UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Genética

TESIS DOCTORAL

Evaluación de la purga genética en poblaciones de censo reducido

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Eugenio López Cortegano

Directora

Aurora García-Dorado García

Madrid, 2018

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE BIOLOGÍA

EVALUACIÓN DE LA PURGA GENÉTICA EN POBLACIONES DE CENSO REDUCIDO

TESIS DOCTORAL

Eugenio López Cortegano

DIRECTORA

Aurora García-Dorado García

MADRID 2017

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE BIOLOGÍA

EVALUACIÓN DE LA PURGA GENÉTICA EN POBLACIONES DE CENSO REDUCIDO

MEMORIA QUE PARA OPTAR AL TÍTULO DE DOCTOR

POR LA UNIVERSIDAD COMPLUTENSE DE MADRID

PRESENTA

Eugenio López Cortegano

Vº Bº del Director

Aurora García-Dorado García

El Doctorando

Eugenio López Cortegano

MADRID 2017

Esta Tesis Doctoral ha sido financiada con una beca predoctoral de Formación de Personal Investigador (BES-2012-055006), y con los proyecto de investigación “Evolución genética del éxito adaptativo en poblaciones amenazadas” (CGL2011-25096) y “Desarrollo de métodos de estimación y predicción de los efectos de la selección natural en poblaciones sometidas a consanguinidad” (CGL2014-53274-P) del Ministerio de Economía y Competitividad (Gobierno de España).

A mis padres, a mis hermanos, y a Qi,

los ejes que definen mi espacio.

AGRADECIMIENTOS

Lo habitual en la investigación científica, pero también en otras muchas otras disciplinas, es que el resultado sea fruto de la colaboración de muchas personas, cada una aportando su grano de arena para construir algo grande a partir de piezas pequeñas, y donde, de forma sinérgica, el resultado final es más grande que la suma de las partes. Este trabajo de tesis no es diferente, sino que es el resultado de la participación de muchas personas que de una forma u otra han contribuido y han permitido que sea posible. Por eso, gracias:

A Aurora García-Dorado, por darme la oportunidad de realizar esta tesis con ella, por su implicación, apoyo, guía y paciencia. Por haber facilitado siempre todos los medios necesarios, y por haberme enseñado tantas cosas bajo su dirección.

A Jinliang Wang, por su permanente atención durante mi estancia en el Instituto de Zoología de Londres, y por enseñarme a optimizar *software*, en particular el desarrollado en esta tesis.

A Armando Caballero, por su enorme hospitalidad durante mi estancia en la Universidad de Vigo, por su implicación, y por su confianza.

A todo el personal del Departamento de Genética, por su apoyo y por su ayuda siempre que ha hecho falta. A Carlos López-Fanjul por enriquecer y ayudar a mejorar los artículos científicos que componen esta tesis con sus comentarios y sugerencias. A Rosario Linacero y Javier Gallego por su gestión. A José por estar siempre dispuesto a echar un cable con el *hardware* y las moscas. A Mónica y a Pablo por su inestimable compañía y amistad. A Diego Bersabé además, por guiarme los primeros años de tesis, estar siempre disponible los posteriores, y por introducirme en el apasionante mundo de la programación en C++ y en GNU/Linux.

A todos los compañeros del Instituto de Zoología de Londres, y de los grupos XB2 y XB5 de la Universidad de Vigo, por su gran sentido del compañerismo, y por hacerme sentir como en casa.

A mi familia, a mis padres y hermanos, por ser siempre consejo sincero y apoyo incondicional, a Qi por cuidarme, inspirarme y acompañarme cada día, y a mi tío Antonio Hernández (Toñito) por su hospitalidad y generosidad.

Finalmente un agradecimiento más general, pero no menos meritorio, para demás amigos y compañeros que me han acompañado a lo largo de estos años, y cuyo apoyo sumado forma también una parte significativa de este trabajo. Gracias.

ÍNDICE

RESUMEN

EVALUACIÓN DE LA PURGA GENÉTICA EN POBLACIONES DE CENSO REDUCIDO

Introducción

La depresión consanguínea puede ser decisiva para la extinción de poblaciones de censo pequeño. Sin embargo, la consanguinidad también incrementa la selección contra los deletéreos responsables de dicha depresión. A este incremento de la selección se le denomina purga genética. Sus consecuencias pueden analizarse utilizando el modelo IP (García-Dorado 2012), que permite predecir la reducción del lastre de consanguinidad y de la depresión consanguínea atribuibles a la purga a través de un coeficiente de consanguinidad purgado (g) que depende del coeficiente de purga (d).

Hasta la fecha se ha publicado una única estima de d para la purga en condiciones no competitivas, obtenida en un experimento con *Drosophila* (Bersabé & García-Dorado 2013). Sin embargo, existen evidencias de que d puede ser mayor en las condiciones más competitivas de las poblaciones silvestres que en cautividad. Además, es necesario estimar d en las propias poblaciones amenazadas de interés, donde no es posible el diseño experimental pero a menudo se dispone de medidas de la eficacia individual y registros genealógicos.

Objetivos

- Estimar experimentalmente el coeficiente de purga en condiciones competitivas.

- Desarrollar una metodología para aplicar el modelo IP a medidas de eficacia en individuos con registros genealógicos., y analizar las propiedades de las estimas de *d* obtenidas usando dicha metodología y su valor predictivo.

- Comparar la capacidad de nuestra metodología para detectar, estimar y predecir la purga con la de modelos previos basados en la consanguinidad ancestral.

Metodología

El objetivo 1 se abordó mediante un experimento de laboratorio con *Drosophila melanogaster*, en el que se evalúo la evolución del lastre de consanguinidad y de la eficacia en dos poblaciones grandes y en conjuntos de líneas de censo reducido ($N \approx 40$) derivadas de ellas, todas mantenidas en condiciones altamente competitivas

El objetivo 2 incluye un análisis teórico que extiende la aplicabilidad del modelo IP previo a cualquier conjunto de datos de eficacia con registros genealógicos, el estudio analítico de los sesgos en la estima de la tasa de depresión consanguínea, el desarrollo de una metodología no lineal de estimación numérica y la programación de un código de acceso libre.

El objetivo 3 se aborda extendiendo la metodología anterior para incluir modelos de purga basados en la consanguinidad ancestral y analizando datos obtenidos mediante simulación.

Resultados

En el experimento con *Drosophila* realizado en condiciones competitivas obtuvimos mucha más purga que la previamente evaluada en condiciones no competitivas, con una estima $d \approx 0.3$.

A continuación obtuvimos ecuaciones genealógicas para el modelo IP y desarrollamos la herramienta informática PURGd que estima los parámetros del modelo, es decir, *d* y la tasa de depresión consanguínea.

Finalmente, al analizar con PURGd datos simulados, encontramos que las estimas IP tienen en general buenas propiedades predictivas. De entre los modelos basados en la consanguinidad ancestral, el modelo de Ballou puede ajustar los datos satisfactoriamente, pero las estimas obtenidas para sus parámetros tienen malas propiedades predictivas en condiciones diferentes de aquellas en que se estimaron.

Conclusiones:

La purga puede ser muy eficaz, revirtiendo la depresión consanguínea en poblaciones silvestres y actuando incluso contra deletéreos de efecto relativamente pequeño, y debe por tanto ser tenida en cuenta en los programas de conservación. El modelo IP y su implementación genealógica a través del programa PURGd es útil para detectar y cuantificar la purga genética a través del coeficiente *d*, y para predecir sus consecuencias en diferentes situaciones, resultando más adecuado que métodos anteriores basados en la consanguinidad ancestral.

SUMMARY

EVALUATION OF GENETIC PURGING IN SMALL SIZED POPULATIONS

Introduction

Inbreeding depression can be an important factor determining the extinction of small sized populations. However, inbreeding also prompts selection against deleterious alleles responsible of inbreeding depression. This increase in selection is referred to as genetic purging. Its consequences can by analyzed by using the IP model (García-Dorado 2012), that allows to predict the reduction of both the inbreeding load and the inbreeding depression due to purging by using a purged inbreeding coefficient (g) which depends on the purging coefficient (d).

So far, only one estimate of d has been published in noncompetitive conditions, obtained in an experiment carried out with *Drosophila* (Bersabé & García-Dorado 2013). However, evidence suggests that d could be higher in the more competitive conditions of wild populations than in captive ones. Furthermore, it is necessary to estimate d in threatened populations, where the experimental approach is not possible but individual measures of fitness and pedigree records are often available.

Objectives

- To estimate the purging coefficient in an experiment where purging occurs in highly competitive conditions.

- To develop a method in order to apply the IP model to fitness measures obtained in individuals with pedigree records, and to analyze the properties of the estimates of *d* obtained using this methodology as well as its predictive value.

- To compare the ability of the previous methodology to detect, estimate and predict purging with that of previous models based on ancestral inbreeding.

Material and methods

Objective 1 was accomplished with a laboratory experiment using *Drosophila melanogaster*, where the evolution of the inbreeding load and fitness was evaluated in two large populations and in small lines ($N \approx 40$) derived from them, maintained in highly competitive conditions.

Objective 2 includes a theoretical analysis that extends the applicability of the IP model to any set of pedigreed fitness data, the analytical study of the bias for the estimates of the inbreeding depression rate, and the development of a numerical nonlineal estimation method and a free access software.

Objective 3 is addressed by extending the previous methodology to purging models based on ancestral inbreeding, and analyzing simulated data.

Results

In the experiments carried out with *Drosophila* in competitive conditions we obtained much more purging than previously evaluated in noncompetitive conditions, with an estimate $d \approx 0.3$.

Furthermore, we deduced genealogical equations for the IP model, and developed the software PURGd that estimates the parameters in the model, *i.e.*, d and the rate of inbreeding depression.

Finally, after analyzing simulated data with PURGd, we found that IP estimates have good predictive properties. Among models based on ancestral inbreeding, Ballou's model can fit data remarkably well, but its parameters have poor predictive properties in conditions different from those where they were estimated.

Conclusions:

Purging can be efficient against inbreeding depression in wild populations, acting even against deleterious alleles of small effect. In consequence, it must be taken into account in conservation programs. The IP model and its genealogical implementation in PUGRd are useful both to detect and quantify genetic purging through the coefficient d, as well as to predict its consequences in different situations, being more appropriate than previous models based on ancestral inbreeding.

INTRODUCCIÓN

EL LASTRE DE CONSANGUINIDAD Y LA DEPRESIÓN CONSANGUÍNEA

La eficacia biológica es el carácter cuantitativo sobre el que actúa la selección natural, y su evolución es determinante para la supervivencia de las poblaciones. Como otros caracteres biológicos, la eficacia media de una población depende de las frecuencias génicas en los loci con efecto sobre la misma, moduladas fundamentalmente por la selección, la deriva, la migración y la mutación. Como consecuencia de la acción de la selección natural sobre la variabilidad surgida por mutación, la gran mayoría de las mutaciones deletéreas que segregan en poblaciones panmícticas grandes son al menos parcialmente recesivas y están a frecuencias bajas. Por tanto, causan un deterioro de la media que puede ser mucho menor que el que correspondería a su expresión en homocigosis. Así pues, una parte del deterioro de la eficacia media que podrían causar no se expresa en la población grande panmíctica, constituyendo el lastre genético oculto en heterocigosis. Habitualmente este lastre se conoce como lastre de consanguinidad porque cualquier proceso que cause un incremento de la consanguinidad y, por tanto, de la homocigosis, revelará una parte de dicho lastre oculto causando un deterioro de la eficacia media conocido como depresión consanguínea. El lastre de consanguinidad se mide comúnmente como número *B* de equivalentes letales por gameto (Morton *et al.* 1956, Charlesworth & Charlesworth 1999) y, en ausencia de selección, representa la tasa a la que se deteriora la eficacia al aumentar la consanguinidad, es decir, la tasa de depresión consanguínea. El fenómeno de la depresión consanguínea constituye un elemento clave

en la evolución de fenómenos esenciales, como el sexo, la recombinación o los sistemas reproductivos (Charlesworth & Charlesworth 1987, Charlesworth & Willis 2009, García-Dorado 2017). Así mismo, como se expone a continuación, la depresión consanguínea puede ser un factor determinante del riesgo de extinción de poblaciones y especies y, por tanto, es un elemento esencial en la genética de la conservación de poblaciones amenazadas.

Una característica común de todas las poblaciones amenazadas es que su censo ha sufrido alguna reducción, a menudo debido a la acción humana, lo cual desencadena diversos procesos estocásticos, tanto demográficos como genéticos, que pueden comprometer su supervivencia dramáticamente (Lande 1988). En concreto, una consecuencia directa de la reducción del censo es el incremento de la consanguinidad y, por tanto, de la frecuencia de los genotipos homocigotos a expensas de los heterocigotos, proceso que también puede desencadenarse debido al incremento del grado de fragmentación poblacional así como de otros patrones de falta de panmixia que promuevan el apareamiento preferente entre individuos emparentados. Por tanto, uno de los peligros a que se enfrentan las poblaciones amenazadas es la depresión consanguínea de la eficacia biológica.

Los efectos perniciosos de la depresión consanguínea sobre la viabilidad y la fecundidad son conocidos desde muy antiguo, y se ha documentado una extensa evidencia de este fenómeno en poblaciones experimentales y naturales, habiéndose obtenido numerosas estimas del lastre de consanguinidad (Ralls *et al.* 1988, Hedrick & Kallinowski 2000, Keller & Waller 2002, Frankham 2005, O'Grady *et al.* 2006, Hedrick & García-Dorado 2016). Por ejemplo, la depresión consanguínea se ha detectado y evaluado repetidas veces en poblaciones humanas (Morton *et al*. 1956, Bittles & Neel 1994). Así,

Morton y colaboradores (1956), en su trabajo pionero, que estableció el modelo genético fundamental que da cuenta de la depresión consanguínea, obtuvieron una estima de 1.5 a 2.5 equivalentes letales en poblaciones humanas, que fue validada por trabajos posteriores (Lee *et al.* 1996). En concordancia cualitativa con estas estimas, se han verificado las graves consecuencias que tienen sobre la eficacia los sistemas endogámicos que han caracterizado a algunas familias y, particularmente a algunas dinastías reales (Ager 2005, Berra *et al*. 2010, Álvarez *et al.* 2015).

El lastre de consanguinidad es particularmente elevado en poblaciones naturales, con un valor medio en torno a $B=6$ equivalentes letales (O'Grady *et al.* 2006), como cuatro veces superior a las estimas de un meta-análisis previo centrado en poblaciones mantenidas en cautividad (Ralls *et al.* 1988). Existen dos razones fundamentales para esta diferencia. Por una parte, las poblaciones silvestres están habitualmente sometidas a ambientes más adversos con condiciones más competitiva, y existen evidencias de que los efectos deletéreos de las mutaciones se exacerban en estas circunstancias (Crnokrak & Roff 1999, Ávila & García-Dorado 2002, Yun & Agrawal 2014), por comparación con las condiciones de mantenimiento en cautividad donde el ambiente es más favorable y menos competitivo, disponiéndose incluso de cuidados veterinarios. Por otra parte, las poblaciones cautivas se mantienen habitualmente con censos mucho menores que las silvestres, y es posible que parte del lastre de consanguinidad se haya perdido como consecuencia de dicho censo o de cuellos de botella ocurridos en el pasado.

En todo caso, el elevado lastre de consanguinidad de las poblaciones naturales implica que la depresión consanguínea de las poblaciones silvestres amenazadas puede comprometer de modo crítico su viabilidad, haciéndose necesario considerar el control de la depresión consanguínea como un elemento esencial de un programa de

conservación. La importancia de la depresión consanguínea en el deterioro de las poblaciones amenazadas es una de las razones por las que el censo poblacional es fundamental para elaborar directrices conservacionistas o determinar la viabilidad de una población amenazada. Así por ejemplo, de los cinco criterios que establece la Unión Internacional para la Conservación de la Naturaleza (IUCN) para determinar el grado de amenaza a que está sometida una población, tres hacen referencia directa al censo poblacional (A, C y D), y uno de ellos (D) propone directamente un umbral de tamaño poblacional que permite por sí solo catalogar una especie como vulnerable (menos de 1000 individuos maduros), amenazada (menos de 250), o en estado crítico de amenaza (menos de 50) (IUCN, 2001).

El establecimiento de el llamado tamaño mínimo viable poblacional (MVP) está también fuertemente condicionado por la depresión consanguínea (Shaffer 1981). El valor aceptado del MVP es de gran trascendencia, pues podría usarse como criterio para dejar de destinar recursos a la conservación de poblaciones que no lo alcancen. Clásicamente, siguiendo las recomendaciones de mejoradores animales, se ha considerado que el tamaño mínimo poblacional necesario para prevenir a corto o medio plazo la extinción causada por depresión consanguínea correspondía a un censo efectivo de 50, lo cual impone este censo como límite inferior del MVP (Franklin 1980, Soulé 1980). Más recientemente algunos autores han recomendado aumentar esta valor a 100 (Frankham *et al.* 2014), precisamente considerando que, con el elevado lastre de consanguinidad ($B=6$) obtenido para poblaciones naturales (O'Grady *et al.* 2006), es necesario un censo efectivo de al menos 100 para evitar que la caída de la eficacia biológica sea mayor del 10% tras 5 generaciones de consanguinidad. De aplicarse este criterio, muchas poblaciones actualmente catalogadas por la IUCN como en estado crítico de amenaza quedarían sin cobertura por su presunta inviabilidad.

LA PURGA GENÉTICA

Como acabamos de describir, la depresión consanguínea se produce fundamentalmente debido a la expresión de los componentes recesivos de los alelos deletéreos en los homocigotos generados por consanguinidad. Sin embargo, esta misma expresión expone dichos componentes recesivos a la acción de la selección natural. Es decir, el aumento de la consanguinidad permite a la selección natural actuar sobre componentes de los efectos deletéreos que estaban previamente ocultos pero que se expresan en los homocigotos generados por consanguinidad (Wang & Hill 1999). Esta selección purificadora desencadenada por la consanguinidad se conoce como selección purgadora, o simplemente purga genética, si bien el término "purga" se utilizó en ocasiones en el pasado de forma más inespecífica, para designar cualquier modo de selección purificadora contra alelos deletéreos. La purga genética reduce pues la frecuencia de los deletéreos parcial o totalmente recesivos, invalidando hasta cierto punto las predicciones de Morton *et al*. (1956), obtenidas bajo un modelo en que el lastre de consanguinidad representa la tasa de depresión consanguínea porque se ignoran los efectos de la selección sobre el cambio en las frecuencias génicas. La reducción de la frecuencia media de los deletéreos atribuible a la purga tiene dos consecuencias fundamentales. Por una parte, el lastre de consanguinidad se reducirá más de lo esperable por simple deriva. Por otra, la depresión consanguínea esperada para la eficacia biológica será menor que la predicción neutra del modelo de Morton *et al*.

