estar adaptada a una gran variedad de hábitats.*

- The general patterns of ecologically-driven divergence that were found in our system, could be blurred by a dynamic landscape configuration by means of: 1)

asymmetric patterns of gene flow (permeability to immigration conditioned by a dynamic adaptive landscape), 2) convergent genetic innovation, and 3) balancing selection of ancestral genetic variability (prior to divergence).

- *Los patrones generales de divergencia ecológica encontrados en nuestro sistema, se ven desdibujados en contextos de inestabilidad ambiental, gracias a: 1) patrones asimétricos de flujo genético (permeabilidad a la inmigración condicionada por una configuración dinámica del paisaje adaptativo), 2) la innovación genética convergente y 3) el mantenimiento de variación genética ancestral (previa a la divergencia).*

Capítulo 4

- Extrapolating genotype-environment association analyses to all possible combinations of alleles at a few loci subject to selection allowed us to explore how much environmental variation could be exploited by our model species given its dispersal limitations. Our results revealed that inferring a species' range from the geographical distribution of SNPs under selection can be not only very accurate, but also truly informative about the genetic dynamics that underlie local adaptation all over a species' range.
- *La extrapolación de los análisis de asociación genotipo-ambiente a todas las combinaciones posibles de alelos en unos pocos loci sujetos a selección, nos permiten explorar cuánta variabilidad ambiental es capaz de explotar nuestra especie modelo, dadas sus limitaciones a la hora de dispersarse. Esta aproximación reveló que la inferencia del área de distribución de una especie a partir de la distribución geográfica de variantes genéticas sujetas a selección puede ser no solo muy precisa, sino también verdaderamente informativa acerca de las dinámicas genéticas que subyacen a la adaptación local a escala global*

(es decir, a escala del área de distribución total).

- We revealed an adaptability threshold defined by the amount of genetic variability subject to selection as the ultimate determinant of the location and shape of range boundaries in our system. Thus, our successful inference of a species' range from the geographical distribution of a few adaptive loci uncovered the role of individual genotypic uniqueness in shaping a species' distribution range.
  - *Encontramos un umbral adaptativo definido por la cantidad de variabilidad genética sujeta a selección como determinante último de la configuración de los límites de distribución de nuestra especie modelo. El éxito a la hora de inferir el área de distribución a partir de la localización geográfica de variantes sujetas a selección, revela el papel de la singularidad genética de los individuos de una especie en la configuración de su área de distribución.*

Capítulo 5

- We have provided empirical evidence of the links among reduced gene flow, increased individual homozygosity, and reduced fitness of individuals coming from anthropogenically fragmented populations.
  - *Se han obtenido pruebas empíricas de la relación entre la reducción del flujo genético, el aumento de la homocigosidad individual y la pérdida de aptitud en individuos provenientes de poblaciones fragmentadas por efecto de la acción antrópica*.
- Our results support the view that negative genetic effects that happen after habitat fragmentation may reduce population viability before population size becomes too small. This is sustained in our line of evidence which uncovered that individual fitness may lie behind the reduced persistence of species in

fragmented landscapes.

- *Nuestros resultados apoyan la idea de que los efectos genéticos negativos que suceden a la fragmentación de hábitat, pueden reducir la viabilidad de las poblaciones antes de que el tamaño de éstas sea demasiado pequeño. Esto se sostiene en la línea de evidencia que pone de manifiesto que la aptitud individual podría resultar un determinante muy importante para la persistencia de las poblaciones en ambientes fragmentados.*

BIBLIOGRAFÍA

Aguillon, S. M., Campagna, L., Harrison, R. G., & Lovette, I. J. (2018) A flicker of hope: Genomic data distinguish Northern Flicker taxa despite low levels of divergence. *The Auk*, **135**, 748-766.

Aguirre-Liguori, J. A., Tenaillon, M. I., Vázquez-Lobo, A., Gaut, B. S., Jaramillo-Correa, J. P., Montes-Hernandez, S., Eguiarte, L. E. (2017) Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular Ecology*, **26**, 4226-4240.

Alexander, D. H., Novembre, J., & Lange, K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**, 1655-1664.

Alleaume-Benharira, M., Pen, I. R., & Ronce, O. (2006) Geographical patterns of adaptation within a species' range: Interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203-215.

Álvarez-Ruiz, L., Megía-Palma, R., Reguera, S., Ruiz, S., Zamora-Camacho, F. J., Figuerola, J., & Moreno-Rueda, G. (2018) Opposed elevational variation in prevalence and intensity of endoparasites and their vectors in a lizard. *Current Zoology*, **64**, 197-204.

Anderson, J. T., Perera, N., Chowdhury, B., & Mitchell-Olds, T. (2015) Microgeographic Patterns of Genetic Divergence and Adaptation across Environmental Gradients in *Boechera stricta* (Brassicaceae). *The American Naturalist*, **1**, 60-73.

Angert, A. L., Bradshaw, H. D., & Schemske, D. W. (2008) Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution*, **62**, 2660-75.

Aparicio, J. M., Ortego, J., & Cordero, P. J. (2006) What should we weigh to estimate heterozygosity, alleles or loci? *Molecular Ecology*, **15**, 4659-4665.

Barrett, R. D. H., & Hoekstra, H. E. (2011) Molecular spandrels: Tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767-780.

Barrett, R. D. H., & Schluter, D. (2008) Adaptation from standing genetic variation.