Desde el punto de vista teórico, se han llevado a cabo diversos análisis destinados a analizar las consecuencias de la purga genética. Por una parte, Glémin (2003) utilizando la teoría de difusión, ha obtenido predicciones sobre la reducción en la frecuencia media de deletéreos atribuible a la purga en una población en equilibrio. Para ello distingue dos

tipos de purga. Por una parte la que actúa sobre el exceso de homocigotos de una población de tamaño reducido, por comparación con la homocigosis de la población infinita. Por otra parte, la que actúa sobre el exceso de homocigotos debido al apareamiento entre parientes en una población no panmíctica respecto de otra panmíctica del mismo tamaño. En ambos casos, evalúa la purga como la reducción de la frecuencia media de los alelos deletéreos con respecto a la frecuencia media de la población en equilibrio no purgada, que se toma como referencia. Su trabajo muestra que, en lo que refiere a la purga atribuible al tamaño finito de una población panmíctica, el proceso estocástico de cambio de frecuencias génicas (es decir, la deriva genética) puede interferir con el proceso de purga, haciendo que en poblaciones pequeñas la purga solo sea eficiente para alelos deletéreos muy recesivos. Por este motivo, la purga así definida solo es eficaz por encima de cierto valor umbral del censo, aunque su eficiencia disminuye cuando los censos son tan elevados que el incremento de homocigosis respecto de la población infinita se vuelve irrelevante. Este modelo contribuye a explicar las diferentes causas de la purga, y las limitaciones observadas en la detección de la misma en algunos trabajos experimentales (Byers & Waller 1999). No obstante, debe notarse que, debido a esta definición de la eficiencia de la purga como reducción de la frecuencia media de deletéreos, una población puede estar sometida a una purga más eficiente que otra de mayor censo, y aun así presentar una eficacia media menor, por tener mayor frecuencia de homocigotos.

Por otra parte, García-Dorado (2012) analizó el efecto de la purga durante un proceso en que se incrementa la consanguinidad, desarrollando ecuaciones sencillas que predicen la evolución de la eficacia media y del lastre de consanguinidad atribuible a los efectos de la consanguinidad y la purga sobre la variabilidad genética inicial (predicciones IP). Estas ecuaciones son función de un coeficiente de consanguinidad purgado g, un

análogo al coeficiente de consanguinidad de Wright (F) que permite predecir la evolución de la homocigosis para los alelos deletéreos incorporando un cálculo determinista de la reducción de la frecuencia génica de dichos deletéreos causada por la purga. A su vez, el coeficiente g es función de un coeficiente purga (d), que depende de la magnitud del efectos de los alelos deletéreos que permanece oculto en heterocigosis a causa de la recesividad. Las predicciones IP son válidas tanto cuando la consanguinidad se produce como consecuencia de una reducción del censo poblacional como cuando se genera por falta de panmixia, pudiendo calcularse en función del censo efectivo en el primer caso y en función de las relaciones genealógicas en ambos, si bien las ecuaciones derivadas hasta el presente trabajo solo manejaban genealogías sin solapamiento de generaciones. Además, en el mismo trabajo se desarrollaron también predicciones bajo un modelo más completo (predicciones Full Model, o FM) que tiene en cuenta la aparición de mutaciones deletéreas nuevas durante el proceso, así como los efectos de la selección estándar no purgadora, y que resulta más exacto cuando la consanguinidad progresa lentamente o estamos interesados en predicciones a largo plazo.

En este análisis IP, se considera que la purga es eficiente cuando la eficacia media de la población consanguínea es mayor a la esperada de acuerdo a la predicción sin purga de Morton y colaboradores (1956), o cuando el lastre de consanguinidad sufre una reducción mayor que la esperada solo por deriva. El modelo IP muestra que cualquier reducción del censo conducirá a un nuevo equilibrio con menor eficacia media que la población original, aunque la diferencia puede ser imperceptible si la reducción del censo es pequeña, pero aun así la purga se considerará eficiente en la medida en que la eficacia del nuevo equilibrio sea superior a la esperada solo por depresión consanguínea. No obstante, el criterio de Glémin determinará que ha habido purga eficiente siempre que la frecuencia media de deletéreos sea menor en el nuevo equilibrio, lo cual nunca se

acompañará de un incremento neto de la eficacia esperada, ya que ésta depende de las frecuencias genotípicas y no de las génicas.

Es interesante notar que la eficiencia de la purga tal como se define en el modelo IP, aumenta con *d* y se reduce con el censo efectivo (N). De hecho, aunque el análisis IP predice cierta purga siempre que *d* sea mayor que cero, se ha comprobado mediante simulación que la eficiencia de la purga se ve anulada por la deriva cuando el producto N*d* es del orden de la unidad o menor, en concordancia cualitativa con las conclusiones de Glémin. Así pues, la purga en poblaciones muy pequeñas solo será eficiente contra alelos letales (o deletéreos severos) de efecto quasi-recesivo. No obstante, cuando la consanguinidad aumenta lentamente, la purga es más eficiente pero también más lenta, pudiendo ser imperceptible durante las primeras generaciones pero causar después una recuperación de la depresión consanguínea inicial. Así pues, la detección de la purga contra deletéreos no severos tras una reducción del censo, requerirá datos de la eficacia biológica durante periodos prolongados de consanguinidad con censos efectivos no demasiado pequeños.

Desde el punto de vista práctico, el modelo IP tiene la ventaja de permitir la predicción de la eficacia biológica de una forma sencilla, empleando una expresión análoga a la de Morton, pero usando la consanguinidad purgada en lugar de la consanguinidad de Wright. El obstáculo principal para el uso de esta expresión es que se precisan estimas del coeficiente de purga, un parámetro del que hasta la fecha se dispone de información muy escasa.

LA PURGA GENÉTICA: MÉTODOS DE DETECCIÓN Y EVIDENCIA EMPÍRICA SIN DATOS GENEALÓGICOS

Existen distintas aproximaciones al problema de la detección de la purga genética basadas en la observación de alguna de sus consecuencias, como la reducción de la depresión consanguínea o del lastre de consanguinidad a valores inferiores a los esperados en ausencia de selección. Por una parte, se han utilizado datos obtenidos de poblaciones naturales o de diseños experimentales en que la purga se estudia en función de los efectos de una reducción del censo o de sistemas regulares de apareamientos no panmícticos. Sin embargo, las evidencias no son del todo consistentes, tanto en lo que refiere a especies animales como vegetales (Byers & Waller 1999, Crnokrak & Barrett 2002, Leberg & Firmin 2008). En general, no suele detectarse purga durante procesos en que la tasa de aumento de consanguinidad es elevada (*i.e.*, el censo efectivo es muy reducido), con excepción de la purga contra deletéreos severos muy recesivos. Por ejemplo, en el caso de líneas mantenidas mediante apareamientos entre hermanos solo es posible detectar selección contra letales quasi-recesivos (Hedrick 1994, Frankham *et al.* 2001). Por este motivo, la capacidad de la purga para actuar contra deletéreos no severos ha sido frecuentemente cuestionada (Frankham *et al*. 2014).

Sin embargo, es común la detección de la purga durante incrementos lentos de la consanguinidad, como es el caso de las poblaciones de censo efectivo moderado (Latter *et al.* 1995, Crnokrak & Barrett 2002, Leberg & Firmin 2008). Así, numerosos trabajos han revelado que la eficiencia de la purga aumenta cuando la consanguinidad progresaba de forma lenta. Por ejemplo, Pedersen y colaboradores (2005) encontraron que, con el mismo nivel de consanguinidad ($F\approx0.67$), la depresión consanguínea para fecundidad en líneas de *Drosophila melanogaster* era significativamente mayor para las líneas obtenidas

mediante apareamientos entre hermanos durante cinco generaciones que en aquellas mantenidas con un censo de cuatro durante nueve generaciones. Por su parte, Pekkala y colaboradores (2012) llegaron a una conclusión similar en un trabajo con *Drosophila littoralis*, donde líneas de censo N=10 presentaban una depresión consanguínea rápida y una tasa de extinción considerable, mientras que la depresión era casi inapreciable en líneas mantenidas con un censo cuatro veces mayor. En otro trabajo, Swindell y Bouzat (2006) mostraron que la depresión consanguínea era sustancialmente menor para líneas con el mismo coeficiente de consanguinidad ($F=0.375$), pero mayor consanguinidad ancestral ($F_a = 0.531$ frente a $F_a=0.250$), como sugería Ballou (Ballou 1977, Swindell & Bouzat 2006). Todo ello apunta a que en efecto la purga puede ser eficaz si la consanguinidad progresa lentamente.

La detección de la purga también puede parecer inconsistente en escenarios supuestamente propicios para valorar el alcance de la purga genética con datos de poblaciones naturales, como son las poblaciones aisladas de animales insulares que han sufrido cuellos de botellas. Por ejemplo, no se ha detectado purga en una población insular aislada de tordos (*P. traversi*), pero si en otra *(Petroica australis rakiura*), si bien ambas contaban con una historia demográfica que habría generado una importante consanguinidad (Laws & Jamieson 2011, Kennedy *et al*. 2014). Sin embargo, en el primer caso la consanguinidad se debe a cuellos de botella drásticos, mientras que en el segundo, en que sí se detectó purga, se había generado lentamente durante un periodo prolongado con censos efectivos moderados.

En general, estos hallazgos son coherentes con las predicciones IP anteriormente mencionadas, según las cuales, cuando la consanguinidad aumenta más lentamente la purga se vuelve más eficiente pero también más lenta, pues su acción requiere la

acumulación previa de cierta consanguinidad. Así pues, a menudo no se detecta purga porque la consanguinidad es demasiado rápida o porque, siendo relativamente lenta, no se observa el proceso durante suficiente tiempo para que se manifiesten los efectos de la purga.

Por otra parte, el modelo IP proporciona una herramienta para la detección y la evaluación de la purga a través de la obtención de la estima de *d* que maximice el ajuste de las predicciones IP a la evolución observada de la eficacia media o del lastre de consanguinidad. Con anterioridad al inicio de esta tesis, sólo un experimento llevado a cabo en *D. melanogaster* por Bersabé y García-Dorado (2013) ha estimado el valor del coeficiente de purga. Este estudio proporcionaba cierto soporte empírico al modelo, pues las estimas que se obtuvieron de *d*, utilizadas en el modelo IP, producían predicciones que se ajustaban a los resultados experimentales mucho mejor que las del modelo clásico sin purga. Sin embargo, las líneas usadas en ese estudio tenían censos efectivos bajos (seis o doce), de modo que solo podía esperarse detectar la acción de la purga contra deletéreos relativamente severos con valores elevados de *d*. Además, esas líneas se mantuvieron en condiciones no competitivas, y como consecuencia, tanto la tasa de depresión consanguínea como la purga pudieron haber sido sustancialmente menores de lo esperado en una población silvestre en condiciones competitivas. Por tanto, es necesario llevar a cabo un experimento que permita evaluar la purga genética en líneas con censos efectivos mayores y en condiciones más adversas o más competitivas.

LA PURGA GENÉTICA: MÉTODOS DE DETECCIÓN Y EVIDENCIA EMPÍRICA UTILIZANDO DATOS GENEALÓGICOS

En lo que respecta a poblaciones amenazadas reales, generalmente no es posible realizar experimentos diseñados para estimar los parámetros genéticos que determinan las consecuencias de la consanguinidad y la purga (δ y d). Sin embargo, es frecuente disponer de información genealógica en programas de conservación tanto *ex situ* como *in situ*. De hecho, existe la posibilidad de reconstruir o completar genealogías a partir de datos moleculares (Lynch & Ritland 1999, Fernández & Toro 2006, Wang 2011), algo que puede ser de gran interés, ya que las genealogías no solo permiten calcular coeficientes de consanguinidad y parentesco sino también entender mejor la estructura poblacional (Pemberton 2008), así como patrones o eventos de dispersión (Norman & Spong 2015) y estrategias reproductoras (Pemberton *et al.* 1992, Wang *et al*. 2011).

Por este motivo se han propuesto algunos métodos capaces de utilizar información genealógica en este sentido, de los cuales el más utilizado hasta la fecha ha sido el desarrollado por Ballou (1997).

Este autor definió un coeficiente de consanguinidad ancestral (F_a) que mide el porcentaje del genoma de un individuo que ha estado expuesto en homocigosis al menos una vez en un ancestro. La utilidad de este coeficiente se debe a que los individuos que tengan más consanguinidad ancestral procederán de linajes en que ha habido más posibilidades de purga que aquellos que tengan el mismo nivel de consanguinidad estándar (F), pero menor consanguinidad ancestral. Esta consideración ha dado lugar al desarrollo de diversos modelos predictivos (Ballou 1997, Boakes & Wang 2005). En principio todos ellos se plantean como modelos lineales en que la eficacia se predice en función de dos variables regresoras (F y F_a) y su correspondiente interacción, o de

diversos subconjuntos de los tres factores, aunque los análisis estadísticos están a menudo basados en regresión logística. Tanto Ballou (1997) como Boakes et al (2007) llevaron a cabo análisis utilizando datos de eficacia para individuos de las genealogías de varias poblaciones animales de zoos, y encontraron un efecto de la purga que fue muy pequeño a nivel global y significativo en pocas ocasiones. No obstante, Boakes y colaboradores hacen notar que las poblaciones de este tipo están sometidas a manejo genético y se mantienen en condiciones benignas, factores que pueden entorpecer notablemente la purga genética.

Por otra parte, Gulisija y Crow (2007), también desarrollaron un modelo para evaluar la purga contra letales recesivos en genealogías cortas en función de otra medida (*O*) de las oportunidades de purga. El valor de *O* en un individuo, calculable a partir de la genealogía, representa la probabilidad de que un alelo de dicho individuo sea copia de otro alelo que estuvo en homocigosis en algún ancestro. Utilizando este método, analizaron un gran número de genealogías cortas de una población de ganado vacuno, y concluyeron que la purga en contra de letales recesivos durante seis generaciones había reducido la depresión consanguínea en un 12.6%. Estos autores consideran que su método solo detecta la purga contra deletéreos de efecto grande, para los que la purga es más eficiente a corto plazo, pero que la purga contra alelos de efectos más moderados podría ser importante a largo plazo.

Tanto los métodos basados en el coeficiente de consanguinidad ancestral como el método de Gulisija y Crow se basan en el ajuste a un modelo (en su planteamiento más sencillo un modelo lineal) en que las variables regresoras son coeficientes definidos de acuerdo a consideraciones heurísticas acerca de la purga. Por el contrario, las predicciones IP se deducen de un modelo genético en función del efecto de la purga genética sobre las

frecuencias génicas, y donde la variable predictora que incorpora la posible purga (g) puede calcularse en función de las genealogías. Así pues, parece razonable esperar que un método basado en el ajuste a estas predicciones tenga mayor sensibilidad, además de proporcionar estimas de parámetros con una utilidad predictiva (δ y d). No obstante, hasta el presente trabajo subsistía el obstáculo práctico de que el cálculo genealógico de g solo estaba desarrollado para genealogías sin solapamiento de generaciones.

OBJETIVOS

La finalidad de esta tesis es ahondar en nuestro conocimiento sobre los efectos de la purga genética en la evolución de la eficacia de las poblaciones, y desarrollar herramientas para su detección, evaluación y predicción que puedan ser útiles en ámbitos aplicados como lo es la conservación de poblaciones amenazadas. Para ello utilizaremos el modelo IP, en el cual la purga depende en última instancia del coeficiente de purga *d* (García-Dorado 2012), y que ha demostrado previamente en estudios de simulación una buena capacidad predictiva. Los objetivos concretos abordados son los que se exponen a continuación.

OBJETIVO 1.- Estimar el coeficiente de purga en condiciones competitivas en una población experimental de *Drosophila melanogaster*.

Este objetivo complementa el trabajo llevado a cabo por Bersabé y García-Dorado (2013) bajo condiciones de mínima competitividad. Dado que, como se ha expuesto anteriormente, los efectos deletéreos parecen ser mayores en condiciones más competitivas, nuestro propósito es estimar el coeficiente de purga cuando las condiciones de mantenimiento durante el aumento de la consanguinidad son más parecidas a las habituales en la naturaleza. Además, aquí utilizamos poblaciones con mayor censo efectivo que Bersabé y García-Dorado, lo cual nos da la oportunidad de estimar la purga atribuible a deletéreos de efecto no severo. La información obtenida es importante para evaluar las expectativas de purga en poblaciones amenazadas mantenidas *in situ*. Este objetivo se aborda en el artículo "*Estimation of genetic purging under competitive conditions*", publicado en el número 70 de la revista *Evolution* (*International Journal of*

Organic Evolution), y constituye íntegramente el primer capítulo en esta tesis. El material suplementario del mismo se incluye en esta tesis como Apéndice 1.

OBJETIVO 2.- Desarrollo de un método para detectar la purga en poblaciones con registros genealógicos y de una herramienta informática para su aplicación práctica.

En primer lugar, Para acometer este objetivo, se adaptó el modelo de cálculo del coeficiente de parentesco purgado (García-Dorado, 2012) para el análisis de genealogías que incluyan generaciones solapadas. Además se dedujo un factor de corrección que permite inferir la tasa de depresión consanguínea a partir del coeficiente de regresión del logaritmo de la eficacia individual en el coeficiente de consanguinidad. Por último se desarrolló una herramienta informática en lenguaje C++, que hemos llamado PURGd, para la detección y estima de la purga en poblaciones genealógicas, y que está disponible para su descarga desde la web del grupo de "Mecanismos genéticos de la evolución, mejora y conservación de las poblaciones" del Departamento de Genética de la Universidad Complutense de Madrid (https://www.ucm.es/genetica1/mecanismos). La guía de usuario de este programa está incluida en la tesis como Apéndice 2. Esta herramienta propone, además del clásico análisis de regresión lineal de un modelo para la eficacia logarítmica, un método numérico de ajuste al modelo predictivo exponencial. Este objetivo constituye la materia del segundo capítulo de esta tesis, contenido en el artículo "*Predictive model and software for inbreeding-purging analysis of pedigreed populations*" publicado en la revista *G3* (*Genes, Genomes, Genetics*).

OBJETIVO 3.- Comparación de diversos modelos de detección de la purga y evaluación de su capacidad predictiva usando PURGd con datos simulados.

Para abordar este objetivo llevamos a cabo un análisis de las propiedades de la herramienta anterior (PURGd) analizando diversos aspectos de su rendimiento en análisis de datos obtenidos mediante simulación. Además, se implementaron en PURGd otros métodos de detección de la purga basados en la consanguinidad ancestral de Ballou con el fin de identificar la metodología de análisis más adecuado. Con objeto de obtener predicciones de la eficacia media utilizando el método de Ballou, se dedujo una expresión para predecir la evolución esperada de la consanguinidad ancestral en poblaciones panmícticas de censo constante. En todos los casos, se analizaron las diversas fuentes de sesgo que pueden afectar las estimas, y se verificó la calidad de las predicciones obtenidas utilizando dichas estimas. Este trabajo constituye el tercer capítulo de esta tesis, también con formato de artículo y actualmente enviado para su publicación. El material suplementario del mismo se incluye en esta tesis como Apéndice 3.

CAPÍTULO PRIMERO

ESTIMATION OF GENETIC PURGING UNDER COMPETITIVE CONDITIONS

Eugenio López-Cortegano[1], Ana Vilas[2], Armando Caballero[2] and Aurora García-Dorado[1]

[1] Departamento de Genética. Facultad de Biología. Universidad Complutense. Madrid 28040. Spain.

[2] Departamento de Bioquímica, Genética e Inmunología. Facultad de Biología. Universidad de Vigo. Vigo 36310 (Pontevedra). Spain.

Running title: Genetic purging in competitive conditions

Key words: Inbreeding load; inbreeding depression; genetic diversity; fitness; productivity

Corresponding author:

Aurora García-Dorado

Departamento de Genética, Facultad de Biología, Universidad Complutense, Ciudad Universitaria, Madrid 28040, Spain.

Phone number: 0034 913 944 975

Email address: augardo@ucm.es

López-Cortegano, E., A. Vilas, A. Caballero, A. García-Dorado. 2016. Estimation of genetic purging under competitive conditions. *Evolution* 70, 1856-1870.

Abstract

Inbreeding depression for fitness traits is a key issue in evolutionary biology and conservation genetics. The magnitude of inbreeding depression, though, may critically depend on the efficiency of genetic purging, the elimination or recessive deleterious mutations by natural selection after they are exposed by inbreeding. However, the detection and quantification of genetic purging for nonlethal mutations is a rather difficult task. Here we present two comprehensive sets of experiments with *Drosophila* aimed at detecting genetic purging in competitive conditions and quantifying its magnitude. We obtain, for the first time in competitive conditions, an estimate for the predictive parameter, the purging coefficient (d), that quantifies the magnitude of genetic purging, either against overall inbreeding depression ($d \approx 0.3$), or against the component ascribed to nonlethal alleles ($d_{NL} \approx 0.2$). We find that competitive fitness declines at a high rate when inbreeding increases in the absence of purging. However, in moderate size populations under competitive conditions, inbreeding depression need not be too dramatic in the medium to short term, as the efficiency of purging is also very high. Furthermore, we find that purging occurred under competitive conditions also reduced the inbreeding depression that is expressed in the absence of competition.

Introduction

The reduction of fitness due to inbreeding is a substantially ubiquitous phenomenon that is relevant to evolutionary and conservation genetics and to animal breeding. This general phenomenon is known as "inbreeding depression", a term that we also use for the amount of fitness decline caused by any given increase of inbreeding. Evidence accumulated during the last decade has shown that inbreeding depression in wild populations under competitive conditions uses to be about fourfold that previously reported for captive populations maintained in benign conditions (Ralls *et al*. 1988; Keller and Waller 2002; Kruuk *et al*. 2002; Liberg *et al*. 2005; O'Grady *et al*. 2006; Walling *et al*. 2011; Kennedy *et al*. 2014), although there is substantial variation among populations (Hedrick and Kalinowsky 2000).

The main cause of inbreeding depression is the large amount of genetic load that is concealed in heterozygosis in noninbred populations due to the recessive components of deleterious effects (Charlesworth and Willis 2009). Since this load is expressed under inbreeding due to increased homozygosity, it is usually denoted inbreeding load *B.* Thus, inbreeding depression is the expression of the previously concealed inbreeding load. According to classical theory, B can be interpreted in terms of number of lethal equivalents (Morton *et al*. 1956) and, under the simplifying assumptions of the model (no linkage or epistasis), it equals the rate δ at which fitness would decline with increasing inbreeding in the absence of selection. In fact, estimates of this rate of inbreeding depression are used as estimates of the inbreeding load, and here both that rate and this inbreeding load will be denoted by δ.

However, the prediction of inbreeding depression requires taking into account genetic purging, which is the selection prompted by inbreeding as it exposes the recessive

component of deleterious effects in the homozygotes. Ignoring purging can have dramatic effect, both when evaluating the evolutionary consequences of inbreeding and when making recommendations in conservation to minimize them. However, under the same simplifying assumptions as in Morton's model (no linkage or epistasis), the joint consequences of inbreeding and purging can be predicted using a simple theoretical model (García-Dorado 2012). This Inbreeding-Purging (**IP**) model provides good approximations for the evolution of the mean and of the inbreeding load for fitness traits, which depend on a purging coefficient (*d*) representing the recessive component of the deleterious effects, responsible for both inbreeding depression and purging. Estimating this parameter is essential in order to obtain predictions of the joint consequences of inbreeding and purging.

The predictions of the **IP** model show that slow inbreeding leads to more efficient purging, because natural selection has more opportunities to operate before a given inbreeding level is attained, but also delays its effects. Thus, under slow inbreeding, fitness depression can progress during some generations at a rate that is very similar to that expected in the absence of selection, but fitness can later recover substantially due to purging. As a consequence, purging could be of critical importance in practical situations, even for populations where its consequences are negligible in experimental conditions. Furthermore, this could explain why purging, although often observed, is not systematically detected (Crnokrak and Barrett 2002; Leberg and Firmin 2008). For example, the efficiency of purging for nonlethal genes is low with fast inbreeding (e.g. full-sib mating; Hedrick 1994; Frankham *et al*. 2001), but this does not imply that it should be also inefficient under slower inbreeding.

Similarly, it is not surprising to observe important inbreeding depression in the

short term for populations whose size has been reduced to a N_e value that is still relatively large, but this does not imply that fitness will not recover later, because purging requires a large number of generations to become appreciable when N_e is large. Thus, Bryant *et al*. (1999) detected substantial inbreeding depression after 5 generations with $N_e = 90$ for housefly, but the depression observed at later generations was much smaller than expected by ignoring purging (neutral prediction) on the basis of the early fitness decline (Bryant *et al*. 1999). Also Kennedy *et al*. (2014) failed to detect purging for robins during a decade of intensive monitoring in the wild with important inbreeding depression, but the rate of inbreeding was small and *F* increased just about 5% during the whole period, so that a decade could have not been enough for purging to occur (Kennedy *et al*. 2014). In general, low purging efficiency has been detected under fast inbreeding, as very small N_e usually induces purging just of lethal or severely deleterious alleles. On the contrary, relevant purging has often been detected under slow inbreeding (Latter *et al*. 1995; Crnokrak and Barrett 2002; Pedersen *et al*. 2005; Swindell and Bouzat 2006; Leberg and Firmin 2008; Pekkala *et al*. 2012). Furthermore, genetic management protocols, as equalization of family contributions or minimum kinship, can also be partially responsible for the inefficiency of purging in other cases (Woodworth *et al*. 2002; Reed *et al* 2003). Very small purging has also been detected in zoo populations with pedigree data (Ballou 1997; Boakes *et al*. 2006), but the authors of these analyses warn that, in those populations, management could diminish the efficiency of purging and, furthermore, purging might have occurred prior to the analyzed period.

In addition, the detection of purging can often be difficult if there is a concurrent adaptive processes. The reason is that adaptation is expected to be more efficient in the large population used as control than in the small experimental populations, so that purging is underestimated. Furthermore, there may be insufficient information about the

amount of inbreeding depression that would be expected in the absence of purging, often due to poor estimates of the initial inbreeding load or of the actual rates of inbreeding (*i.e.*, of the effective population size). In that case, it may not be possible to check whether the observed inbreeding depression is smaller than predicted by ignoring purging.

Furthermore, the larger rates of inbreeding depression detected in the wild, compared to captive populations, are likely to be due to larger recessive deleterious components (d), either because alleles with large d values were already purged in captive conditions due to a previous history of inbreeding, or because deleterious effects are larger when expressed in harsh environments. One of the main differences between wild and captive conditions is that competition is usually larger in the wild, where population size is heavily constrained by resources limitation. In fact, the average effect of deleterious mutation has been found to be particularly large when expressed in competitive conditions (Ávila and García-Dorado, 2002), and it has been reported that the inbreeding load in *Drosophila* is larger for fitness assayed in more competitive conditions than in less competitive ones (Yun and Agrawal 2014). Therefore, purging can be less efficient and more difficult to detect when operating in captive or noncompetitive conditions.

Summarizing, the occasional failure to detect purging of nonlethal alleles, can usually be ascribed to at least one of the following features: (i) inbreeding increases too fast to allow efficient purging (*i.e.*, N_e is exceedingly small so that $N_e d < 1$ and genetic drift overwhelms purging); (ii) inbreeding increases slow enough to allow for efficient purging (N_e is large), but the number of generations of inbreeding analyzed is too small since, for large N_e, it takes more generations for purging to act; (iii) inbreeding occurs under managing strategies that reduce purging efficiency; (iv) inbreeding occurs in conjunction

with adaptive processes; (v) there is no solid information about the effective population size N_e and/or the inbreeding load δ in the base population; (vi) important previous inbreeding is being disregarded, so that most purging may have occurred in the past; (vii) purging was inefficient because it occurred under benign or noncompetitive conditions, where both the rate of inbreeding depression and purging are smaller than in wild competitive conditions.

Therefore, when predicting the consequences of inbreeding, there is no ground to disregard purging just because it has not been consistently detected in every situation studied. Detecting purging is experimentally demanding, as it requires monitoring the evolution of fitness in the absence of substantial adaptive processes, for not too small populations, during a considerable number of generations and in the absence of genetic management. Optimally, we should also have good information on the inbreeding load in the base population, and reliable information on the rate of inbreeding (*i.e.*, on N_e). In addition, the efficiency of purging should be evaluated in environmental conditions that are similar to those of practical interest, particularly regarding competitive conditions. It is difficult to obtain data that meet all these requirements, but this does not mean that, in real cases, purging will not occur in its time given appropriate population sizes.

Thus, it is necessary to evaluate purging efficiency and to estimate the purging coefficient operating in competitive conditions in experiments designed in the light of theory. So far, the only estimate available for the purging coefficient comes from a *Drosophila* experiment carefully designed for that purpose, where the consequences of inbreeding and purging on egg to pupae viability were investigated in lines derived from a wild population (Bersabé and García-Dorado 2013, B&GD hereafter). However, these lines were maintained under low density conditions, with no competition regarding

fecundity or mating ability (single mating vials), and with small effective size N_e = 6 or 12. The results provided strong evidences for purging, but suggested that purging of nonlethal alleles was modest, so that a larger effective population size might have been more efficient to estimate the purging coefficient applying to the nonlethal inbreeding load.

Here, we present results from two parallel and highly repeated experiments, each performed using a different large laboratory population of *Drosophila melanogaster* that, at some time, was used to derive a large set of lines maintained with moderate and stable effective population size (N_e about 50), which continue being maintained under crowded competitive conditions. In these large populations and small lines, we investigate the evolution of the inbreeding load and of the mean for two fitness traits. One trait (noncompetitive pupae productivity *P*, which includes fecundity and egg to pupae viability components of fitness) is measured in noncompetitive conditions under moderate culture density (single pairs per vial). The other trait (competitive productivity *W*, which includes fecundity and egg to adult viability components) provides a measure of competitive productivity relative to that of a marker strain under crowded culture density. The analyses carried out provide the opportunity to assay the consequences of purging operating in competitive conditions for two independent populations (*i.e.*, genetic backgrounds), and to estimate the corresponding purging coefficient *d*. Furthermore, one of these populations analyzed is the same as that previously used by B&GD to estimate purging under noncompetitive conditions, allowing particularly direct inferences on the relevance of competition on inbreeding and purging. In addition, the effective size of the small lines of these experiments is larger than in B&GD, allowing for efficient purging even for mutations with only mild *d* values. On the light of theory, the period analyzed is long enough for purging to have an important effect on the evolution of mean fitness for lines of this effective size.

We show that purging operating in competitive conditions is efficient even against nonlethal deleterious alleles, both for competitive and for noncompetitive productivity traits, and we discuss some consequences of these findings.

Methods

The essential features of the predictive model, experimental design and methods are outlined in the following sections. More details are given in the Supporting Information. Table 1 gives a glossary of terms used in the study.

Models	
IP	Inbreeding-Purging model: it takes into account the consequences of inbreeding and purging
FM	Full-model: it takes into account the consequences of inbreeding, purging, mutation and standard selection
Traits	
ω	Fitness
P	Noncompetitive pupae productivity
W	Competitive productivity
Parameters	
N	Population size
N_e	Effective population size
δ_t	Rate of inbreeding depression used as an estimate of the inbreeding load (in the absence of selection and for independent loci with no epistasis, the rate of inbreeding depression equals the inbreeding load). Subscript *t* stands for generation number and no subscript *t* implies it refers to a base population.
δ_L	Lethal component of the inbreeding load, either for *W* (*i.e.*, overall lethal inbreeding load) or for P (inbreeding load due to lethals expressed in the egg to pupae phase)
δ_{NL}	Component of the inbreeding load not ascribed to lethal alleles
d	Purging coefficient against overall inbreeding load
d_{NL}	Purging coefficient against inbreeding load ascribed to nonlethal alleles
F	Inbreeding coefficient
F_{ST}	Fixation index
g	Purged inbreeding coefficient

Table 1. Glossary of parameters and subscripts used in the paper.

THE INBREEDING –PURGING (IP) MODEL

When a stable population undergoes a reduction in effective size to a new stable value N_e, it experiences an increase in inbreeding that causes inbreeding depression, but also experiences purging. *i.e.*, some increase in the intensity of selection against the (partially) recessive deleterious alleles responsible for inbreeding depression. According to the Inbreeding-Purging (**IP**) approach (García-Dorado 2012), the joint consequences of inbreeding and purging upon a fitness trait ω can be approximately predicted as a function of the purged inbreeding coefficient g, which is equivalent to the standard Wright's inbreeding coefficient (F) corrected for the reduction of the frequency of deleterious alleles induced by purging. In this experiment, two different fitness traits will be used for ω (P and W; see below). The evolution of g depends on the effective population size (N_e) and on a purging coefficient d. Regarding a single locus, d equals the recessive component of the deleterious allele, that is, it equals half the difference between the fitness disadvantage of the homozygotes and twice the disadvantage of the heterozygotes (see Supporting Information). Therefore, it amounts to $d_L = 0.5$ for recessive lethal alleles. This purged inbreeding coefficient can be predicted through generations (t) as

$$g_t \approx \{ 1/(2N_e) + [1 - 1/(2N_e)]\, g_{t-1} \} (1 - 2d\, F_{t-1}) ,$$

where F_t is the standard Wright's inbreeding coefficient, $F_t = 1 - [1 - 1/(2N_e)]^t$. Then, the evolution of average fitness, with initial value ω_0, can be predicted through generations as

$$\omega_t = \omega_0 \exp[-\delta g_t],$$

where δ is the rate at which fitness declines with increasing inbreeding in the absence of

selection and, for independent loci and no epistasis, equals the load concealed in heterozygosis in the initial population, often denoted inbreeding load *B*. The evolution of the inbreeding load can be predicted as

$$\delta_t = \delta g_t (1 - F_t) / F_t.$$

It has been shown through extensive computer simulation, using a large variety of distributions for the deleterious effects, that this **IP** model formulated for a single locus provides a good approximation for overall purging caused by many loci with different fitness effects (García-Dorado 2012). This requires using an effective purging coefficient *d* (purging coefficient hereafter) that applies to the set of loci responsible for inbreeding depression and that can be empirically estimated by fitting fitness data to **IP** predictions. As shown in the Supporting Information, such overall **IP** predictions can be improved by separately accounting for purging against the lethal inbreeding load (δ_L , with purging coefficient $d_L \approx 0.5$) and the nonlethal inbreeding load (δ_{NL}, with overall nonlethal purging coefficient d_{NL}). Our aim is to estimate both d and d_{NL} purging coefficients.

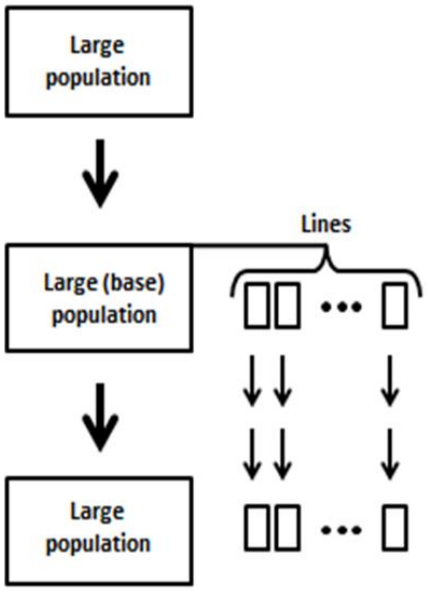
Furthermore, we must note that **IP** predictions are approximations that ignore both the standard nonpurging natural selection that would operate for a noninbred equilibrium population of the same size (standard selection hereafter) and the continuous appearance of deleterious mutations. These two factors determine the inbreeding load under the new mutation-selection-drift (MSD) balance to be attained in the long term for the new population size. The consequences of new deleterious mutation together with standard selection, inbreeding and purging, can be predicted using the Full-Model (**FM**) equations, also provided by García-Dorado (2012). The difference between **IP** and **FM** predictions regarding the evolution of mean for fitness traits in our lines is expected to be negligible, as these lines are not so small that fixation of new deleterious mutations is relevant to

fitness decline, or that the population can retain substantial inbreeding load in the new MSD equilibrium. However, the effective size of our large laboratory populations, although smaller than that of the original wild population, is still substantially large (~1000), so that these populations could harbor a nonnegligible inbreeding load at the MSD balance (δ^*) (García-Dorado *et al*. 2006, Amador *et al*. 2010). Therefore, **FM** predictions for the evolution of the inbreeding load in our large populations (δ_t) could in principle be substantially different from **IP** predictions. Thus, when analyzing the evolution of δ_t in a large population, we will also obtain **FM** predictions. Since, as explained in the Supporting Information, this requires extrapolating the asymptotic inbreeding load from a different laboratory experiment, the d and d_{NL} estimates obtained using the **FM** approach should be interpreted as rough approximations, but are useful to illustrate possible biases derived from the use of the simpler **IP** approach.

BASIC EXPERIMENTAL DESIGN

In what follows we will describe two experiments carried out in two laboratories (Vigo and Madrid) under similar designs using *Drosophila melanogaster* (Fig. 1). For each experiment, a wild population was captured, maintained in the lab with large size for a long period, and used at a given generation to obtain a large set of small populations (lines) that were thereafter maintained synchronously to the large population. In order to evaluate the consequences of inbreeding and purging and to estimate the purging coefficient, we analyzed the evolution of the inbreeding load (δ) and of the mean for two fitness traits in these populations and lines.

For the large laboratory populations, a subscript denotes the number of generations since their capture from the wild. For the lines, the subscript gives the number of generations since they were founded from the large laboratory population. When analyzing the evolution of the inbreeding load in a large population, the base population was assumed to be the original wild one. However, when analyzing the evolution of the lines, the base population was assumed to be the corresponding large laboratory population at the generation in which the lines were derived (Fig. 1). To avoid ambiguity, the inbreeding load in these "base populations" is denoted by δ (or δ_L and δ_{NL}) with no subscript (eqs. S1-S7).