Trends in Ecology and Evolution, **23**, 38-44.

Barton, N. H., & Etheridge, A. M. (2018) Establishment in a new habitat by polygenic adaptation. *Theoretical Population Biology*, **122**, 110-127.

Beerli, P., & Felsenstein, J. (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763-773

Beerli, P., & Felsenstein, J. (2001). Maximum-likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the USA*, **98**, 4563-4568.

Belliure, J., Smith, L., & Sorci, G. (2004) Effect of testosterone on T cell-mediated immunity in two species of Mediterranean lacertid lizards. *Journal of Experimental Zoology Part A: Comparative Experimental Biology*, **301**, 411-8.

Berg, J. J., & Coop, G. (2014) A Population Genetic Signal of Polygenic Adaptation. *PLoS Genetics*, **10**, e1004412.

Berlocher, S. H., Roelofs, W., Feder, J. L., Dambroski, H. R., Linn, C., & Nojima, S. (2003) Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences*, **100**, 11490-11493.

Bloor, P., & Dávila, J. A. (2008) Polymorphic microsatellite markers for the lizard *Psammodromus algirus* (Squamata: Lacertidae). *Molecular Ecology Resources*, **8**, 631-633.

Blount, J. D., Metcalfe, N. B., Birkhead, T. R., & Surai, P. F. (2003) Carotenoid modulation of immune function and sexual attractiveness in zebra finches. *Science*, **300**,125-7.

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., & SanCristobal, M. (2010) Detecting selection in population trees: The Lewontin and Krakauer test extended. *Genetics*, **186**, 241-262.

Bons, J. & Geniez, P. (1996) Anfibios y Reptiles de Marruecos (incluido el Sahara

Occidental): Atlas biogeográfico – Amphibiens et Reptiles du Maroc (Sahara Occidental Compris). *Barcelona: Asociación Herpetológica española*.

Brace, A. J., Sheikali, S., & Martin, L. B. (2015) Highway to the danger zone: Exposure-dependent costs of immunity in a vertebrate ectotherm. *Functional Ecology*, **29**, 924-930.

Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2013) Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution*, **67**, 3258-3273.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, R. M., Ramdoss, Y., & Buckler, E. S. (2007) TASSELL Software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633-2635.

Bridle, J. R., & Vines, T. H. (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution*, **22**, 140-147.

Brown, J. H. (1984) On the Relationship between Abundance and Distribution of Species. *The American Naturalist*, **124**, 255-279.

Buckley, L. B. (2010) The range implications of lizard traits in changing environments. *Global Ecology and Biogeography*, **19**, 452-464.

Burbrink, F. T. (2010) Historical versus contemporary migration in fragmented populations. *Molecular Ecology*, **19**, 5321-5323.

Campagna, L., Gronau, I., Silveira, L. F., Siepel, A., & Lovette, I. J. (2015) Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Molecular Ecology*, **24**, 4238-4251.

Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018) What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, **124**, 561-583.

Carbayo, J., Martín, J., & Civantos, E. (2018) Habitat type influences parasite load in Algerian *Psammodromus* (*Psammodromus algirus*) lizards. *Canadian Journal of Zoology*, **97**, 172-180.

Caro, T. M., & Laurenson, K. (1994) Ecological and Genetic Factors in Conservation :

A Cautionary Tale. *Science*, **263**, 485-486.

Carranza, S., Harris, D. J., Arnold, E. N., Batista, V., & Gonzalez De La Vega, J. P. (2006) Phylogeography of the lacertid lizard, *Psammodromus algirus*, in Iberia and across the Strait of Gibraltar. *Journal of Biogeography*, **33**, 1279-1288.

Casagrande, S., Pinxten, R., Zaid, E., & Eens, M. (2015) Birds receiving extra carotenoids keep singing during the sickness phase induced by inflammation. *Behavioral Ecology and Sociobiology*, **69**, 1029–1037.

Case, T. J., & Taper, M. L. (2000) Interspecific Competition, Environmental Gradients, Gene Flow, and the Coevolution of Species' Borders. *The American Naturalist*, **155**, 583-605.

Casher, L., Lane, R., Barrett, R., & Eisen, L. (2002) Relative importance of lizards and mammals as hosts for ixodid ticks in northern California. *Experimental and Applied Acarology*, **26**, 127-143.

Chapman, J. R., Nakagawa, S., Coltman, D. W., Slate, J., & Sheldon, B. C. (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. *Molecular Ecology*, **18**, 2746-2765.

Charlesworth, B. (2009) Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, **10**, 195-205.

Chemnitz, J., Jentschke, P. C., Ayasse, M., & Steiger, S. (2015) Beyond species recognition: Somatic state affects long-distance sex pheromone communication. *Proceedings of the Royal Society B: Biological Sciences*, **282**, 20150832.

Cheptou, P. O., Hargreaves, A. L., Bonte, D., & Jacquemyn, H. (2017) Adaptation to fragmentation: evolutionary dynamics driven by human influences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160037.

Chiu, K. W., & Maderson, P. F. A. (1975) The microscopic anatomy of epidermal glands in two species of gekkonine lizards, with some observations on testicular activity. *Journal of Morphology*, **147**, 23-39.

Clausen, J., Keck, D., & Hiesey, W. (1948) Experimental studies on the nature of species. III. Environmental responses of climatic races of *Achillea*. *Carnegie Institution of Washington*.