Large population	Vigo	Madrid
Census size	3000	2600
Number of generations	136	123
Generation for the assay of:		
Productivity (*P*)	22,50,103,111	112
Productivity (*W*)	----	83
Lethal analysis	128	57
Lines	**Vigo**	**Madrid**
Number of lines	20	64
Census size per line	100	80
Number of generations	42	40
Generation for the assay of:		
Productivity (*P*)	25	30
Productivity (*W*)	----	10,20,30,40
Lethal analysis	42	----
Microsatellites	5,10,25	10,20

Figure 1. Experimental design. In each of the experiments, a large population was maintained for a long period. At generations 86 (Vigo) and 83 (Madrid), a set of lines of small size was founded from each large population and subsequently maintained. Analyses of egg to pupa noncompetitive productivity (*P*), competitive productivity (*W*) relative to a marker strain, lethal chromosomes and genotyping for microsatellites was carried out at different moments in the large populations or small lines.

ORIGIN AND MAINTENANCE OF POPULATIONS AND LINES

Vigo experiment

A laboratory population was founded from about 1000 females captured in a wine cellar close to Vigo and maintained in 30 bottles. About 50 males and 50 females mated and lay eggs in each bottle per generation. Progeny was sampled from these bottles and mixed according to a circular scheme to produce the next generation, such that the *i*th bottle was formed by about 50 flies from the *i*th bottle and 50 from the *i*th+1 bottle from the previous generation. Thus, the large population was maintained with about 3000 flies per generation.

At generation 86, 1000 males and 1000 females were sampled from the large population to establish 20 lines, each maintained thereafter in a single bottle with exactly 50 male and 50 female parents during 42 generations synchronously to the large population.

Madrid experiment

A population was founded from 276 females captured in Segura Viudas cellar (Penedés) and maintained in similar conditions as in the Vigo experiment, in 32 bottles with 40 males and 40 females per bottle (thus, a total of 2560 flies per generation). At generation 83, 64 lines were founded, and each was thereafter similarly maintained in a single bottle with 40 male and 40 female parents during 40 generations, synchronously to the large population.

FITNESS TRAITS ASSAYED

Noncompetitive pupae productivity P

In both experiments, noncompetitive pupae productivity (*P*) was assayed for single 4-days old females mated in individual vials to single males. *P* was measured as the number of pupae produced in the vial after 11 days. This trait includes egg to pupae viability and female fecundity fitness components, assayed under relatively low density, and in the absence of competition regarding fecundity or mating success.

Competitive productivity W

Competitive productivity (*W*) was assayed for groups of 20 females in Madrid experiment. Previously, groups of four males and four females, all of them four days old, were mated for three days in a single vial. Then groups of 20 inseminated females were placed in a bottle with 20 inseminated females from a curly (*Cy*/*If*) laboratory strain in a single evaluation bottle. Then *W* was computed as the ratio of the number of offspring contributed by the assayed population or line (wild progeny) to the number of offspring contributed by the marker strain plus 1. This trait includes egg to adult viability and female fecundity fitness components, both assayed in crowded competitive conditions. Since each female had ample opportunity of being inseminated before being transferred to the evaluation bottles, *W* does not include competitive components for mating success.

ESTIMATES OF THE EFFECTIVE POPULATION SIZE

Both in Vigo and Madrid experiments, the effective population size (N_e) of the lines was inferred from the evolution of F_{ST} for nine microsatellite loci at several generations (generations 5, 10 and 25 for Vigo lines; 10 and 20 for Madrid lines).

EVALUATION OF THE OVERALL INBREEDING LOAD

The inbreeding load for noncompetitive pupae productivity (P) (*i.e.*, the rate δ of inbreeding depression expected in the absence of selection) was estimated in the large population of Vigo experiment at generations 22, 50, 103 and 111. In each of these generations (t), average P was estimated using outbred (P_O) and inbred (P_I) individuals, as described in the Supporting Information, so that $\delta_t = \ln(P_O / P_I) / F_t$.

An analogous evaluation was made at generation 112 in the large population of Madrid experiment. The inbreeding load for W was estimated in the large Madrid population at generation 83 (*i.e.*, in the base population of the lines).

ESTIMATION OF THE LETHAL AND NONLETHAL COMPONENTS OF THE INBREEDING LOAD

Vigo experiment

At generation 128, 549 chromosomes II were sampled from the large Vigo population and tested for lethality using a classical design with the *Cy/If* marker strain, in order to estimate the proportion of lethal chromosomes II. As explained in the Supporting Information, this allows to estimate the lethal component of the inbreeding load for P (*i.e.*, the δ_L inbreeding load ascribed to alleles that are lethal during the egg to pupae phase). Synchronously, three randomly selected lines were assayed in a similar way (149, 169 and 166 chromosomes, respectively).

In addition, during this test we registered the ratio of the number of wild (+/+) to Curly (*Cy*/+) offspring in the vials corresponding to nonlethal chromosomes, which measures the mean fitness of nonlethal wild chromosomes II in homozygosis relative to that of *Cy*/+ heterozygous individuals.

Madrid experiment

At generation 57 after the capture of the Madrid laboratory population, 447 chromosomes II were sampled from the large population and tested for lethality using the same protocol as in Vigo experiment. This allows to estimate the lethal inbreeding load (δ_L), both for W (*i.e.*, the overall lethal inbreeding load) and for P (*i.e.*, the inbreeding load ascribed to alleles that are lethal during the egg to pupae phase). The overall lethal inbreeding load at generation $t = 57$, both for W and P, were used as proxies for those at generation 83 corresponding to the base population of the lines.

In both experiments, the nonlethal component of the inbreeding load for P or W at any generation was obtained by subtraction ($\delta_{NL} = \delta - \delta_L$), using the appropriate lethal inbreeding load for each trait.

EVALUATION OF THE FITNESS DECLINE IN THE LINES

Vigo experiment

Noncompetitive pupae productivity (P) was synchronously evaluated for the large population and the lines 25 generations after their foundation. Both for the population and for each line, sampled individuals were randomly mated in single pair vials for three generations and P was assayed for the last two generations (assays 1 and 2).

Madrid experiment

Noncompetitive pupae productivity (P) was assayed at generation 30 for the lines and, synchronously, at generation 113 for the large population.

Competitive productivity (W) was assayed in each line at generations 10, 20, 30 and 40. In each case, it was synchronously assayed in the large population, which was

used as a control. In each of these four assays, the mean for competitive productivity W is given as the ratio of the mean of the lines to the synchronous estimate in the large population.

INFERENCE OF THE PURGING COEFFICIENTS

Inference for trait P, obtained from the evolution of δ_t in the large Vigo population

We computed **IP** predictions for the evolution of the inbreeding load δ_t (eq. S4) for a grid of δ (*i.e.*, the initial inbreeding load) and d values (see Supporting Information). From this grid, we obtained the joint Least Square (LS) estimates for these two parameters that better fit the observed evolution of δ_t. The procedure was repeated by accounting separately for lethal and nonlethal depression (eq. S5), in order to obtain estimates of the nonlethal purging coefficient d_{NL}.

In addition, using the previous estimates of the initial inbreeding load, and additional assumptions on its asymptotic value (δ^*; see Supporting Information), rough LS estimates for the purging coefficients d and d_{NL} were obtained in a similar way using the **FM** approach (eqs. S6 and S7).

Statistical contrasts and confidence intervals (CI) for the estimates of d were obtained using the F distributed statistic derived from the likelihood ratio test (Casella and Berger 2001). This gives only approximate results, due to the limited number of δ_t values and to the likely departures from normality for their sampling errors.

Inference for trait W, obtained from the evolution of the trait's mean in Madrid lines

A LS estimate for the overall purging coefficient *d* was obtained by computing **IP** predictions (eq. S1 for *W*, with $W_0 = 1$) for a grid of *d* values, searching for the *d* estimate that produced the best fitting between the mean relative *W* observed in generations 10, 20, 30 and 40 and the corresponding **IP** predictions. Similarly, a LS estimate was obtained for the nonlethal purging coefficient, d_{NL}, by fitting observed values of relative *W* to predictions separately accounting for purging against the lethal and nonlethal inbreeding load (eq. S3).

Statistical contrasts and approximate confidence intervals for the estimates of *d* were again performed using the F statistic derived from the likelihood ratio test.

Results

THE EFFECTIVE SIZE OF THE LINES AND OF THE LARGE POPULATIONS

Table 2 gives the effective population size estimated from microsatellite analysis in Vigo and Madrid lines. We estimated the per line effective population size under the maintenance conditions of the lines (one bottle per line) as the average of the three estimates in the case of Vigo and of the two estimates in Madrid experiment, which gives effective population sizes of $N_e = 52$ and 43, respectively. This gives $N_e \approx 1000$ in the large Vigo population (maintained with 20 bottles and ~3000 individuals), and $N_e \approx 1376$ in the large Madrid population (maintained with 32 bottles and 2560 individuals).

	Generation	N_e	Confidence limits
Vigo	5	59	44 - 88
	10	34	23 - 50
	25	63	57 - 69
Madrid	10	45	40 - 51
	20	42	32 - 54

Table 2. Estimates of the effective population size (Ne) in Vigo (N = 100) and Madrid (N = 80) lines, obtained from microsatellite data at different generations since the start of the lines, and their 95% confidence bootstrap limits.

PURGING FOR NONCOMPETITIVE PUPAE PRODUCTIVITY *P*

The lethal Inbreeding load in Vigo experiment

Ninety-six lethal chromosomes II were detected out of 549 chromosomes sampled from the large laboratory population at generation 128, which gives a lethal inbreeding load $\delta_{L128} = 0.316$ for *P* (see Supporting Information).

Analogously, the synchronous estimates of the lethal component of the inbreeding load for *P* in the three lines analyzed ($t = 42$) were 0.091, 0.142 and 0.200. Thus, the average lethal inbreeding load in the lines (0.14) was about half that of the large population.

In addition, the mean relative fitness for nonlethal chromosomes II in homozygosis, estimated in the large population during the lethal analysis, was 0.442 ± 0.009, significantly smaller than that of the lines (0.485 ± 0.009, $P < 0.00018$). This implies that, excluding lethal alleles, the fitness of homozygous chromosomes II was about 10% larger in the lines than in the large population, which should be ascribed to purging in the lines.

Evolution of the Inbreeding load for P in the large Vigo population: Evidence of purging

Figure 2 gives the estimates obtained for the overall inbreeding load for noncompetitive pupae productivity (*P*) in the large Vigo population, plotted against generation number *t*. The observed decline of δ_t was practically linear on *t* so that, at any generation within the time interval corresponding to these estimates, the expected rate of inbreeding depression in Vigo large population could reasonably be inferred using the linear regression of δ_t on *t* estimated from these observations. This gives

$$E(\delta_t) = 2.04520 - 0.01308\ t, \qquad (1)$$

with standard errors 0.03579 and 0.00044 for the intercept and the slope, respectively.

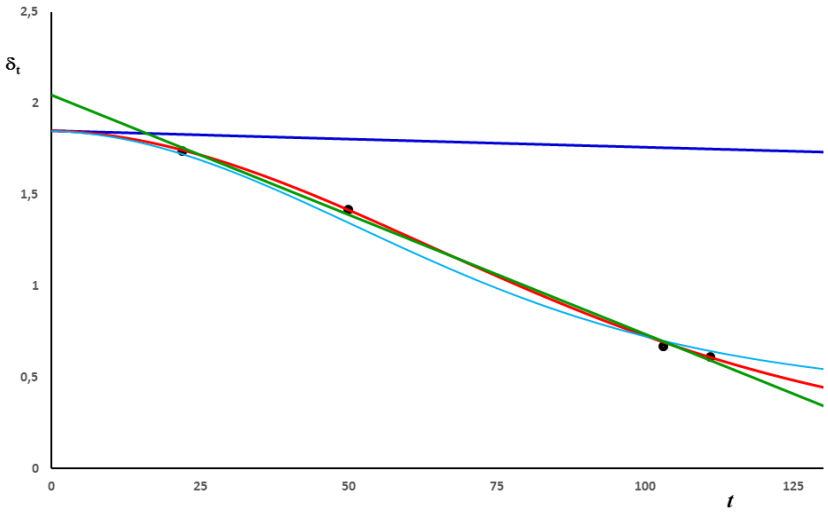


Figure 2. Evolution of the inbreeding load for *P* in Vigo large laboratory population. Dots: experimental estimates; green: linear regression fitting experimental estimates; dark blue: neutral prediction; red: Inbreeding-Purging (**IP**) prediction considering overall purging upon lethal and nonlethal components; light blue: Full-Model (**FM**) prediction also based on overall purging (**IP** and **FM** predictions obtained by separately accounting for purging against the lethal and nonlethal inbreeding load are not shown but virtually overlap the corresponding **IP** and **FM** predictions plotted here).

Figure 2 also shows different predictions for the evolution of δ_t, computed using our estimate of the effective population size (N_e = 1000). It shows that, for N_e = 1000 and in the absence of purging (d = 0), δ_t is expected to decline almost linearly on t during all the experiment with a small slope $\sim\delta/2N_e = -9 \times 10^{-4}$ (where δ is the initial inbreeding load). However, the estimated δ_t declined much faster, the slope (–0.01308 ± 0.000445) being significantly larger than that expected in the absence of selection ($p < 6.6 \times 10^{-4}$), which implies substantial purging.

Inference of the purging coefficients for P in the large Vigo population

The estimates of purging coefficients obtained by LS fitting of the evolution of δ_t to **IP** predictions are given in Table 3. Using overall **IP** predictions (eq. S4), the LS estimate of the initial inbreeding load is δ = 1.85 and that for the purging coefficient is d = 0.30, with a narrow approximate 95% CI (0.28-0.33). **IP** predictions using this estimate fit very well the observed values (Fig. 2).

Predictions were also obtained by separately accounting for the lethal and nonlethal components of the inbreeding load for *P*. Since we do not have an estimate of the initial lethal inbreeding load in Vigo population, we assumed that, based on consistent empirical evidence for *Drosophila* viability (Simmons and Crow 1977), 50% of the inbreeding load in the original wild population was due to lethal alleles. LS estimates fitting the **IP** model (eq. S5) give d_{NL} = 0.19, and the same initial inbreeding load as before (δ = 1.85; *i.e.*, δ_L = 0.925, δ_{NL} = 0.925). This gives predictions that virtually overlap those computed in terms of overall d and are therefore not shown in Figure 2.

Trait	Purging coefficient	Estimate (95% CI)	Experiment	Data	Generations assayed	Effective size	Initial δ	Assumptions
P	d	0.30 (0.28;0.33)	Vigo	δ_t decline (large population)	22, 50, 103, 111	1000[a]	1.85[b]	
P	d_{NL}	0.19 (0.14;0.26)	Vigo	δ_t decline (large population)	22, 50, 103, 111	1000 [a]	1.85[b]	Initially $\delta_{L(P)} = \delta_{(P)}/2$
W	d	0.27 (0.08;0.5)	Madrid	W in the lines (*vs* large population)	10, 20, 30, 40	43 [a]	2.884	
W	d_{NL}	0.24 (0.06;0.5)	Madrid	W in the lines (vs large population)	10, 20, 30, 40	43 [a]	2.884 [a]	In the base population of the lines δ_L = 0.441[c]

Table 3: IP estimates of the purging coefficient for noncompetitive (*P*) and competitive (*W*) productivity.

[a] Independently estimated in this experiment

[b] This is the estimate of the initial δ obtained by least square simultaneously to that of d

[c] This a.ssumes that δ_L in the large population at t = 83 (lines foundation) is approximately the estimate obtained at generation t = 57

Figure 2 also shows the **FM** prediction (eq. S6), computed using the corresponding LS estimate of the overall effective purging coefficient d = 0.47 (95% CI 0.34-0.50). Accounting separately for purging against the lethal and nonlethal inbreeding load of P, this model gives d_{NL} = 0.44 (95% CI 0.25-0.50; predictions not shown as they virtually overlap overall **FM** predictions for d = 0.47). These large **FM** predictions rely on extrapolations regarding the inbreeding load expected in the long term at the new MSD balance: (δ^* = 0.33, with δ_{NL}^* = 0.15 and δ_L^* = 0.18; see Supporting Information). Therefore, they should be taken with caution, although it is worth noticing that the **FM** approach predicts $\delta_{L(P)128}$ = 0.28, which is consistent with the estimate of the lethal inbreeding load obtained at generation 128 from the lethal analysis (0.316). In any case, these **FM** estimates illustrate that purging coefficients estimated by fitting the decline of δ_t to **IP** predictions can be considered conservatively low.

Inference of the initial inbreeding load for P in Vigo lines

Based on the remarkable linearity for the decline of δ_t observed in the large Vigo population, the inbreeding load at t = 86 (when the lines were derived) can be reasonably inferred using the estimated linear regression (eq. (1)), which gives E(δ_{86}) = 0.92. Therefore, we will use δ = 0.92 as the inbreeding load in the base population of Vigo lines. Its lethal component was predicted using the **FM** approach, also for the large population at generation 86 as shown in the Supporting Information, and the nonlethal one can be obtained by subtraction. These values are given in Table 4.

	δ^a	$\delta_L{}^b$	$\delta_{NL}{}^c$
Noncompetitive pupae productivity (*P*)			
Vigo	0.920	0.425	0.495
Madrid	1.402	0.333	1.070
Competitive productivity (*W*)			
Madrid	2.884	0.441	2.443
	(0.696)	(0.044)	(0.698)

Table 4: Inbreeding load for noncompetitive productivity and for competitive productivity in the base populations of Vigo and Madrid lines.

[a] Overall inbreeding load; [b] lethal inbreeding load; [c] nonlethal inbreeding load. Standard errors derived from bootstrap analyses are given in parenthesis.

The decline of noncompetitive pupae productivity (P) in Vigo lines

Means for noncompetitive pupae productivity (*P*), synchronously evaluated for the large population and the lines in samples obtained 25 generations after the lines were founded, are given in Table 5, which also shows predictions for the mean of the lines computed using the synchronous mean of the large population as a noninbred control and assuming δ = 0.920, δ_{NL} = 0.495 (Table 4) and N_e = 52. In these lines, both the consequences of standard selection and the decline expected from fixation of new deleterious mutation by generation 25 should be negligible, due to their modest effective size. Therefore, the **IP** approach is expected to give satisfactory predictions for the decline of *P*, so that **FM** predictions are not discussed.

All the lines survived through the whole experiment and, on the average, they showed no decline for *P* compared to the large population. The table also shows that the mean productivity *P* of the lines was highly significantly larger than neutral predictions (those computed ignoring purging) and than **IP** predictions computed using the purging coefficients formerly estimated by B&GD under noncompetitive conditions. On the

Table 5. Mean for noncompetitive pupae productivity *P* (± standard error) in the two assays of Vigo lines at generation 25, and in one assay in Madrid lines at generation 30, and synchronous mean for their corresponding large population. Neutral predictions (ignoring purging) and **IP** expectations are also given (see text for explanations).