Cole, C. J. (1966) Femoral glands of the lizard, *Crotaphytus collaris*. *Journal of Morphology*, **118**, 119-136.

Coltman, D. W., Pilkington, J. G., Smith, J. A., & Pemberton, J. M. (1999) Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution*, **53**, 1259-1267.

Coltman, D. W., & Slate, J. (2003) Microsatellite measures of inbreeding: a meta-analysis. *Evolution*, **57**, 971-983.

Connallon, T., & Sgrò, C. M. (2018) In search of a general theory of species' range evolution. *PLoS Biology*, **16**, e2006735.

Consuegra, S., Verspoor, E., Knox, D., & García De Leániz, C. (2005) Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. *Conservation Genetics*, **6**, 823-842.

Crnokrak, P., & Roff, D. A. (1999) Inbreeding depression in the wild. *Heredity*, **83**, 260-270.

Dawson, M. N., Grosberg, R. K., Stuart, Y. E., & Sanford, E. (2010) Population genetic analysis of a recent range expansion: Mechanisms regulating the poleward range limit in the volcano barnacle *Tetraclita rubescens*. *Molecular Ecology*, **19**, 1585-1605.

De Kort, H., Vandepitte, K., Bruun, H. H., Closset-Kopp, D., Honnay, O., & Mergeay, J. (2014) Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Molecular Ecology*, **23**, 4709-4721.

Delhey, K., & Peters, A. (2008) Quantifying variability of avian colours: Are signalling traits more variable? *PLoS ONE*, **3**, e1689.

Delph, L. F. (2018) The Study of Local Adaptation: A Thriving Field of Research. *Journal of Heredity,* **109**, 1-2.

Diaz, J. A. (1993) Breeding coloration, mating opportunities, activity, and survival in the lacertid lizard *Psammodromus algirus*. *Canadian Journal of Zoology*, **71**, 1104-1110.

Díaz, J. A. (1997) Ecological correlates of the thermal quality of an ectotherm's habitat: A comparison between two temperate lizard populations. *Functional Ecology*, **11**, 79-89.

Diaz, J. A., Alonzo-Gomez, A. L., & Delgado, M. J. (1994) Seasonal variation of gonadal development, sexual steroids, and lipid reserves in a population of the lizard *Psammodromus algirus*. *Journal of Herpetology*, **28**, 199-205.

Díaz, J. A., Carbonell, R., Virgós, E., Santos, T., & Tellería, J. L. (2000) Effects of forest fragmentation on the distribution of the lizard *Psammodromus algirus*. *Animal Conservation*, **3**, 235-240.

Díaz, J. A., Pérez-Tris, J., Bauwens, D., Pérez-Aranda, D., Carbonell, R., Santos, T., & Tellería, J. L. (2007) Reproductive performance of a lacertid lizard at the core and the periphery of the species' range. *Biological Journal of the Linnean Society*, **92**, 87-96.

Díaz, J. A., Pérez-Tris, J., Tellería, J. L., Carbonell, R., & Santos, T. (2005) Reproductive investment of a lacertid lizard in fragmented habitat. *Conservation Biology*, **19**, 1578-1585.

Díaz, J. A., Verdú-Ricoy, J., Iraeta, P., Llanos-Garrido, A., Pérez-Rodríguez, A., & Salvador, A. (2017) There is more to the picture than meets the eye: adaptation for crypsis blurs phylogeographical structure in a lizard. *Journal of Biogeography*, **44**, 397-408.

Dixo, M., Metzger, J. P., Morgante, J. S., & Zamudio, K. R. (2009) Habitat fragmentation reduces genetic diversity and connectivity among toad populations in the Brazilian Atlantic Coastal Forest. *Biological Conservation*, **142**, 1560-1569.

Duckworth, R. A. (2008) Adaptive Dispersal Strategies and the Dynamics of a Range Expansion. *The American Naturalist*, **1**, S4-S17.

Dudaniec, R. Y., Yong, C. J., Lancaster, L. T., Svensson, E. I., & Hansson, B. (2018) Signatures of local adaptation along environmental gradients in a range-expanding damselfly (*Ischnura elegans*). *Molecular Ecology*, **27**, 2576-2593.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008) Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170-1188.

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., Wolf, J. B. W. (2012) The genomic landscape of species divergence in Ficedula flycatchers. *Nature*, **491**, 756-60.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Endler, J. A. (1977) Geographic variation, speciation, and clines. *Monographs in Population Biology*, **10**, 1-246.

Endler, J. A., & Houde, A. E. (1995) Geographic Variation in Female Preferences for Male Traits in *Poecilia reticulata*. *Evolution*, **49**, 456-468.

Excoffier, L., & Lischer, H. E. L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.

Fahrig, L. (2003) Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 487-515.

Feulner, P. G. D., Chain, F. J. J., Panchal, M., Huang, Y., Eizaguirre, C., Kalbe, M., Milinski, M. (2015) Genomics of Divergence along a Continuum of Parapatric Population Differentiation. *PLoS Genetics*, **11**, e1005414.

Fitze, P. S., Gonzalez-Jimena, V., San-Jose, L. M., San Mauro, D., Aragón, P., Suarez, T., & Zardoya, R. (2011) Integrative analyses of speciation and divergence in *Psammodromus hispanicu*s (Squamata: Lacertidae). *BMC Evolutionary Biology*, **11**, 347.

Foley, J. A., Defries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Snyder,

P. K. (2005) Global Consequences of Land Use, **8**, 570-574.

Foll, M., & Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, **180**, 977-993.

Folstad, I., & Karter, A. J. (2002) Parasites, Bright Males, and the Immunocompetence Handicap. *The American Naturalist*, **3**, 603-622.

Forstmeier, W., Schielzeth, H., Mueller, J. C., Ellegren, H., & Kempenaers, B. (2012) Heterozygosity-fitness correlations in zebra finches: Microsatellite markers can be better than their reputation. *Molecular Ecology*, **21**, 3237-3249.

Frankham, R. (2005) Genetics and extinction. *Biological Conservation*, **126**, 131-140.

Frankham, R. (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919-1927.

Frankham, R. (2010) Inbreeding in the wild really does matter. *Heredity*, **104**, 124.

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admettla, A., Pattini, L., & Nielsen, R. (2011) Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics*, **7**, e1002355.

Gallet, R., Froissart, R., & Ravigné, V. (2018) Experimental demonstration of the impact of hard and soft selection regimes on polymorphism maintenance in spatially heterogeneous environments. *Evolution*, **72**, 1677-1688.

Giery, S. T., & Layman, C. A. (2015) Interpopulation Variation in a Condition-Dependent Signal: Predation Regime Affects Signal Intensity and Reliability. *The American Naturalist*, **186**, 187-95.

Groot, M. P., Wagemaker, N., Ouborg, N. J., Verhoeven, K. J. F., & Vergeer, P. (2018) Epigenetic population differentiation in field- and common garden-grown *Scabiosa columbaria* plants. *Ecology and Evolution*, **8**, 3505-3517.

Halbritter, A. H., Billeter, R., Edwards, P. J., & Alexander, J. M. (2015) Local adaptation at range edges: Comparing elevation and latitudinal gradients. *Journal of Evolutionary Biology*, **28**, 1849-1860.

Haldane, J. B. S. (1956) The Relation between Density Regulation and Natural Selection. *Proceedings of the Royal Society B: Biological Sciences*, **145**, 306-308.

Hamilton, W. D., & Zuk, M. (1982) Heritable true fitness and bright birds: A role for parasites? *Science*, 218, 384-387.

Han, F., Lamichhaney, S., Rosemary Grant, B., Grant, P. R., Andersson, L., & Webster, M. T. (2017) Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Research*, **27**, 1004-1015.

Hanski, I. (1998) Metapopulation dynamics. *Nature*, **396**, 41-49.

Hanski, I., & Saccheri, I. (2006) Molecular-level variation affects population growth in a butterfly metapopulation. *PLoS Biology*, **4**, 719-726.

Hansson, B., & Westerberg, L. (2002) On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology*, **11**, 2467-2474.

Hardie, D. C., & Hutchings, J. A. (2010) Evolutionary ecology at the extremes of species' ranges. *Environmental Reviews*, **18**, 1-20.

Hargreaves, A. L., Samis, K. E., & Eckert, C. G. (2014) Are Species' Range Limits Simply Niche Limits Writ Large? A Review of Transplant Experiments beyond the Range. *The American Naturalist*, **183**, 157-173.

Harris, R. B., Alström, P., Ödeen, A., & Leaché, A. D. (2018) Discordance between genomic divergence and phenotypic variation in a rapidly evolving avian genus (*Motacilla*). *Molecular Phylogenetics and Evolution*, **120**, 183-195.

Hedrick, P. W., & Garcia-Dorado, A. (2016) Understanding Inbreeding Depression, Purging, and Genetic Rescue. *Trends in Ecology and Evolution*, **31**, 940-952.

Hereford, J. (2009) A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The American Naturalist*, **173**, 579-588.

Herrera, C. M., & Bazaga, P. (2008) Adding a third dimension to the edge of a species' range: Altitude and genetic structuring in mountainous landscapes. *Heredity*, **100**, 275-285.

Hill, G. E. (1991) Plumage coloration is a sexually selected indicator of male quality. *Nature*, **350**, 337-339.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Whitlock, M. C. (2016) Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions (Supplemental Material). *The American Naturalist*, **188**, 379-397.

Hoffman, J. I., Forcada, J., Trathan, P. N., & Amos, W. (2007) Female fur seals show active choice for males that are heterozygous and unrelated. *Nature*, **445**, 912-914.

Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A. S., Dasmahapatra, K. K. (2014) High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences*, **111**, 3775-3780.

Holt, R. D. (2003) On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, **5**, 159-178.

Holt, R. D., & Keitt, T. H. (2005) Species' borders: A unifying theme in ecology. *Oikos*, **108**, 3-6.

Huisman, J., Kruuk, L. E. B., Ellis, P. A., Clutton-Brock, T., & Pemberton, J. M. (2016) Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, **113**, 3585-3590.

Hutchinson, G. E. (1957) Concluding remarks on the cold Spring Harbor Symposia on Quantitative Biology. *Population Studies: Animal Ecology and Demography. Cold Spring Harbor SYmposia on Quantitative Biology*.

Iraeta, P., Monasterio, C., Salvador, A., & Díaz, J. A. (2006) Mediterranean hatchling lizards grow faster at higher altitude: A reciprocal transplant experiment. *Functional Ecology*, **20**, 865-872.