	Observed	Observed	Neutral prediction	**IP** predictions for lines average *P* obtained using *d* or d_{NL}, either estimated by B&GD or in Vigo experiment			
Assay	Large population	Lines	Lines	$d = 0.09$ [a] B&GD	$d = 0.3$ [b] Vigo	$d_{NL} = 0.08$ [c] B&GD	$d_{NL} = 0.19$ [d] Vigo
Vigo 1	79.23 (±2.14)	78.63 (±1.37)	65.03***	68.61***	73.68 [ns]	72.34**	73.90 [ns]
Vigo 2	87.30 (±2.08)	89.08 (±1.87)	71.66***	75.60***	81.19**	79.71***	81.43**
Madrid	81.14 (±1.99)	66.92 (±2.27)	56.84***	61.18 *	73.51 *	65.70 [ns]	70.50 [ns]

*** stands for p< 0.001, ** for p<0.01, * for p<0.05 and ns for nonsignificant (always for the tests about whether the mean observed in the lines is larger than expected with no purging or with the purging coefficients estimated in B&GD, or is different than predicted using the purging coefficients estimated here from the decline in δ_t in the large Vigo population)

[a] Bersabé & García-Dorado estimate for *d* for noncompetitive conditions

[b] our **IP** estimate for overall *d*

[c] Bersabé & García-Dorado inference for the upper bound of d_{NL} for noncompetitive conditions

[d] our **IP** estimate of d_{NL}

contrary, the average productivity P of the lines in the first assay was not significantly different from the **IP** predictions computed with any of the purging coefficients estimated from the decline of δ_t in the large population of this experiment, although it was significantly larger in the second assay (in the edge of significance when using the **FM** estimates of d or d_{NL}).

Overall, these results show that purging has completely erased the negative impact of inbreeding depression on P in Vigo lines. They are in agreement with purging coefficients larger than B&GD estimates obtained for purging under noncompetitive conditions, and on the order of those estimated from the decline of δ_t in the Vigo population maintained under competitive conditions, and they suggest that d and d_{NL} estimates were conservative when obtained under the **IP** approach.

Inference of the initial inbreeding load for P in Madrid lines

The inbreeding load for noncompetitive productivity, estimated at generation 112 in the large Madrid population, was 0.848 (with bootstrap error 0.142), but we do not have an estimate obtained at the time the lines were founded. Due to the similitude between this design and that of Vigo experiment, it seems reasonable to consider that the inbreeding load for productivity should decline in Madrid large population roughly at the same rate as in Vigo experiment. Therefore, from the linear regression of δ_t on generation number estimated in Vigo experiment (eq. 1), δ_t would be expected to drop from generation 83 to generation 112 by a 0.604 factor. Thus, for the base population of Madrid lines, we could infer $\delta \approx 0.8483 / 0.6040 = 1.402$ (Table 4). We used the lethal component of the inbreeding load obtained for P at generation 57 in the large Madrid population ($\delta_{L57} = 0.333$) as an approximation for that of generation 83, although this can induce some underestimation of the nonlethal inbreeding load and of the corresponding purging coefficients. Thus, in the base population of Madrid lines, we obtained $\delta_L = 0.333$ and $\delta_{NL} = \delta - \delta_L = 1.070$ (Table 4).

The decline of noncompetitive pupae productivity (P) in Madrid lines

Madrid lines were assayed for productivity at generation 30, synchronously to the control. Means are given in Table 5. Predictions were computed assuming $N_e = 43$ and the initial inbreeding load of *P* inferred for Madrid lines ($\delta = 1.403$, $\delta_{NL} = 1.070$; Table 4). The mean for *P* in Madrid lines was larger than the neutral expectation ($p < 5 \times 10^{-6}$) and also larger than predictions computed using B&GD estimates of the purging coefficient, nonsignificantly when using the upper bound $d_{NL} = 0.08$, although significantly when using the corresponding point estimate $d_{NL} = 0.02$ ($p < 0.017$; not shown in the Table). The mean for *P* in the lines was nonsignificantly different from predictions computed accounting separately for purging against lethal and nonlethal loads by using our Vigo estimate $d_{NL} = 0.19$, but was smaller than predictions computed considering overall purging ($d = 0.3$) or using purging coefficients estimated under the **FM** approach (not shown).

Overall, the decline for average *P* in Madrid lines is more consistent with the purging coefficients estimated for the Vigo population, maintained in similar competitive conditions, than with the smaller purging coefficients estimated by B&GD in lines from the same Madrid population but maintained in noncompetitive conditions.

PURGING FOR COMPETITIVE PRODUCTIVITY *W*

Initial inbreeding load for W in Madrid lines

The estimate of the inbreeding load for *W* in the large Madrid population was directly obtained at generation $t = 83$, when the lines were derived. Thus, this estimate ($\delta_{83} = 2.884 \pm 0.696$, see Table 4) is the inbreeding load at the base population of Madrid lines. Since, out of the 447 chromosomes II assayed at generation 57, 82 were lethal, the lethal inbreeding load was $\delta_{L57} = 0.44$. Again, we used this estimate as an approximation of that for $t = 83$ which, as explained above, implies a conservatively low estimate for d_{NL}. Therefore, for the base population of the lines, the

estimates of the inbreeding load for W were $\delta = 2.88$, $\delta_L = 0.44$ and $\delta_{NL} = \delta - \delta_L \approx 2.44$ (Table 4).

Evolution of the mean and estimation of the purging coefficients for W in Madrid lines.

Figure 3 gives mean competitive productivity for the lines, relative to the synchronous estimate in the large laboratory population. None of the lines was extinct through the experiment. The figure shows neutral predictions obtained ignoring purging (solid line) using $N_e = 43$ and $\delta = 2.884$. The figure also shows **IP** predictions (dashed line) obtained by separately considering purging against the lethal and nonlethal fractions of the inbreeding load ($\delta_L = 0.441$ and $\delta_{NL} = 2.443$, respectively). They are computed using the LS estimate $d_{NL} = 0.24$ obtained from these data. This d_{NL} estimate would be significantly larger than 0.06 ($p < 0.05$). The LS estimate of the overall effective purging coefficient was $d = 0.27$. This approach fits the data almost as well as the previous one, and produces predictions (not shown) that virtually overlap those shown in Figure 3 for $d_{NL} = 0.24$. This $d = 0.27$ estimate would be significantly larger than 0.09 ($p < 0.05$). Therefore, purging coefficients in Madrid lines were larger than those estimated by B&GD for smaller lines maintained under noncompetitive conditions, and are within the range of estimates obtained for P in Vigo lines.

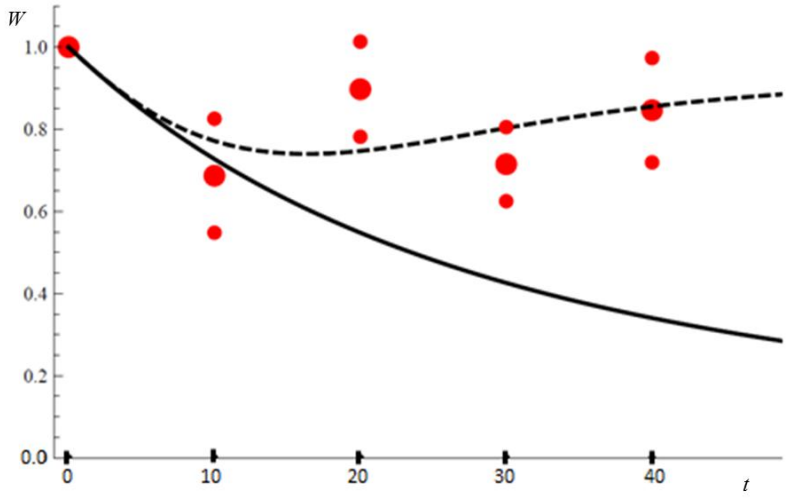


Figure 3. Mean of the averages of competitive productivity, *W*, relative to the control population against generation number for Madrid lines. Large dots give observed means; small dots delimit one standard error intervals. Solid line: neutral prediction. Dashed line: inbreeding-purging (**IP**) prediction obtained by separately accounting for purging against the lethal (δ_L = 0.4406, d_L = 0.5) and nonlethal (δ_{NL} = 2.4434, d_{NL} = 0.24) inbreeding load.

Discussion

INBREEDING DEPRESSION IN COMPETITIVE *vs* NONCOMPETITIVE CONDITIONS

We have analyzed the consequences of purging on the evolution of the inbreeding load and of the mean of two fitness traits (*P* and *W*) in two different *Drosophila* populations maintained under crowded competitive conditions. We will begin considering whether the inbreeding load (δ) for these two traits is representative of that for overall fitness under captive or wild-like competitive conditions. In Madrid large population, 83 generations after its capture, the δ value for competitive productivity *W* was twice that inferred for noncompetitive one *P* (Table 4). In

principle, this could be partly ascribed to inbreeding depression for pupae to adult viability, as this fitness component is included in W but not in P. However, δ had been estimated for this same population just after its capture, both for egg to pupae viability (EPV) and for egg to adult viability (EAV), and the two estimates were quite similar ($\delta = 1.6$ and $\delta = 1.8$, respectively; the first estimate reported by B&GD while the second one is a personal communication of the authors). This implies that the rate of inbreeding depression for pupae to adult viability was small, so that the larger inbreeding load for W should be ascribed to the competitive nature of this trait, in agreement with previous experimental evidence (Yun and Agrawal 2014).

Furthermore, the rate of decline for δ_t observed in Figure 2 suggests that the overall inbreeding load in Madrid large population could have been about twice when it was captured than when estimated at generation 83. In fact, if δ_t was reduced in Madrid population by the same factor as in Vigo one (both for P and W), equation (1) would imply an initial inbreeding load $\delta = 2.99$ for P and $\delta = 6.15$ for W in Madrid large population, both values well above the initial estimate reported by B&GD for EPV in the same population ($\delta = 1.6$). This implies that P and W are more comprehensive fitness measures than EPV. Our estimate $\delta \approx 6$ for competitive productivity is in agreement with the values reported for fitness in the wild (O'Grady *et al*. 2006), suggesting that high competition is a main determinant of the larger inbreeding load found in the wild, compared to captive conditions.

EVIDENCE FOR PURGING

Next we will consider the evidences for purging in our populations and lines maintained in competitive conditions, and the corresponding estimates of the purging coefficients, and we will compare these results with those reported by B&GD in a different set of lines of the same Madrid population. These authors obtained $d \approx 0.09$ and, taking into account the small effective size of their lines ($N_e = 6$ or 12), they concluded that the purging coefficient against nonlethal alleles

should have been in the range $0.02 < d_{NL} < 0.08$. However, in B&GD experiment, lines were maintained by mating individual couples in separate, relatively large vials. Then, the expected number of offspring contributed by each single mated couple was made proportional to its pupae productivity. Thus, in B&GD experiment, purging acted upon pupae productivity under noncompetitive conditions.

Purging for noncompetitive pupae productivity *P* was inferred in Vigo experiment from the evolution of δ in the large population, which declined almost linearly through the experiment (Fig. 2). The slope of this decline is not consistent with the neutral expectation, but agrees with inbreeding-purging predictions (**IP**) for purging coefficient $d \approx 0.3$ against the pool of nonlethal and recessive lethal alleles. The purging coefficient against nonlethal alleles was estimated by assuming that these alleles contributed half the initial inbreeding load. This supposition is based on the empirical observation, consistent through *Drosophila* literature, that lethal inbreeding load usually accounts for 40% - 50% of the overall viability inbreeding load of wild populations (Simmons and Crow 1977; Bersabé and García-Dorado 2013). In our case, the actual contribution of δ_L to initial δ could have been somewhat below 50%, both because this value is in the upper end of the observed range and because our trait *P* includes fecundity components, while the lethal inbreeding load estimated is ascribed just to alleles that have lethal effects on viability. However, if δ_L contributed less than 50% to the initial overall δ, by assuming a 50% contribution we will underestimate δ_{NL} and d_{NL}. Therefore, we consider that our $d_{NL} \approx 0.2$ value is a conservative estimate. Thus, it can be concluded that purging coefficients estimated in Vigo experiment were larger than those previously reported by B&GD.

Considerably larger estimates were obtained under the **FM** approach (d =0.47; d_{NL} =0.44), which takes into account standard selection and continuous mutation, asymptotically leading to a new equilibrium with nonnull inbreeding load. These **FM** estimates involve extrapolations

regarding the amount of inbreeding load expected in the long term, and, in the case of the estimate of d_{NL}, regarding the lethal and nonlethal components of that load. However, these **FM** estimates illustrate that the estimates of d and d_{NL} obtained using the **IP** model are expected to be biased downwards. In contrast to **FM** estimates, the **IP** estimate of d involves no extrapolation of parameters, and that for d_{NL} only involves the conservative assumption that δ accounts for 50% of the initial load. Therefore, the **IP** estimates of d and of d_{NL} obtained in Vigo experiment for P can be considered reliable and conservative estimates of the corresponding purging coefficients.

Furthermore, after 25 generations with N_e = 52, average P in Vigo lines was larger than predicted by ignoring purging or by assuming B&GD purging coefficients (Table 5). In the first evaluation (assay 1), this mean was close to the predictions computed using our **IP** estimates of d and d_{NL}. In the second evaluation (assay 2) it was even larger, suggesting purging coefficients closer to the **FM** estimates. However, these results should be taken with caution, as the predictions for average P in the lines rely on the inference of the initial δ of the lines. Furthermore, the estimate of d_{NL} also depends on using the estimate obtained at generation 57 for the δ_L value at generation 83, although this is expected to underestimate d_{NL} so that it can be considered a conservative decision. In any case, the evolution of the mean for noncompetitive pupae productivity in Vigo lines is consistent with the purging coefficients estimated from the decline of δ_t in the large population since its capture from the wild. This is so despite the 86 generations elapsed between the capture of the large populations and the foundation of the lines. Thus, during those 86 generations δ was roughly halved due to slow purging (Fig. 2), but the purging coefficient was not substantially reduced.

As an additional proof of efficient purging against nonlethal alleles in Vigo lines, the fitness of nonlethal chromosomes II in homozygosis in these lines at generation 42 was about 10% larger than in the synchronous large population. Thus, the fitness of an individual homozygote for the

whole autosomal genome but carrying no lethal alleles, would be expected to be about 21% higher in the lines than in the large population. This should be ascribed to purging which, during this period, was more efficient in reducing the average frequencies of (partially) recessive deleterious alleles in the lines than in the large laboratory population (Glémin 2003), although it would be expected to be more efficient preventing long term decline for outbred fitness in the large population (Wang *et al*. 1999; García-Dorado 2012; Bersabé and García-Dorado 2013).

The mean productivity P of Madrid lines, maintained with N_e = 43 and evaluated at generation 30, showed a relatively larger decline, but was still more consistent with purging coefficients on the order of those estimated for this trait in Vigo experiment than with predictions assuming no purging or based on B&GD estimates obtained for the Madrid population in noncompetitive conditions.

Purging for competitive productivity W was inferred in Madrid experiment from the evolution of mean W in the small lines (N_e = 43) over 40 generations, compared to the large population that is used as a control. The evolution of mean W was, again, inconsistent with neutral predictions, and implied important purging coefficients (d = 0.27, d_{NL} = 0.24) on the order of those estimated for P in Vigo experiment and larger than those estimated by B&GD for the same genetic background but under noncompetitive conditions (Fig. 3).

It must be noted that all d estimates have been obtained using our estimated effective population sizes as if they were known values. Furthermore, the estimate of d obtained for W was conditional to the δ value previously estimated in the base population. Thus, although our estimates of d are conservatively small, the corresponding confidence intervals we report are conditional to our estimates of the effective population size (and, in the case of W, on the estimate of the initial δ), and unconditional confidence intervals should be somewhat larger. Even so, our estimates consistently indicate that purging was a very important force reducing the inbreeding load and the

depression of mean fitness in competitive conditions.

The efficiency of purging in the present experiments could have been larger than in B&GD because alleles with $d < 0.08$ could have escaped purging in B&GD due to the smaller N_e and larger drift in that experiment. However, this seems unlikely, as the nonlethal effective purging coefficient estimated from the decline of δ_t in the large Vigo population was $d_{NL} \approx 0.2$ and the overall purging coefficient was $d \approx 0.3$, for both fitness traits and for both small and moderate population sizes. Furthermore, no decline of mean *P* was observed in Vigo lines, which suggests that most inbreeding load was due to deleterious alleles with individual purging coefficient such that $N_e d > 1$ (*i.e.*, $d > 0.02$).

It is also worth noticing that, in B&GD, **IP** predictions for the evolution of the mean computed by considering overall purging, fitted the data worse than those computed by taking separately into account purging of lethal and nonlethal alleles. However, our data for mean *W* fitted both predictions similarly. This should be due to a smaller contribution of recessive lethals to overall δ for *W* than for viability, partly because the estimate of δ_L includes just alleles with lethal effect on viability, while δ_{NL} includes inbreeding load from most fitness components, excluding mating success. Therefore, the proportional contribution of recessive lethal alleles to the inbreeding load for competitive *W* is small, *d* and d_{NL} are relatively similar, and the data fit similarly both **IP** predictions. This supports the use of the simpler **IP** method, based on overall δ and *d* values, as a reasonable approximation when dealing with fitness measures that are more comprehensive than viability and are assayed at competitive conditions.

Overall, our results imply that, in more crowded (wild-like) conditions, inbreeding depression was larger, but purging was also more efficient. These results are in agreement with experimental evidence, showing that deleterious mutations arisen in mutation accumulation experiments have larger average effects on competitive fitness than on noncompetitive viability,

and can be efficiently purged under competitive conditions in populations with effective size similar to that of the lines of the present experiments (Avila and García-Dorado 2002).

Interestingly, we found that purging in competitive conditions is efficient against inbreeding depression expressed both in competitive (*W*) and noncompetitive (*P*) conditions. Thus, the larger inbreeding load estimated in more competitive conditions should be ascribed, to a good extent, to the same deleterious alleles as in noncompetitive conditions but with larger effects, rather than to a different genetic basis. This is in apparent contradiction with the common notion that adaptation to captive conditions should entail some misadaptation in the wild, due to adaptive tradeoffs (Agrawal *et al*. 2010, Woodworth *et al*. 2002, Frankham 2008). However, our data are not inconsistent with current views. First, tradeoffs between adaptations to alternative environments are not ubiquitous. To take an example particularly relevant to our case, they were not detected in *Drosophila* populations maintained under different crowding conditions (Sánchez-Molano and García-Dorado 2011). Second, a small fraction of the alleles responsible for inbreeding load can have opposite fitness effects under competitive and noncompetitive conditions. In this case, a change in the competitive conditions would promote an increase in the frequency of rare alleles prompting adaptation to the new situation. As the frequency of those alleles increases, they will make a much larger contribution to the fitness additive variance than the bulk of rare alleles that are unconditionally deleterious, leading to a transitory negative correlation between fitness in competitive and noncompetitive conditions. Third, alleles with opposing fitness effects under different competitive conditions can show no associations between the sign of their effect and that of their dominance, and therefore could not contribute to inbreeding depression. Furthermore, these alleles could determine most genetic correlation because they can be relatively common, due to some kind of environmental heterogeneity for competitive conditions (Agrawal 2010). Thus, our results are not in contradiction with previous experimental data, and can be consistent with the existence of adaptive tradeoffs caused by alleles with fitness effect of opposite sign under different

competitive conditions.