Iraeta, P., Monasterio, C., Salvador, A., & Díaz, J. A. (2011) Sexual dimorphism and interpopulation differences in lizard hind limb length: Locomotor performance or chemical signalling? *Biological Journal of the Linnean Society*, **104**, 318-329.

Iraeta, P., Salvador, A., & Díaz, J. A. (2008) A reciprocal transplant study of activity, body size, and winter survivorship in juvenile lizards from two sites at different altitude. *Écoscience*, **15**, 298-304.

Iraeta, P., Salvador, A., & Díaz, J. A. (2013) Life-history traits of two Mediterranean lizard populations: A possible example of countergradient covariation. *Oecologia*, **172**, 167-176.

Iraeta, P., Salvador, A., Monasterio, C., & Díaz, J. A. (2010) Effects of gravidity on the locomotor performance and escape behaviour of two lizard populations: The importance of habitat structure. *Behaviour*, **147**, 133-150.

Iwasa, Y., Pomiankowski, A., & Nee, S. (1991) The Evolution of Costly Mate Preferences II. The "Handicap" Principle. *Evolution*, **45**, 1431-1442.

Jacot, A., Scheuber, H., Kurtz, J., & Brinkhof, M. W. G. (2005) Juvenile immune status affects the expression of a sexually selected trait in field crickets. *Journal of Evolutionary Biology*, **18**, 1060-1068.

Johansson, M., Primmer, C. R., & Merilä, J. (2007) Does habitat fragmentation reduce fitness and adaptability? A case study of the common frog (*Rana temporaria*). *Molecular Ecology*, **16**, 2693-2700.

Joo, J. W. J., Kang, E. Y., Org, E., Furlotte, N., Parks, B., Hormozdiari, F., Eskin, E. (2016) Efficient and accurate multiple-phenotype regression method for high dimensional data considering population structure. *Genetics*, **204**, 1379-1390.

Kalinowski, S. T., Wagner, A. P., & Taper, M. L. (2006) ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, **6**, 576–579

Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016) Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications,* **9**, 1205–1218.

Kawecki, T. J. (2008) Adaptation to Marginal Habitats. *Annual Review of Ecology, Evolution, and Systematics.* **39**, 321-342.

Kearney, M., & Porter, W. (2009) Mechanistic niche modelling: Combining

physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334-350

Kekäläinen, J., Pirhonen, J., & Taskinen, J. (2014) Do highly ornamented and less parasitized males have high quality sperm? - an experimental test for parasite-induced reproductive trade-offs in European minnow (*Phoxinus phoxinus*). *Ecology and Evolution,* **22**, 4237-4246.

Keller, L.F., & DeeM, W. (2002) Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, **17**, 230–241.

Kim, Y., & Stephan, W. (2002) Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*, **160**, 765-77.

Kirkpatrick, M., & Barton, N. H. (1997) Evolution of a Species' Range. *The American Naturalist,* **150**, 1-23.

Kokko, H., Chaturvedi, A., Croll, D., Fischer, M. C., Guillaume, F., Karrenberg, S., Stapley, J. (2017) Can Evolution Supply What Ecology Demands? *Trends in Ecology and Evolution,* **32**, 187-197

Kolbe, J. J., Leal, M., Schoener, T. W., Spiller, D. A., & Losos, J. B. (2012) Founder effects persist despite adaptive differentiation: A field experiment with lizards. *Science*, **335**, 1086-1089.

Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., Foster, S. A. (2009) Relaxed selection in the wild. *Trends in Ecology and Evolution*, **24**, 487-496.

Lande, R. (1988) Genetics and Demography in Biological Conservation. *Science*, **241**, 1455–1460.

Lande, R., & Shannon, S. (1996) The Role of Genetic Variation in Adaptation and Population Persistence in a Changing Environment. *Evolution*, **50**, 434-437.

LaRue, E. A., Holland, J. D., & Emery, N. C. (2018) Environmental predictors of dispersal traits across a species' geographic range. *Ecology*, **99**, 1857-1865.

Laurent, S., Pfeifer, S. P., Settles, M. L., Hunter, S. S., Hardwick, K. M., Ormond, L., Rosenblum, E. B. (2016) The population genomics of rapid adaptation:

Disentangling signatures of selection and demography in white sands lizards. *Molecular Ecology*, **25**, 306-323.

Lawson, L. P., & Petren, K. (2017) The adaptive genomic landscape of beak morphology in Darwin's finches. *Molecular Ecology*, **26**, 4978-4989.

Lee-Yaw, J. A., Fracassetti, M., & Willi, Y. (2018) Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. *Ecography*, **41**, 622-634.

Lee, B. H., Bae, S.-W., Shim, J. J., Park, S. Y., & Park, and H. Y. (2016) Imaging Single-mRNA Localization and Translation in Live Neurons. *Molecules and Cells*, **39**, 841-846.

Levinst, R. (1962) Theory of Fitness in a Heterogeneous Environment . I. The Fitness Set and Adaptive Function. *The American Naturalist*, **96**, 361-373.

Lewontin, R. C., & Krakauer, J. (1975) Testing the heterogeneity of F values. *Genetics*, **80**, 397-398.

Llanos-Garrido, A., Díaz, J. A., Pérez-Rodríguez, A., & Arriero, E. (2017) Variation in male ornaments in two lizard populations with contrasting parasite loads. *Journal of Zoology*, **303**, 218-225.

López, P., Gabirot, M., & Martín, J. (2009) Immune challenge affects sexual coloration of male iberian wall lizards. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, **331A**, 96-104.

Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R., Cherney, J. H., Casler, M. D., Costich, D. E. (2013) Switchgrass Genomic Diversity, Ploidy, and Evolution: Novel Insights from a Network-Based SNP Discovery Protocol. *PLoS Genetics*, **9**, e1003215.

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981-994.

Madsen, T., Shine, R., Olsson, M., & Wittzell, H. (1999) Restoration of an inbred adder population. *Nature*, **402**, 34–35.

Maisano Delser, P., Corrigan, S., Duckett, D., Suwalski, A., Veuille, M., Planes, S.,

Mona, S. (2018) Demographic inferences after a range expansion can be biased: the test case of the blacktip reef shark (*Carcharhinus melanopterus*). *Heredity*, **18**, 36:42.

Malo, A. F., Roldan, E. R. S., Garde, J. J., Soler, A. J., Vicente, J., Gortazar, C., & Gomendio, M. (2008) What does testosterone do for red deer males? *Proceedings of the Royal Society B: Biological Sciences*, **276**.

Martín, J., & Forsman, A. (1999) Social costs and development of nuptial coloration in male *Psammodromus algirus* lizards: an experiment. *Behavioral Ecology*, **10**, 396-400.

Martín, J., & López, P. (2014) Condition-dependent chemosignals in reproductive behavior of lizards. *Hormones and Behavior*, **68**, 14-24.

McEntee, J. P., Tobias, J. A., Sheard, C., & Burleigh, J. G. (2018) Tempo and timing of ecological trait divergence in bird speciation. *Nature Ecology and Evolution*, **2**, 1120-1127.

McGirr, J. A., & Martin, C. H. (2018) Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. *Evolution Letters*, **2**, 62-75.

Milinski, M., & Bakker, T. C. M. (1990) Female sticklebacks use male coloration in mate choice and hence avoid parasitized males. *Nature*, **344**, 330-333

Møller, A. P. (1988) Female choice selects for male sexual tail ornaments in the monogamous swallow. *Nature*, **332**, 640-642.

Moller, A. P., & de Lope, F. (1994) Differential Costs of a Secondary Sexual Character: An Experimental Test of the Handicap Principle. *Evolution*, **48**, 1676-1683.

Moody, K. N., Hunter, S. N., Childress, M. J., Blob, R. W., Schoenfuss, H. L., Blum, M. J., & Ptacek, M. B. (2015) Local adaptation despite high gene flow in the waterfall-climbing Hawaiian goby, Sicyopterus stimpsoni. *Molecular Ecology*, **24**, 545-563.

Nathan, L. R., Kanno, Y., & Vokoun, J. C. (2017) Population demographics influence genetic responses to fragmentation: A demogenetic assessment of the ‘one

migrant per generation' rule of thumb. *Biological Conservation*, **210**, 261–272.

Nosil, P., & Crespi, B. J. (2004) Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. *Evolution*, **58**, 102-112.

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375-402.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013) Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983-5999.

Palmer, D. H., & Kronforst, M. R. (2015) Divergence and gene flow among Darwin's finches: A genome-wide view of adaptive radiation driven by interspecies allele sharing. *BioEssays*, **37**, 968-974.

Parmentier, H. K., De Vries Reilingh, G., & Nieuwland, M. G. B. (1998) Kinetic and immunohistochemical characteristics of mitogen-induced cutaneous hypersensitivity in chickens selected for antibody responsiveness. *Veterinary Immunology and Immunopathology*, **66**, 367-376.

Pavlidis, P., Jensen, J. D., Stephan, W., & Stamatakis, A. (2012) A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution*, **29**, 3237-3248.

Peterson, A. T. (2011) Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography*, **38**, 817-827.

Pfenninger, M., & Posada, D. (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): Fragmentation, corridor migration, and secondary contact. *Evolution*, **56**, 1776-1788.

Phillips, B. L., Brown, G. P., Travis, J. M. J., & Shine, R. (2008) Reid's Paradox Revisited: The Evolution of Dispersal Kernels during Range Expansion. *The American Naturalist*, **172**, 534-548.

Phillips, P. C. (1996) Maintenance of polygenic variation via a migration-selection

balance under uniform selection. *Evolution*, **50**, 1334–1339.

Phillipsen, I. C., Kirk, E. H., Bogan, M. T., Mims, M. C., Olden, J. D., & Lytle, D. A. (2015) Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology*, **24**, 54-69.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Sexton, J. O. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**, 1246752.

Pimm, S. L., & Raven, P. (2000) Extinction by numbers. Nature, **403**, 843–845.

Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B., & Thompson, J. D. (2017) Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews*, **92**, 1877-1909.

Pleguezuelos, J. M. (1997) Distribución y Biogeografía de los Anfibios y Reptiles en España y Portugal. *Monográfica Tierras del Sur. Universidad de Granada*, **19**, 535.

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

Polechová, J. (2018) Is the sky the limit? On the expansion threshold of a species' range. *PLoS Biology*, **16**, e2005372.

Polechová, J., & Barton, N. H. (2015) Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences*, **112**, 6401-6406.

Polechová, J., Barton, N., & Marion, G. (2009) Species' Range: Adaptation in Space and Time. *The American Naturalist*, **174**, e186-e204.

Pollock, N. B., Vredevoe, L. K., & Taylor, E. N. (2012) How do host sex and reproductive state affect host preference and feeding duration of ticks? *Parasitology Research*, **111**, 897-907.