The larger *d* value detected for purging in competitive conditions, compared to noncompetitive estimates by B&GD, might partly be explained if alleles responsible for the inbreeding depression of EPV had smaller *d* than those responsible for the inbreeding depression of fecundity components included in *P* and *W* but not in EPV. However, empirical evidence indicates that fecundity traits show small inbreeding depression in *Drosophila* (Fernández *et al*. 2003). Thus, most likely, alleles determining inbreeding depression for noncompetitive EPV have larger *d* values productivity under competitive conditions. This may occur because some deleterious alleles have a larger effect on competitive than on noncompetitive viability. In addition, it may occur because some alleles that are deleterious for viability have larger pleiotropic side effects on fecundity under competitive than under noncompetitive conditions, and they could also have pleiotropic side effects on mating ability that would scape purging in the noncompetitive B&GD experiment. Disentangling these possible direct and pleiotropic effects would be relevant to evolutionary issues. For example, the existence of pleiotropic side effects on mating success for alleles affecting viability would be useful to assess the role of the "good genes" hypothesis in sexual selection (Agrawal 2001; Siller 2001; Lumley *et al.* 2015). Our results imply that alleles responsible for inbreeding depression of noncompetitive viability have larger concealed deleterious effects on competitive fitness (larger *d*) but, unfortunately, they do not allow to ascertain whether this is due to larger effect on competitive viability or to larger pleiotropic side effects for other fitness components.

Summarizing, our experiments reveal that the inbreeding load is larger for competitive than for noncompetitive fitness measures, and that most inbreeding load is due to alleles that are deleterious both in competitive and noncompetitive conditions. However, we also find that purging is much more efficient when operating in competitive conditions. In the case of our lines, with

effective population sizes about 40-50, that efficiency halts or reverts fitness decline up to values very close to those of the original population. The evolution of the inbreeding load and of the mean imply that purging is efficient against nonlethal alleles, with $d_{NL} \approx 0.2$, but they reasonably fit simple **IP** predictions computed considering purging upon the overall set of lethal and nonlethal alleles with purging coefficients $d \approx 0.3$. Furthermore, slow purging can cause considerable depletion of the inbreeding load with little reduction of the purging coefficient, so that the potential for purging in the future should not be assumed to be irrelevant just because a population had a recent history of moderate demographic decline.

IMPLICATIONS FOR CONSERVATION AND EVOLUTION

The results obtained here are relevant to conservation practice. For example, it has been recently suggested that the genetic rule of thumb for the minimum viable population size should be doubled from 50 to 100 in order to prevent inbreeding depression (Franklin 1980; Hedrick and Kalinowsky 2000; ; Jamieson and Allendorf 2012; Frankham *et al*. 2014; Franklin *et al*. 2014). However, our results support the view that the $N_e = 50$ rule remains appropriate as far as the initial reproductive potential is not too low (García-Dorado 2015). Thus, although larger population sizes should always be intended, our results support conservation efforts even in small populations, and emphasize the convenience of breeding endangered populations in competitive conditions that are similar to those found in the wild, encouraging the intensification of *in situ* conservation.

From an evolutionary point of view, the remarkable efficiency of purging detected in our experiments suggests that the mutational load in sexual populations of small or moderate size can be substantially smaller than in asexual ones, which can account for part of the proposed advantages of sexual reproduction (Haag and Roze 2007). The same phenomenon can be expected when inbreeding is caused by spatial population structure or by breeding strategies that impose some restriction to panmixia, with the advantage that these situations would not necessarily induce

drift (Agrawal and Chasnov 2001; Roze and Rousset 2004; Ávila *et al*. 2010). Furthermore, our results imply that inbreeding depression in competitive conditions is largely due to alleles with a large recessive deleterious component (*i.e.*, large *d* value that favors purging), in agreement with the estimates of the average deleterious effects and coefficient of dominance of deleterious mutations (García-Dorado and Caballero 2000; García-Dorado *et al*. 2004). This means that each individual may carry not too many (partially) recessive deleterious alleles, an scenario that involves weak linkage and is therefore favorable to explain the evolution of diploidy, since an allele extending the diploid phase will enjoy some advantage from the masking of recessive deleterious effects without becoming increasingly loaded with too many linked deleterious alleles (Otto and Gerstein 2008). In summary, our results show that purging needs to be considered as an important factor for the evolution of main biological properties, as diploidy, sex, population structure or breeding strategies, as well as for the persistency of endangered populations.

Acknowledgements

We thank Carlos López-Fanjul for helpful discussion, Diego Bersabé and Daniel Kleinman for their support in estimating the lethal rate of inbreeding depression for Madrid experiment, Raquel Sampedro and Nieves Santamaría for technical assistance in Vigo experiment, and two anonymous referees and the Editor Denis Roze for useful comments. E. López-Cortegano is currently funded by a FPI research fellowship (BES-2012-055006) and A. Vilas was funded by a FPI research fellowship (BES-2010-032537) from Ministerio de Economía y Competitividad (MINECO, Spain). This work was founded by grants CGL2014-53274-P and CGL2012-39861-C02-01 from MINECO, grant GPC2013-011 from Xunta de Galicia, and Fondos Feder: "Unha maneira de facer Europa".

Data archiving

The doi for our data is 10.5061/dryad.7c68n1.

Literature Cited

Agrawal, A. F. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature 411 (6838)*, 692-695.

Agrawal, A. F. 2010. Ecological determinants of mutation load and inbreeding depression in subdivided populations. *The American Naturalist 176*(2): 111-122.

Agrawal, A. F., and J. R. Chasnov. 2001. Recessive mutations and the maintenance of sex in structured populations. *Genetics 158*(2): 913-917.

Agrawal, A. A., J. K. Conner, and S. Rasmann. 2010. Tradeoffs and negative correlations in evolutionary ecology. *Evolution since Darwin: the first 150*: 243-268.

Amador, C., A. García-Dorado, D. Bersabé, and C. López-Fanjul. 2010. Regeneration of the variance of metric traits by spontaneous mutation in a *Drosophila* population. *Genetics Research 92*(02): 91-102.

Ávila, V., and A. García-Dorado. 2002. The effects of spontaneous mutation on competitive fitness in *Drosophila melanogaster*. *Journal of Evolutionary Biology 15*(4): 561-566.

Ávila, V., C. Amador, and A. García-Dorado. 2010. The purge of genetic load through restricted panmixia in a *Drosophila* experiment. *Journal of Evolutionary biology 23*(9): 1937-1946.

Ballou, J. 1997. Ancestral inbreeding only minimally affects inbreeding depression in mammalian populations. *Journal of Heredity 88*(3): 169-178.

Bersabé, D., and A. García-Dorado. 2013. On the genetic parameter determining the efficiency of purging: an estimate for *Drosophila* egg-to-pupae viability. *Journal of Evolutionary Biology 26*(2): 375-385.

Boakes, E., J. Wang, and W. Amos. 2007. An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity 98*(3): 172-182.

Bryant, E. H., V. L. Backus, M. E. Clark, and D. H. Reed. 1999. Experimental tests of captive breeding for endangered species. *Conservation Biology 13*(6): 1487-1496.

Casella, G. and R. L. Berger. 2002. *Statistical Inference*. Second Ed. Pacific Grove, CA: Duxbury.

Crnokrak, P., and S. C. Barrett. 2002. Perspective: purging the genetic load: a review of the experimental evidence. *Evolution 56*(12): 2347-2358.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. *Nature Reviews Genetics 10*(11): 783-796.

Fernández, J., S. T. Rodríguez-Ramilo, A. Perez-Figueroa, C. Lopez-Fanjul, and A. Caballero. 2003. Lack of nonadditive genetic effects on early fecundity in *Drosophila melanogaster*. Evolution 57(3): 558-565.

Frankham, R. 2008. Genetic adaptation to captivity in species conservation programs. *Molecular Ecology 17*(1): 325-333.

Frankham, R., C. J. Bradshaw, and B. W. Brook. 2014. Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation 170*: 56-63.

Frankham, R., C. M. Gilligan, D. Morris, and D. A. Briscoe. 2001. Inbreeding and extinction: effects of purging. *Conservation Genetics 2*(3): 279-284.

Franklin, I. R. 1980. Evolutionary change in small populations. Pp. 135-149 M. E. Soule and B. A. Wilcox (eds.) in *Conservation Biology: an Evolutionary-Ecological Perspective. Sinauer Associates, Sunderland, Massachusetts*.

Franklin, I. R., F. W. Allendorf, and I. G. Jamieson. 2014. The 50/500 rule is still valid–Reply to Frankham *et al*. *Biological Conservation 176:* 284-285.

García-Dorado, A. 2012. Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics 190*(4): 1461-1476.

García-Dorado, A. 2015. On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. *Heredity* 115: 185-187.

García-Dorado, A., C. López-Fanjul and A. Caballero, 2004 Rates and effects of deleterious mutations and their evolutionary consequences, pp. 20–32 in *Evolution: From Molecules to Ecosystems*, edited by A. Moya and E. Font. Oxford University Press, Oxford.

García-Dorado, A., and A. Caballero. 2000. On the average coefficient of dominance of deleterious spontaneous mutations. *Genetics 155:* 1991-2001.

Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution 57*(12): 2678-2687.

Haag, C. R., and D. Roze. 2007. Genetic load in sexual and asexual diploids: segregation, dominance and genetic drift. *Genetics 176*(3): 1663-1678.

Hedrick, P. W. 1994. Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity 73*(4): 363-372.

Hedrick, P. W. and. S. T. Kalinowsky. 2000. Inbreeding Depression in Conservation Biology. *Annual Review of Ecology and Systematics 31*: 139-162.

Jamieson, I. G., and F. W. Allendorf,. 2012. How does the 50/500 rule apply to MVPs? *Trends in Ecology and Evolution 27*(10): 578-584.

Keller, L.F., and D. M. Waller. 2002. Inbreeding effects in wild populations. *Trends in Ecology and Evolution* 17: 230-241.

Kennedy, E. S., C. E. Grueber, R. P. Duncan, and I. G. Jamieson. 2014. Severe inbreeding depression and no evidence of purging in an extremely inbred wild species—the chatham island black robin. *Evolution 68*(4): 987-995.

Kruuk, L. E., B. C. Sheldon, and J. Merilä. 2002. Severe inbreeding depression in collared flycatchers (*Ficedula albicollis*). *Proceedings of the Royal Society of London. Series B: Biological Sciences 269*(1500): 1581-1589.

Latter, B., J. Mulley, D. Reid, and L. Pascoe. 1995. Reduced genetic load revealed by slow inbreeding in *Drosophila melanogaster*. *Genetics 139*(1): 287-297.

Leberg, P. L., and B. D. Firmin. 2008. Role of inbreeding depression and purging in captive breeding and restoration programmes. *Molecular Ecology 17*(1): 334-343.

Liberg, O., H. Andrén, H.-C. Pedersen, H. Sand, D. Sejberg, P. wabakken, M. Kesson, and S. Bensch. 2005. Severe inbreeding depression in a wild wolf *Canis lupus* population. *Biology Letters 1*(1): 17-20.

Lumley, A. J., L. Michalczyk, J. J. Kitson, L. G. Spurgin, C. A. Morrison, J. L. Godwin, M. E. Dickinson, O. Y. Martin, B. C. Emerson, T. Chapman *et al*. 2015. Sexual selection protects against extinction. *Nature* 522: 470-473.

Morton, N. E., J. F. Crow, and H. Muller. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. *Proceedings of the National Academy of Sciences of the United States of America 42*(11): 855.

O'Grady, J. J., B. W. Brook, D. H. Reed, J. D. Ballou, D. W. Tonkyn, and R. Frankham. 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation 133*(1): 42-51.

Otto, S. P., and A. C. Gerstein. 2008. The evolution of haploidy and diploidy. *Current Biology 18*(24): R1121-R1124.

Pedersen, K. S., T. N. Kristensen, and V. Loeschcke. 2005. Effects of inbreeding and rate of inbreeding in *Drosophila melanogaster*–Hsp70 expression and fitness. *Journal of Evolutionary Biology 18*(4): 756-762.

Pekkala, N., K. Emily Knott, J. S. Kotiaho, and M. Puurtinen. 2012. Inbreeding rate modifies the dynamics of genetic load in small populations. *Ecology and Evolution 2*(8): 1791-1804.

Ralls, K., J. D. Ballou, and A. Templeton. 1988. Estimates of lethal equivalents and the cost of inbreeding in mammals. *Conservation Biology 2*(2): 185-193.

Reed, D. H., E. H. Lowe, D. A. Briscoe, and R. Frankham. 2003. Inbreeding and extinction: effects of rate of inbreeding. *Conservation Genetics 4*(3): 405-410.

Roze, D., and F. Rousset. 2004. Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. *Genetics* 167(2): 1001-1015.

Sánchez-Molano, E., and A. García-Dorado. 2011. The consequences on fitness of equating family contributions: inferences from a *Drosophila* experiment. *Conservation genetics 12*(2): 343-353.

Siller, S. 2001. Sexual selection and the maintenance of sex. *Nature*, *411*(6838), 689-692.

Simmons, M. J., and J. F. Crow. 1977. Mutations affecting fitness in *Drosophila* populations. *Annual Review of Genetics 11*(1): 49-78.

Swindell, W., and J. Bouzat. 2006. Reduced inbreeding depression due to historical inbreeding in *Drosophila melanogaster*: evidence for purging. *Journal of Evolutionary Biology 19*(4): 1257-1264.

Walling, C. A., D. H. Nussey, A. Morris, T. H. Clutton-Brock, L. E. Kruuk, and J. M. Pemberton. 2011. Inbreeding depression in red deer calves. *BMC Evolutionary Biology 11*(1): 318.

Wang, J., W. G. Hill, D. Charlesworth, and B. Charlesworth. 1999. Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genetical Research 74*(02): 165-178.

Woodworth, L. M., M. E. Montgomery, D. A. Briscoe, and R. Frankham. 2002. Rapid genetic deterioration in captive populations: causes and conservation implications. *Conservation Genetics 3*(3): 277-288.

Yun, L., and A. F. Agrawal. 2014. Variation in the strength of inbreeding depression across environments: Effects of stress and density dependence. *Evolution 68*(12): 3599-3606.

CAPÍTULO SEGUNDO

PREDICTIVE MODEL AND SOFTWARE FOR INBREEDING-PURGING ANALYSIS OF PEDIGREED POPULATIONS

Aurora García-Dorado[1], Jinliang Wang[2] and Eugenio López-Cortegano[1]

[1] Departamento de Genética. Facultad de Biología. Universidad Complutense. Madrid 28040. Spain.

[2] Institute of Zoology. Zoological Society of London. London NW1 4RY. United Kingdom.

Running title: Purging in pedigreed data

Key words: Inbreeding depression; purging coefficient; rate of inbreeding depression; inbreeding load; logarithmic fitness

Corresponding author:

Aurora García-Dorado

Departamento de Genética, Facultad de Biología, Universidad Complutense, Ciudad Universitaria, Madrid 28040, Spain.

Phone number: 0034 913 944 975

Email address: augardo@ucm.es

García-Dorado, A., J. Wang and E. López-Cortegano. 2016. Predictive model and software for inbreeding-purging analysis of pedigreed populations. *Genes, Genomes, Genetics* 6 (11), 3593-3601.

Abstract

The inbreeding depression of fitness traits can be a major threat for the survival of populations experiencing inbreeding. However, its accurate prediction requires taking into account the genetic purging induced by inbreeding, which can be achieved using a "purged inbreeding coefficient". We have developed a method to compute purged inbreeding at the individual level in pedigreed populations with overlapping generations. Furthermore, we derive the inbreeding depression slope for individual logarithmic fitness, which is larger than that for the logarithm of the population fitness average. In addition, we provide a new software PURGd based on these theoretical results that allows analyzing pedigree data to detect purging and to estimate the purging coefficient, which is the parameter necessary to predict the joint consequences of inbreeding and purging. The software also calculates the purged inbreeding coefficient for each individual, as well as standard and ancestral inbreeding. Analysis of simulation data show that this software produces reasonably accurate estimates for the inbreeding depression rate and for the purging coefficient that are useful for predictive purposes.

Introduction

Due to the increase in the frequency of homozygous genotypes for (partially) recessive deleterious alleles under inbreeding, inbreeding depression is a major threat for the survival of small populations (Falconer and Mackay 1996, Saccheri *et al*. 1998, Hedrick and Kalinowski 2000, Frankham 2005). However, as these alleles become more exposed under inbreeding, an increase in the efficiency of natural selection against them is also expected, which is known as genetic purging and tends to reduce the frequency of deleterious alleles and, consequently, the fitness decline induced by inbreeding (Templeton and Read 1984, Hedrick 1994, Ballou 1997, García-Dorado 2012, 2015).

The first models developed to detect the consequences of purging on inbreeding depression from pedigree data accounted for purging by using an ancestral purging coefficient F_a that represents the proportion of an individual's genome that is expected to have been exposed to homozygosis by descent in at least one ancestor (Ballou 1997, Boakes and Wang 2005). The rational is that, due to genetic purging, inbred individuals with inbred ancestors would have fewer deleterious alleles than individuals with the same inbreeding but noninbred ancestors.

More recently, a theoretical Inbreeding-Purging (**IP**) approach has been developed that predicts the evolution of fitness under inbreeding by taking purging into account by means of a purged inbreeding coefficient *g*. This **IP** model considers that purging acts against a purging coefficient (*d*) that quantifies the component of the deleterious effects that are only expressed under inbreeding (García-Dorado 2012). For a single locus model, *d* represents the per copy excess of the deleterious effect in the homozygous over that expected on an additive hypothesis, and its value ranges from d=0 (no purging) to d=0.5 (purging against recessive lethal alleles). In practice, as *d* varies across loci, a single value, known as the effective purging coefficient (denoted by d_e in

García-Dorado 2012; here denoted by *d* for simplicity), can be used to compute approximate predictions for the overall consequences of purging over the whole genome. Estimating this effective *d* value is of main interest as it will provide a measure of the purging occurred, and will allow us to use the model to predict the expected evolution of fitness.

Until now, the only empirical estimates of the purging coefficient *d* have been obtained from the evolution of fitness average in *Drosophila* bottlenecked populations (Bersabé and García-Dorado 2013; López-Cortegano *et al*. 2016). However, in conservation practice, fitness data are often available for pedigreed populations. Two versions of the **IP** model were originally proposed, one aimed to predict mean fitness as a function of the number of generations under a reduced effective population size N_e, the other one aimed to predict individual fitness from pedigree information. Nonetheless, the latter version was only developed for data with nonoverlapping generations, which imposes serious limitations to its use in experimental and conservation practice.

Here we extend the **IP** model to compute the purged inbreeding coefficient *g* for individuals in pedigrees with overlapping generations. Furthermore, we derive a new expression that gives the expected individual log-fitness as a function of *g* and of the initial inbreeding load δ, deriving the slope of inbreeding depression for individual logarithmic fitness, which is larger than that for the logarithm of average population fitness. In addition, we present the new free software PURGd, based on this **IP** approach, that is able to use data for fitness traits in pedigreed samples to test for purging and to estimate the corresponding effective purging coefficient *d*. This software also estimates the inbreeding depression rate for individual fitness, and computes the standard (F), ancestral (F_a) and purged (g) inbreeding coefficients for the pedigreed individuals.