Przeworski, M. (2002) The signature of positive selection at randomly chosen loci. *Genetics*, **59**, 2312-2323.

Pujol, B., Zhou, S.-R., Sanchez Vilas, J., & Pannell, J. R. (2009) Reduced inbreeding depression after species range expansion. *Proceedings of the National Academy of Sciences*, **106**, 15379-15383.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Sham, P. C. (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, **83**, 559-575.

Putman, A. I., & Carbone, I. (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution*, **4**, 4399-4428.

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

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Westram, A. M. (2017) Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, **30**, 1450-1477.

Reed, D. H., & Frankham, R. (2003) Correlation between Fitness and Genetic Diversity. *Conservation Biology*, **17**, 230–237.

Reed, D. H., Teoh, V. H., Stratton, G. E., & Hataway, R. A. (2011) Levels of gene flow among populations of a wolf spider in a recently fragmented habitat: Current versus historical rates. *Conservation Genetics*, **12**, 331–335.

Reedy, A. M., Cox, C. L., Chung, A. K., Evans, W. J., & Cox, R. M. (2016) Both sexes suffer increased parasitism and reduced energy storage as costs of reproduction in the brown anole, *Anolis sagrei*. *Biological Journal of the Linnean Society*, **117**, 516-527.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348-4370.

Rhode, C., Bester-Van Der Merwe, A. E., & Roodt-Wilding, R. (2017) An assessment

of spatio-temporal genetic variation in the South African abalone (Haliotis midae), using SNPs: implications for conservation management. *Conservation Genetics*, **18**, 17-31.

Rice, W. R., & Salt, G. W. (1990) The Evolution of Reproductive Isolation as a Correlated Character Under Sympatric Conditions: Experimental Evidence. *Evolution*, **44**, 1140-1152.

Roff, D. A. (2007) Contributions of genomics to life-history theory. *Nature Reviews Genetics*, **8**, 116-125.

Rosenblum, E. B. (2006) Convergent Evolution and Divergent Selection: Lizards at the White Sands Ecotone. *The American Naturalist*, **167**, 1-15.

Rosenblum, E. B., Hickerson, M. J., & Moritz, C. (2007) A multilocus perspective on colonization accompanied by selection and gene flow. *Evolution*, **61**, 2971-2985.

Rosenblum, E. B., Rompler, H., Schoneberg, T., & Hoekstra, H. E. (2010) Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings of the National Academy of Sciences*, **107**, 2113-2117.

Rousset, F. (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.

Rundle, H. D., & Nosil, P. (2005) Ecological speciation. Ecology Letters, **8**, 336-352.

Safran, R. J., Scordato, E. S. C., Wilkins, M. R., Hubbard, J. K., Jenkins, B. R., Albrecht, T., Kane, N. C. (2016) Genome-wide differentiation in closely related populations: the roles of selection and geographic isolation. *Molecular Ecology*, **25**, 3865-3883.

Sagarin, R. D., Gaines, S. D., & Gaylord, B. (2006) Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology and Evolution*, **21**, 524-530.

Salvador, A., Veiga, J. P., Martin, J., Lopez, P., Abelenda, M., & Puerta, M. (1996) The cost of producing a sexual signal: Testosterone increases the susceptibility of male lizards to ectoparasitic infestation. *Behavioral Ecology*, **7**, 145-150.

Sanford, E., Holzman, S. B., Haney, R. A., Rand, D. M., & Bertness, M. D. (2006)

Larval tolerance, gene flow, and the northern geographic range limit of fiddler crabs. *Ecology*, **87**, 2882-2894.

Santos, T., Díaz, J. A., Pérez-tris, J., Carbonell, R., & Tellería, J. L. (2008) Habitat quality predicts the distribution of a lizard in fragmented woodlands better than habitat fragmentation. *Animal Conservation*, **11**, 46–56.

Santos, T., Pérez-Tris, J., Carbonell, R., Tellería, J. L., & Díaz, J. A. (2009) Monitoring the performance of wild-born and introduced lizards in a fragmented landscape: Implications for ex situ conservation programmes. *Biological Conservation*, **142**, 2923–2930.

Savolainen, O., Lascoux, M., & Merilä, J. (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807-820.

Schield, D. R., Adams, R. H., Card, D. C., Perry, B. W., Pasquesi, G. M., Jezkova, T., Castoe, T. A. (2017) Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. *Ecology and Evolution*, **7**, 3951-3966.

Schlötterer, C. (2003) Hitchhiking mapping - Functional genomics from the population genetics perspective. *Trends in Genetics*, **19**, 32-38.

Schluter, D. (2000) Ecology and the origin of species. *Trends in Ecology and Evolution*, **16**, 372-380.

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014) Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, **68**, 1-15.

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009) Evolution and Ecology of Species Range Limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415-436.

Sexton, J. P., Strauss, S. Y., & Rice, K. J. (2011) Gene flow increases fitness at the warm edge of a species' range. *Proceedings of the National Academy of Sciences*, **108**, 11704-11709.

Shaner, P. J. L., Tsao, T. H., Lin, R. C., Liang, W., Yeh, C. F., Yang, X. J., Li, S. H.

(2015) Climate niche differentiation between two passerines despite ongoing gene flow. *Journal of Animal Ecology*, **84**, 829-839.

Simmons, A. D., & Thomas, C. D. (2004) Changes in Dispersal during Species' Range Expansions. *The American Naturalist*, **164**, 378-395.

Sinervo, B., & Licht, P. (1991) Proximate constraints on the evolution of egg size, number, and total clutch mass in lizards. *Science*, **252**, 1300–1302.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Smith, J. M. (1991) Theories of sexual selection. *Trends in Ecology and Evolution*, **6**, 146-151.

Spielman, D., Brook, B. W., & Frankham, R. (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences*, **101**, 15261–15264.

StaSoft Inc. (2016) StatSoft.

Stevens, M., Párraga, C. A., Cuthill, I. C., Partridge, J. C., & Troscianko, T. S. (2007) Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, **90**, 211-237.

Svardal, H., Rueffler, C., & Hermisson, J. (2015) A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. *Theoretical Population Biology*, **99**, 76-97.

Szulkin, M., Bierne, N., & David, P. (2010) Heterozygosity-fitness correlations: a time for reappraisal. *Evolution*, **64**, 1202–1217.

Telleria, J. L., Díaz, J. A., Perez-Tris, J., de Juana, E., de la Hera, I., Iraeta, P., Santos, T. (2011) Barrier effects on vertebrate distribution caused by a motorway crossing through fragmented forest landscape. *Animal Biodiversity and Conservation*, **2**, 331–340.

Tenan, S., Fasola, M., Volponi, S., & Tavecchia, G. (2017) Conspecific and not performance-based attraction on immigrants drives colony growth in a waterbird. *Journal of Animal Ecology,* **86**, 1074-1081.

Tiffin, P., & Ross-Ibarra, J. (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology and Evolution*, **29**, 673-680.

Tigano, A., & Friesen, V. L. (2016) Genomics of local adaptation with gene flow. *Molecular Ecology*, **25**, 2144-2164.

Tigano, A., Shultz, A. J., Edwards, S. V., Robertson, G. J., & Friesen, V. L. (2017) Outlier analyses to test for local adaptation to breeding grounds in a migratory arctic seabird. *Ecology and Evolution*, **7**, 2370-2381.

Toräng, P., Wunder, J., Obeso, J. R., Herzog, M., Coupland, G., & Ågren, J. (2015) Large-scale adaptive differentiation in the alpine perennial herb Arabis alpina. *New Phytologist*, **206**, 459-470.

Trerotola, M., Relli, V., Simeone, P., & Alberti, S. (2015) Epigenetic inheritance and the missing heritability. *Human Genomics*, **9**, 17.

Turesson, G. (1922) The genotypical response of the plant species to the habitat. *Hereditas*, **3**, 211-350.

Uller, T., Isaksson, C., & Olsson, M. (2006) Immune challenge reduces reproductive output and growth in a lizard. *Functional Ecology*, **20**, 873-879.

Veiga, J. P., & Salvador, A. (2001) Individual consistency in emergence date, a trait affecting mating success in the lizard *Psammodromus algirus*. *Herpetologica*, **57**, 99-104.

Verdú-Ricoy, J., Carranza, S., Salvador, A., Busack, S. D., & Díaz, J. A. (2010) Phylogeography of *Psammodromus algirus* (Lacertidae) revisited: Systematic implications. *Amphibia Reptilia*, **31,** 576-582.

Verdú-Ricoy, J., Iraeta, P., Salvador, A., & Díaz, J. A. (2014) Phenotypic responses to incubation conditions in ecologically distinct populations of a lacertid lizard: A tale of two phylogeographic lineages. *Journal of Zoology*, **292**, 184-191.

Via, S. (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences*, **106**, 9939-9946.

Westemeier, R. L., Brawn, J. D., Simpson, S. A., Esker, T. L., Jansen, R. W., Walk, J. W, Saige, K. N. (1998) Tracking the long-term decline and recovery of an

isolated population. *Science*, **282**, 1695-1698.

Whitlock, M. C., & Lotterhos, K. E. (2015) Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of FST. *The American Naturalist*, **186**, S24-36.

Whitney, J. L., Bowen, B. W., & Karl, S. A. (2018) Flickers of speciation: Sympatric colour morphs of the arc-eye hawkfish, *Paracirrhites arcatus*, reveal key elements of divergence with gene flow. *Molecular Ecology*, **27**, 1479-1493.

Williams, G.C. (1966) Adaptation and Natural Selection. *Princeton University Press, Princeton.*

Wootton, J. T., & Pfister, C. A. (2013) Experimental separation of genetic and demographic factors on extinction risk in wild populations. *Ecology*, **94**, 2117-2123.

Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Int. Congr. Genet.,* **1**, 356-366

Yates, M. C., & Fraser, D. J. (2014) Does source population size affect performance in new environments? *Evolutionary Applications*, **7**, 871–882.

Young, A., Boyle, T., & Brown, T. (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, **11**, 413-418.

Zahavi, A. (1975) Mate selection-A selection for a handicap. *Journal of Theoretical Biology*, **54**, 205-214.

Zamora-Camacho, F. J., Reguera, S., Rubiño-Hispán, M. V., & Moreno-Rueda, G. (2014) Eliciting an immune response reduces sprint speed in a lizard. *Behavioral Ecology*, **26**, 115-1brown20.

Zuk, M., Thornhill, R., Ligon, J. D., & Johnson, K. (1990) Parasites and mate choice in red jungle fowl. *Integrative and Comparative Biology,* **30**, 235-244.