The Model

THE RATE OF INBREEDING DEPRESSION ESTIMATED FROM INDIVIDUAL FITNESS

In order to analyze and interpret the consequences of inbreeding and purging at an individual level, we must first consider the relationship between individual fitness and inbreeding in a neutral model with no natural selection.

Assume a population where a number of deleterious alleles segregate at a low frequency q at different loci acting multiplicatively on fitness. Here onwards we will concentrate just on (partially) recessive deleterious alleles, which are assumed to be responsible for inbreeding depression. Each locus has two alternative alleles, the wild one and the mutant deleterious allele. It has three genotypes, with average fitness 1, 1-hs and 1-s for the wild homozygous genotype, the heterozygous genotype and the deleterious homozygous genotype, respectively. Therefore, the population inbreeding load, which can be measured by the number of lethal equivalents (Morton *et al*. 1956), is

$$\delta=\Sigma 2dq(1-q), \qquad (1)$$

where $d=s(1/2-h)$, and the sum is over all the relevant loci.

For simplicity, we will assume that the initial frequency of each deleterious allele is small enough that homozygous genotypes are only produced due to inbreeding. Furthermore, in this section, we will also assume completely recessive gene action (h=0; $s=2d$). This assumption smooths the explanation below, but is not necessary for the validity of the conclusions.

After some inbreeding, the fitness of an individual that is homozygous by descent for

deleterious alleles at *n* loci is

$$W = W_{max}(1-\varepsilon)(1-2d)^n, \quad (2)$$

where W_{max} is the maximum possible fitness value and ε is the proportional reduction of the fitness of that individual due to all kind of environmental and genetic factors, excluding inbreeding depression.

If the inbreeding load is due to many loosely linked deleterious loci and deleterious alleles segregate at low frequency, the number n_i of deleterious alleles in homozygosis for an individual i with standard Wright´s inbreeding coefficient F_i should be Poisson distributed. Since the probability of being homozygous for a deleterious allele in noninbred individuals is assumed to be negligible, the expected value of this number should be is $E(n_i) = \Sigma F_i q(1\text{-}q)$ (Falconer and Mackay 1996). Thus, substituting $\Sigma q(1\text{-}q)$ from Equation (1), we obtain that the mean of this Poisson distribution is

$$\lambda = E(n_i) = F_i \, \delta / 2d \quad (3)$$

Therefore, from Equation 2, and assuming that ε and F are independent, the expected fitness of an individual i that has genealogical inbreeding F_i is

$$E(W_i) = E(W_0) \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!}(1-2d)^n$$

where $E(W_0) = E[W_{max}(1\text{-}\varepsilon)]$ is the expected fitness of a noninbred individual. The equation above can be rewritten as

$$E(W_i) = E(W_0) \; e^{-\lambda 2d} \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n \, e^{\lambda 2d}}{n!}(1-2d)^n,$$

and can be rearranged as

$$E(W_i) = E(W_0) \quad e^{-\lambda 2d} \sum_{n=0}^{\infty} \frac{e^{-\lambda(1-2d)}[\lambda(1-2d)]^n}{n!}.$$

Noting that $\sum_{n=0}^{\infty} e^{-\lambda(1-2d)} [\lambda(1-2d)]^n/n!$ adds up all the probabilities for a Poisson distribution with mean λ(1-2*d*) (*i.e.*, it equals 1), and since λ = F_i δ / 2*d* (Equation 3), we obtain the exponential expression

$$E(W_i) = E(W_0) \quad e^{-\delta F_i} , \tag{4}$$

and, similarly, the average fitness of a population with average inbreeding F_t in generation *t* , as far as the number of loci homozygous for a deleterious allele per individual can be assumed to be Poisson distributed with mean λ = F_t δ / 2*d*, is

$$E(W_t) = E(W_0) \quad e^{-\delta F_t} , \tag{5}$$

In order to estimate δ from observed inbreeding depression, logarithms are usually taken in Equations 4 or 5 to obtain a linear model of the kind ln(*W*) = ln(W_0) - δ *F*. However, since the average of the logarithms of a variable is smaller than the logarithm of the average (see Jensen's inequality), applying this procedure to individual fitness values can produce important upwards bias in the estimate of δ. Thus, from Equation 2, the logarithm of fitness (log-fitness hereafter) for an individual that is homozygous by descent for *n* deleterious alleles is

$$\ln(W) = \ln[W_{max}(1-\varepsilon)] + \ln[(1-2d)^n],$$

so that, using the Poisson distribution of n_i, the expected value for log-fitness for an individual *i* that has genealogical inbreeding F_i is

$$E[\ln(W_i)] = E[\ln(W_0)] + \sum_{n=0}^{\infty} \ln[(1-2d)^n] \frac{e^{-\lambda}\lambda^n}{n!}, \tag{6}$$

where the intercept E[ln(W_0)] = E{ln[W_{max} (1-ε)} represents the average of individual log-fitness

at the noninbred population. Since the second term equals ln(1-2d)E(n_i), using Equation 3, Equation 6 gives

$$E[\ln(W_i)] = E[\ln(W_0)] + \frac{\ln(1-2d)}{2d}\,\delta F_i\,, \qquad (7)$$

On the other hand, in agreement with classical theory (Morton *et al*. 1956), Equations 4 and Equation 5 imply

$$\ln [E(W_i)] = \ln [E(W_0)] - \delta F_i \qquad (8)$$

and

$$\ln [E(W_t)] = \ln [E(W_0)] - \delta F_t \qquad (9)$$

It is interesting to note that, as indicated by Morton *et al*., the two equations above produce good approximation as far as each individual locus makes a small contribution to the overall expected inbreeding load.

Equation 8 allows to estimate δ from the decline in average fitness for a given inbreeding level, as in designs where fitness is measured in a sample of outbred and a sample of inbred individuals (for example, full sib offspring). Equation 9 allows to estimate δ, generally using linear regression, from the decline in average fitness through generations of inbreeding, as in a population that has experienced a reduction in size. Both approaches induce no bias in the estimate of δ, as far natural selection can be ignored and sample sizes are sufficiently large that the expected value of the logarithm of the sample's average is close to the logarithm of the expected average (*i.e*., to $\ln[E(W_t)]$ or $\ln[E(W_i)]$.

However, Equation 5 shows that the slope of linear regression for the logarithm of individual fitness on individual inbreeding is

$$b = \frac{\ln(1-2d)}{2d}\delta, \quad (10)$$

where the limit of $\frac{\ln(1-2d)}{2d}$ as *d* approaches 0 is -1. Therefore, unless *d* is very small, -*b* provides an upwardly biased estimate for the inbreeding load δ.

Here we present a software package (PURGd) that estimates the purging coefficient and the inbreeding load from the relationship between individual fitness and individual inbreeding using two alternative approaches. The first approach estimates *b* from the linear regression of log individual fitness on individual genealogical inbreeding. The second approach estimates δ by numerical least squares (**LS**) from untransformed fitness, directly using Equation 4. In addition to allowing the use of individual fitness data including 0 values (as in the case of a dichotomous 0-1 variable for dead-alive records), this procedure allows to directly estimate δ, instead of *b*.

THE INBREEDING-PURGING (IP) MODEL: COMPUTING PURGED INBREEDING AND PURGED COANCESTRY FROM PEDIGREES

According to the IP approach, in order to incorporate the consequences of purging, the evolution of fitness under inbreeding should be predicted by replacing the standard inbreeding coefficient *F* with a purged inbreeding coefficient *g* where *F* is weighted by the reduction in frequency of deleterious alleles induced by purging. Thus, Equations 4 and Equation 5 become:

$$E(W_i) = E(W_0) \quad e^{-\delta g_i}, \quad (11)$$

$$E(W_t) = E(W_0) \quad e^{-\delta g_t}, \quad (12)$$

García-Dorado (2012) derived equations allowing to compute g_i for individuals in pedigrees with nonoverlapping generations. These g_i values depend on the pedigree and on the *d* value defined above as *d*=*s*(1/2-*h*), which here represent the purging coefficient. For multilocus models where *d* varies across loci, it has been empirically shown using extensive simulations that

d can be replaced with an effective purging coefficient that accounts for purging across the whole genome to a good approximation. This effective purging coefficient was denoted d_e in García-Dorado (2012) but here, for simplicity, it will be denoted *d* and referred to just as purging coefficient.

In what follows we derive more general expressions to compute approximate g_i values for individuals in arbitrary pedigrees that can include overlapping generations.

The purged inbreeding coefficient g_i is defined as $g_i = E(F_i\ q_i)/q_0$, were E stands for "expected value" and q_0 (q_i) is the frequency of the deleterious alleles in the base population (expected in individual *i*). In other words, ($q_0\ g_i$) is the probability that individual *i* is homozygous by descent for the deleterious allele. In order to settle notation, we will use A and B to denote individual X's parents, C and D to denote individual A's parents, and E and H to denote individual B's parents, as shown in Figure 1.

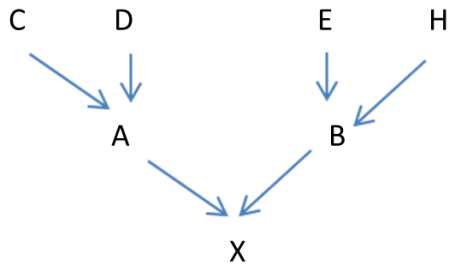


Figure 1: General pedigree notation

Let *f*(A,B) be Malécot's coancestry between individuals A and B; *i.e.*, the probability that a random allele from a neutral locus in A and, independently, a random allele from the same locus in B are identical by descent (IBD) (Malécot 1948). By analogy to García-Dorado 2012, we will assume that the probability that two copies sampled from different individuals are IBD is unaffected by the fitness values of the copies.

As in García-Dorado 2012, let $\gamma(A,B)$ be the purged coancestry between A and B, which are assumed to have survived purging selection. In other words, $[q_0\,\gamma(A,B)]$ is the probability that two alleles, one randomly sampled from A and the other independently sampled from B at the same locus, are identical by descent for the deleterious allele. Therefore, the purged inbreeding coefficient for an individual X that has still not undergone purging, can be computed as the purged coancestry between their parents; *i.e.*, $g_x = \gamma(A,B)$.

Note that $q_0 \cdot \gamma(A, B)$ could be defined as the probability that an allele randomly sampled from A is deleterious and identical by descent to another allele randomly sampled from B; *i.e.*, $q'_A\, f(A,B)$, where q'_A denotes the frequency of the deleterious allele in individual A conditional to it having survived purging selection. Alternatively, $[q_0 \cdot \gamma(A, B)]$ could also be defined as the probability that an allele randomly sampled from B is deleterious and identical by descent to an allele randomly sampled from A (*i.e.*, $q'_B\, f(AB,)$. Therefore, by averaging both alternatives we obtain

$$q_0 \cdot \gamma(A, B) = \tfrac{1}{2}\,(q'_A + q'_B)\, f(A,B) \qquad (13)$$

Finally, let $\gamma(A,B|E)$ be the purged coancestry between A and B conditional to sampling from B the copy inherited from E. In other words, $[q_0\,\gamma(A,B|E)]$ is the probability that one allele randomly sampled from A is deleterious and identical by descent to the copy that B inherited from E.

Therefore, $[q_0 \cdot \gamma(A, B)]$ is the probability that the copy sampled from B was inherited from E (*i.e.*, ½) and then the two copies (one sampled from A and the other one from B) are identical by descent for a deleterious allele, plus the analogous probability corresponding to sampling from B the copy inherited from H, *i.e.*:

$$q_0\ \gamma(A,B) = \tfrac{1}{2}\, q_0\ \gamma(A, B|E)\ + \tfrac{1}{2}\, q_0\ \gamma(A, B|H) \qquad (14)$$

Thus, we need a systematic procedure to compute $\gamma(A, B|E)$ that can be recurrently used to obtain $\gamma(A,B)$ and g_x. To achieve this, we note that the probability that one allele randomly sampled from A and the copy that B inherited from E are identical by descent for the deleterious allele can be computed in two ways:

i) After B survives purging, the copy in B inherited from E is the deleterious allele. Since purging is expected to reduce deleterious frequency in B by a factor $(1\text{-}2d\ F_B)$ (García-Dorado 2012), this occurs with probability $q'_E \cdot [1\text{-}2\cdot d\cdot F_B]$. Furthermore, this copy is identical by descent to that sampled from A. Taking into account that $f(A,E)$ is assumed to be independent on the allelic state (*i.e.*, is the same for deleterious as for wild alleles), this occurs with probability $q'_E \cdot [1\text{-}2\cdot d\cdot F_B] \cdot f(A,E)$

ii) The copy sampled from A is deleterious and is identical by descent to the copy that B has inherited from E. This occurs with probability $q'_A \cdot f(A,E)$

Thus, we compute $[q_0 \cdot \gamma(A, B|E)]$ by averaging these two probabilities above, which gives

$$q_0 \cdot \gamma(A, B|E) = \tfrac{1}{2} (q'_A + q'_E) f(A, E,) - q'_E \cdot f(A, E,)\, d\cdot F_B \qquad (15)$$

Now, if inbreeding progresses slowly, the last $q'E$ in the above expression can be replaced with $\tfrac{1}{2} (q'_A + q'_E)$ to a good approximation, and Equation 15 approaches

$$q_0 \cdot \gamma(A, B|E) \approx [\tfrac{1}{2} (q'_A + q'_E) f(A, E)] (1- d\cdot F_B) \qquad (16)$$

which, applying Equation 13 to A and E, gives the approximate expression

$$\gamma(A, B|E) = \gamma (A, E) (1- d\cdot F_B) \qquad (17)$$

Therefore, substituting the conditional purged coancestry given by Equation 17 into Equation 14, we obtain

$$\gamma(A,B)= \tfrac{1}{2}\ [\gamma\ (A, E) + \gamma\ (A, H)]\ (1-\ d{\cdot}F_B\ )\ , \qquad (18)$$

As in the case of classical Malécot's coancestry (f), purged inbreeding arises from the pedigree knots where $\gamma(A,B)$ happens to represent a self-coancesty (A and B are the same individual). In those cases, as previously shown (García-Dorado 2012),

$$\gamma(A,A) = \tfrac{1}{2}\ [1 + g_A][1 - 2d\ F_A] \qquad (19)$$

Equation 18 is analogous to the classical recurrent expression that gives the coancestry between A and B as the average coancestry between A (which should not be younger than B) and both parents of B ($f(A, B) = [f(A, E) + f(A, H)]$), except that Equation 18 accounts for the purging occurred in B. Thus, Equation 18 can be recurrently used together with Equation 19 to compute purged coancestry between pairs of individuals that have survived purging, which equates the purged inbreeding expected for their offspring ($g_x = \gamma(A,B)$)

To compare this approach with that previously derived for nonoverlapping generations, we note that, analogously to Equation 18, we can write

$$\gamma(A, E) = \tfrac{1}{2}\ [\gamma(C,E)+ \gamma(D,E)]\ (1 - d\ F_A), \qquad (20)$$

and

$$\gamma(A, H) = \tfrac{1}{2}\ [\gamma(C,H)+ \gamma(D,H)]\ (1 - d\ F_A), \qquad (21)$$

And, substituting Equations 20 and 21 into Equation 18, we obtain

$$\gamma(A, B) = \tfrac{1}{4}\ [\gamma(C,E)+ \gamma(D,E)+ \gamma(C,H) + \gamma(D,H)](1-d\ F_A)(1-d\ F_B)\ .$$

This expression slightly overrates the purged coancestries (and, therefore, the purged inbreeding coefficients) derived by García-Dorado (2012) for nonoverlapping generations, which gave

$$\gamma(A, B) = ¼ [\gamma(C,E) + \gamma(D,E) + \gamma(C,H) + \gamma(D,H)] [1 - d (F_A+F_B)] \quad (22)$$

The overrate is due to the use of the approximation $q'_E \approx ½ (q'_A + q'_E)$ to derive Equation 16, which, on the average, underrates the deleterious frequency against which purging is operating. The bias should however be small, since the squared term ($d^2 F_A F_B$) can only be important for large d and F values, which implies small is γ and g values. Using simulated pedigrees in bottlenecked populations with nonoverlapping generations, we have found that the correlation between $\gamma(A, B)$ computed from Equation 18 and from García-Dorado 2012 were always larger than 0.999 for a wide range of different purging coefficients from $d=0$ to $d=0.5$ (results not shown).

Finally, it must be noted that, for **IP** predictions to be reliable, drift should be relatively unimportant, compared to purging. Thus, when considering the consequences of inbreeding and purging on average fitness, predictions are reliable for $dN_e > 1$, where N_e is the drift effective population size (García-Dorado 2012). For panmictic populations of constant size, drift effective size is equal to inbreeding effective size ($N_e = 1 / 2\Delta F$, where ΔF is the per generation inbreeding rate), so that we can expect **IP** predictions to be reliable if, through the whole process, $d > 2\Delta F$. This rate can be computed for consecutive time periods with length equal to the average generation interval. Thus, at each interval $\Delta F = (F' - F) / (1 - F)$, where F and F' are the average inbreeding in the population at the beginning and the end of the interval.

DATA AVAILABILITY

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. PURGd software and example data are available in https://www.ucm.es/genetica1/mecanismos.

The Software

We present a new software package (PURGd, available from https://www.ucm.es/genetica1/mecanismos) that uses the **IP** model to jointly estimate the effective purging coefficient, *d*, and the inbreeding load in the base population, or its related parameter, *b*, defined in Equation 10, that better account for the fitness values of a set of pedigreed individuals. Additional details are given in the user's guide included in the package.

The program computes standard coancestry and inbreeding (*f* and *F* values), as well as Ballou's ancestral inbreeding coefficient (F_a) for each individual. Furthermore, for each *d* value considered, it recurrently uses Equation 18, Equation19 and Equation 22 to compute the corresponding purged inbreeding coefficients (*g*). Using these coefficients, the program obtains LS estimates for the *d* value, and for the remaining parameters in the model. As the predictive model may incorporate additional factors potentially affecting fitness, and since fitness is assumed to be a multiplicative trait, Equation 11 is generalized to include an arbitrary number of additional factors (say *x*, z…), giving the general model

$$E(W_i) = W_{max}(1-\varepsilon_i)\ e^{\beta_1 g_i + \beta_2 x_i + \beta_3 z_i \ldots}\ , \qquad (23)$$

where $\beta_1 = -\delta$ is the regression coefficient on purged inbreeding *g*, *g* is a function of *d*, and the remaining β_j values measure the effect of the corresponding additional factors, which may include the maternal purged inbreeding coefficient.

This software numerically searches the *d* value that minimizes the squared deviations from observed fitness to model predictions (*i.e.*, for the least square LS estimate). However, regarding the remaining parameters, the model can be fitted using two different approaches, as explained below. In the first one (linear regression method, LR), for each *d* value considered, a linear regression model is fitted for log-transformed fitness. In the second one (numerical nonlinear

regression method, NNLR), the above model for untransformed fitness (Equation 23) is numerically explored searching for the joint numerical LS estimates of *d* and of the nonlinear regression coefficients. Although this NNLR method is computationally more demanding, the program runs quickly and has low RAM requirements under both approaches. Optionally, the initial average for fitness or log-fitness and/or for the regression coefficient on *g* can be introduced by the user, allowing to incorporate independent estimates of these parameters when available.

Additionally, the software will also give the results for the corresponding analysis conditional to d=0, so that the user can observe the consequences of considering/ignoring purging in the analysis and can check how the model improves under the estimate of *d*, compared to the assumption of no purging (d=0).

LR METHOD

To perform LR analysis, the model represented by Equation 23 is linearized by taking logarithms. This leads to the linear predictive equation

$$\ln(W_i) = b_0 + b_1 g_i + b_2 x_i + b_2 z_i \ldots ,$$

where the different *b* values estimate the corresponding regression coefficients. Since logarithms are taken for individual fitness, instead of for average fitness, by analogy to Equation 7, the intercept b_0 estimates $E[\ln(W_0)]$, and b_1 estimates $[\ln(1-2d)/2d]\delta$ (Equation 10).

However, as it has been noticed (García-Dorado 2012), the IP model is a conservative approach that tends to underrate the long term fitness expected from inbreeding and purging. For this reason, when the estimate of the expected log-fitness for noninbred individuals (b_0) is obtained jointly with b_1 and with the purging coefficient (d), the method tends to overfit the model by estimating a too low initial fitness and, simultaneously, too small values for the decline of log

fitness with F_i (*i.e.*, for $-b_1$) and for the purging coefficient *d*. Thus, this procedure tends to give b_1 and *d* estimates that will produce poor predictions when extrapolated to populations with different rates of inbreeding, or to periods of different length. On the contrary, when $E[\ln(W_0)]$ is not simultaneously estimated, the estimates *b* and *d* have much smaller bias and good predictive properties.

Therefore, b_0 is obtained by PURGd in a previous step as the average of log-fitness for noninbred individuals with noninbred ancestors ($F = F_a = 0$), or is introduced by the user as a known value. Then, in a second step, the software searches for the *d* value that optimizes the fitting of the data to the linear regression equation

$$Y_i = b_1 g_i + b_2 x_i + b_2 z_i \ldots \quad ,$$

where the dependent variable is $Y_i = \ln(W_i) - b_0$, so that regression is forced through the origin.

Regression analysis is performed for all the possible *d* values in a grid corresponding to the interval $0 \leq d \leq 0.5$ with step 0.01, which is the default accuracy. If higher accuracy is requested, PURGd first finds a preliminary estimate with precision 0.01 as before, and then uses the Golden Section Search (GSS) algorithm in an interval ± 0.01 around that estimate (Press *et al* 1992).

Finally, the software returns the *d* estimate that minimizes the residual sum-of-squares in the corresponding LR analysis of individual log-fitness. For each analysis, the program also gives the corresponding results of the above LR, with statistic contrasts assuming normality and independence of residual errors, and with the adjusted determination coefficient and the corrected Akaike information criterion, computed taking into account how many parameters are being estimated in the whole process.

Table 1 reproduces the software's output for LR approach, where estimates have been averaged for the analysis of a set of 50 simulated lines. Each line is derived from a large panmictic

population at the Mutation-Selection-Drift balance (N=1000), and is maintained with size N=10 during 50 generations. Completely recessive deleterious mutations with homozygous effect s=0.3 occur at a rate of 0.1 new mutations per gamete and generation in unlinked sites. Since h=0, this implies that the theoretical value for the purging coefficient is d=0.15. The simulation details can be found in Bersabé *et al*. 2016. Output is presented for two different simulation sets. In the first one, natural selection is operating during the maintenance of the lines, so that purging is expected to occur. In the second set, natural selection is relaxed, implying no purging. To achieve this, when simulating each offspring all individuals had the same probability of being sampled as parents of the next generation regardless their fitness values. The software estimates a purging coefficient d = 0.102 ± 0.009 in the first case, and d = 0.003 ± 0.001 in the second one (SE computed from the 50 replicates). Therefore, the method has discriminated between situations with or without purging, although it has underestimated the actual purging coefficient. Furthermore, for lines undergoing purging, the data fit much better the **IP** model prediction, computed using the corresponding estimate of d, than that conditional to d=0 that assumes no purging, as shown by the higher determination coefficient and the smaller residual sum of squares and Akaike criterion.

Pedigree file	*Analysis*	*d coefficient*	*RSS*	*p-value (F)*	*aR2*	*AICc*	*lnW0*	*SD(lnW0)*	*b(g)*	*SD[b(g)]*	*p-value(t)*
Purged_lines	**IP** *model*	0.102	147.291	<1.0e-16	0.758	804.642	-0.124	0.206	-3.298	0.081	<1.0e-16
	No- purging model	0	253.130	<1.0e-16	0.586	1069.500	-0.124	0.206	-1.222	0.041	<1.0e-16
Relaxed_lines	**IP** *model*	0.003	188.396	<1.0e-16	0.966	921.812	-0.122	0.201	-5.177	0.040	<1.0e-16
	No-purging model	0	195.72	<1.0e-16	0.964	944.204	-0.122	0.201	-4.965	0.039	<1.0e-16

Table 1. **Averaged results obtained using the linear regression method (LR) for the set of 50 simulated lines described in the main text that were maintained with size N=10during 50 generations, where the true values for the inbreeding load and the purging coefficients in the base population are δ=4.217 and *d*=0.15, respectively.** These results are shown in the same format as in the PURGd output. Pedigree File, name of the data file; Analysis, the model used in the analysis; *d* coefficient, the purging coefficient estimated in the IP analysis or assumed by the No-purging model; RSS, residual sum of squares; P-value (F), the P-value in the F-test for the regression analysis; aR2, adjusted determination coefficient; AICc, the corrected Akaike Information Criterion; $\ln W_0$, the estimate of the expected log-fitness in the base noninbred population; SD($\ln W_0$), SD of $\ln W_0$; b(g), linear regression coefficient on *g* (it is denoted b_1 in the predictive equation and estimates [ln(1-2d)/2d]δ, as defined in Equation 10; its expected value in this case is -5.014, very close to the IP estimate obtained for the relaxed lines); SD[b(g)], SD of b(g); P-value(t), P-value for the t-test on the significance of this linear regression coefficient.

The analysis of additional simulated lines maintained with size N=50 (not shown) produced similar results, again discriminating between purged and relaxed lines and providing better fitting for purged lines when using the corresponding estimates of *d*. For purged lines, the estimate for the regression coefficient of fitness on purged inbreeding was $b(g) = -3.590 \pm 0.276$ which, solving Equation 10, gives an estimate $\delta = 3.019$ for the inbreeding load, close to the value obtained for N=10 ($\delta = 2.774$), but the estimate for the purging coefficient was larger ($d = 0.218 \pm 0.029$).

NNLR METHOD

The previous logarithmic transformation cannot be applied to fitness traits presenting null values, as in the case binary of 0/1 variables for death/alive records. In such cases, inbreeding depression has previously been analyzed using a logit transformation of fitness in order to perform multiple logistic regression (Ballou, 1997; Boakes and Wang 2005). However, that statistical approach assumes a model of the kind $\ln[W_i/(1-W_i)] = \beta_0 - \beta_1 g_i$, while our genetic model has the form $\ln(W_i) = \beta_0 - \beta_1 g_i$,. Therefore, PURGd gives the user the option of obtaining LS estimates for the parameters in the genetic model given by Equation 23 by numerically optimizing the fitting of the untransformed fitness data to the predictions of the nonlinear regression equation given by

$$W_i = W_0 e^{b_1 g_i + b_2 x_i + b_3 z_i \ldots} ,$$

where the different *b* values are the estimates of the corresponding β parameter in Equation 23, so that b_1 estimates -δ, and W_0 is the estimate of the expected fitness value for the noninbred base population. For the same reasons as in the LR method, W_0 is obtained in a previous step as the average *W* for the set of individuals with $F = F_a = 0$, or

is introduced by the user.

After estimating W_0, the Numerical Least Square option of PURGd uses the Artificial Bee Colony (ABC) algorithm (Karaboga and Basturk 2007) to search simultaneously for the LS estimate of the purging coefficient *d* (where each *d* value considered determines a set of g_i values) and for the set of *b* coefficients that produces the lowest residual sum of squares (RSS), calculated as:

$$\text{RSS} = \sum_i (W_i - W_0 e^{b_1 g_i + b_2 x_i + b_3 z_i \ldots})^2$$

This algorithm has been successfully used for estimating parameters in nonlinear systems in different kinds of disciplines such as image processing, engineering or neural networks among others (Karaboga *et al*. 2014) using ~500 generations and 250 bees in the colony. Although we have always found consistent solutions, it is recommended to repeat analysis several times to check for the stability of the method, and to change running parameters and range values, looking for a consistent solution.

Therefore, the output gives a LS estimate for *d* and for the remaining β_j parameters in the model (Equation 23). An important advantage of this approach is that, besides allowing to deal with 0 fitness values, $-b_1$ directly estimates the inbreeding load δ, instead of estimating $-[\ln(1-2d)/2d]\delta$. Furthermore, although LS estimates for nonlinear regression are not expected to be unbiased, preliminary unpublished simulated results suggest that this method usually gives estimates of the purging coefficient and of the inbreeding load that produce predictions at least as accurate as those obtained using estimates computed from linear regression on log-fitness data, though it is computationally more demanding. Although this approach does not allow to perform standard *F*-Tests for statistical significance, the RSS and the corrected Akaike

information criterion values (the later again relying on the assumption of normality and independence for residual errors) are reported in the output as a measure of the fitting quality.

Table 2 reproduces the software's output for this NNLR approach, where estimates have been averaged for the analysis of the same sets of simulated lines analyzed in Table 1. In this case the estimates of the purging coefficient for lines maintained with natural selection is $d = 0.092 \pm 0.007$, and that obtained for lines maintained under relaxed selection is $d = 0.007 \pm 0.001$, again discriminating between purging and no purging cases but underestimating the purging coefficient (SE again empirically estimated from the 50 replicated lines). As in the LR method, the data for simulated lines undergoing purging fit much better the **IP** model than the d=0 no-purging model.

For simulated lines maintained with size N=50 (not shown), NNLR analysis of data discriminated between purged and relaxed lines, and provided better fitting for purged lines when using the corresponding estimates of d, as in the case of the LR analysis. Again, the estimate for the inbreeding load for purged lines ($\delta = -b(g) = 2.756 \pm 0.241$), was very close to that estimated for N=10, but the estimate for the purging coefficient was larger (d=0.190 ± 0.005).

Pedigree file	*Analysis*	*d coefficient*	*RSS*	*AICc*	*W0*	*SD(W0)*	*b(g)*
Purged_lines	**IP** *model*	0.092	16.996	-326.399	0.902	0.152	-2.898
	No- purging model	0	28.387	-71.356	0.902	0.152	-1.202
Relaxed_lines	**IP** *model*	0.007	4.072	-1037.943	0.903	0.154	-4.533
	No-purging model	0	4.145	-1033.899	0.903	0.154	-4.443

Table 2. Averaged results obtained using the numerical nonlinear regression method (NNLR) for the set of 50 simulated lines described in the main text that were maintained with size N=10 during 50 generations, where the true values for the inbreeding load and the purging coefficients in the base population are δ=4.217 and *d*=0.15, respectively. These results are shown in the same format as in the PURGd output. Pedigree File, name of the data file; Analysis, the model used in the analysis; *d* coefficient, the purging coefficient estimated in the IP analysis or assumed by the No-purging model; RSS, residual sum of squares; AICc, the corrected Akaike Information Criterion; W_0, the estimate of the expected fitness in the base noninbred population; SD(W_0), SD of W_0; b(g), nonlinear regression coefficient on *g* that estimates the inbreeding load (n(g), denoted b_1 in the predictive equation, estimates -δ).

PREDICTIVE VALUE OF THE ESTIMATES

Figure 2 gives the evolution of fitness against generation number and the corresponding IP predictions, computed for each set of lines using in Equation 12 the corresponding estimates of δ and *d* obtained by the software. Good fitting is observed for N=10 and for N=50 regardless whether LR or NNLR are used, both for the relaxed lines and for those maintained under purging.

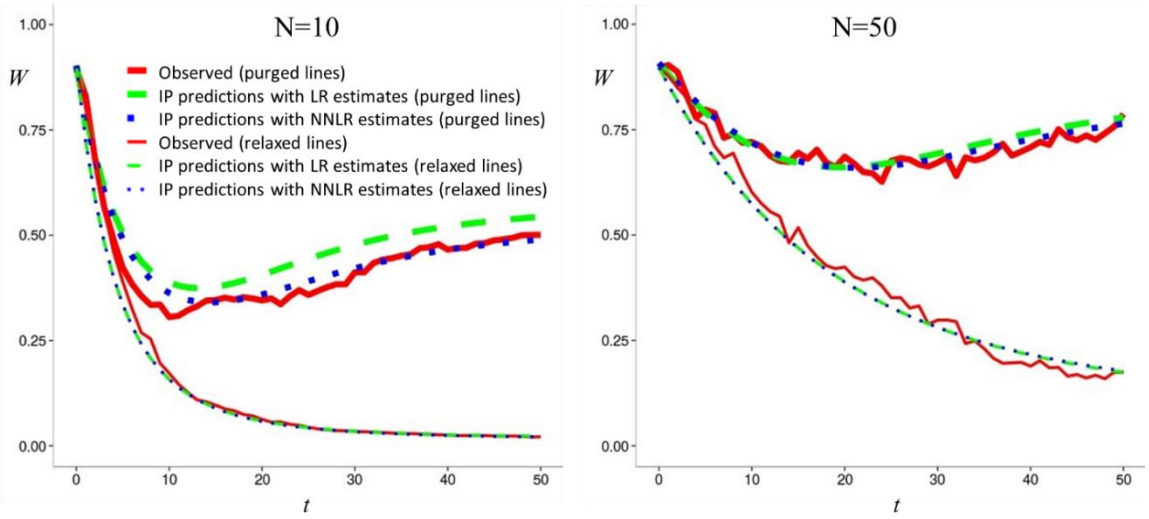


Figure 2. Evolution of mean fitness through generations for simulated lines maintained with size N=10 (analysis given in Tables 1 and 2) or N=50 during 50 generations (red solid lines), together with IP predictions computed using the estimates obtained by PURGd from the linear regression method (LR, green dashed lines) or the numerical nonlinear regression method (NNLR, blue dotted lines). Results are given both for lines that have undergone purging (thick lines), and for lines for which natural selection was relaxed while they were maintained with reduced size (thin lines, which largely overlap with each other).

Discussion

In the present work we derive a theoretical approach to analyze the fitness data for pedigreed individuals in order to estimate the inbreeding load δ and the purging coefficient *d* necessary to predict the joint consequences of inbreeding and purging. Furthermore, we present PURGd, a free software implementing this theoretical approach, and illustrate its performance analyzing some results obtained by the software for simulated data.

In the first place, since the inbreeding depression rate is usually estimated from log-fitness data, we derive the expected regression slope of individual log-fitness on

individual inbreeding in the absence of selection, which amounts to $b = [\ln(1-2d)/2d]\delta$. Therefore, using $-b$ as an estimate of the inbreeding load δ implies upwardly biased estimation. This first result is interesting because increased effort in field studies related to conservation of endangered species, together with molecular technics, allow to record and/or reconstruct pedigrees in wild populations and offers an interesting opportunity to study inbreeding depression in the wild (Keller and Waller 2002), but can induce upwardly biased estimates due to the use of log transformed individual fitness. The bias is expected to be small if *d* values are low, but the large inbreeding depression rate estimated in wild populations are likely to be associated to relatively large *d* values and, therefore, to substantial bias (Kruuk *et al.*, 2002; Liberg *et al.*, 2005; O'Grady *et al*. 2006; Walling *et al.*, 2011; Kennedy *et al.*, 2014; Hedrick *et al*. 2016). This phenomenon can contribute to enhance the perceived difference between the inbreeding load expressed in wild populations compared to estimates based on the assay of mean fitness for groups of individuals with different average inbreeding, as is often the case in experimental conditions. In order to avoid this bias, an alternative estimation approach is suggested, based on the numerical LS analysis of the original predictive IP model for untransformed fitness. This approach is implemented in the PURGd software, and is used to analyze some simulated data.

In the second place, in order to estimate the purging coefficients (*d*) from individual fitness data, we present general expressions to compute purged inbreeding (*g*) from pedigrees with overlapping generations. Although these expressions involve some approximations, we have found that they produce reliable values for individual *g*.

Other methods for detecting purging from fitness measured in pedigreed individuals have been previously devised, based on the idea that the ancestral purging F_a of an

individual is someway related to the opportunities of purging upon its genome in previous generations. Using F and F_a, different linear models have been proposed that have, in some occasions, detected small levels of purging in simulated and real pedigrees of captive breeding populations (Ballou 1997, Lacy and Ballou 1998, Boakes and Wang 2005, Swindell *et al*. 2006, Boakes *et al*. 2007, Ceballos and Álvarez 2013). However, these methods were based on the analysis of statistical models that are not supported by a predictive genetic model. In addition, a logit transformation was applied to fitness, just on statistical grounds. Therefore, these models could fit fitness data poorly. More importantly, they do not allow to estimate a purging parameter that can be used for predictive purposes. On the contrary, our method is based on the predictive IP model that was derived on the basis of the genetic mechanisms of inbreeding depression and purging, so that it is expected to fit the data better, and to allow the estimation of a parameter that can be used for predictive purposes: the effective purging coefficient *d*. However, the model involves some approximations and usually produces conservative predictions underrating the consequences of purging. Therefore, statistical methods based on this IP model can overfit the model by inducing some bias in the estimates.

For illustrative purposes we have presented here the analysis of a set of simulated data for a simple situation where inbreeding and purging occur due to a reduction in population size (Table 1 and Table 2). For N=10, the inbreeding load computed using Equation 1 in the base simulated population was δ=4.217. The LR method estimates d = 0.102 ± 0.009 and b=-3.298 ± 0.096 (SE computed from the 50 replicates analyzed), which using the true simulated value for d (0.15) into Equation 6 gives an estimate of the inbreeding load δ=2.774. Thus both the inbreeding load and the purging coefficient are underestimated when they are jointly estimated. The δ and d estimates obtained using the numerical method are very similar (2.898 ± 0.115 and 0,092 ± 0.007, respectively). Under

both methods, the data fit the IP model much better than the no-purging (d=0) model. In parallel, we present the analysis for a similar set of simulated lines where selection and, therefore, purging, had been relaxed during the inbreeding period. It is worth noticing that the estimates of the purging coefficient d given by PURGd for these relaxed lines are virtually zero, showing that the method detects whether purging is occurring or not. Furthermore, when natural selection is relaxed during the maintenance of the reduced size lines, the LR approach gives $b = -5.177 \pm 0.165$ so that the estimate of δ is 4.354, and the δ estimate obtained using the numerical approach is very similar (4.533). Thus, the underestimation of δ observed when purging is operating in the lines, can be ascribed to regression overfitting the data through the underestimation of both δ and d, due to the approximate nature of the IP model. It should be noted that some underestimation of d could also occur because, for Nd on the order of 1 or smaller, purging efficiency may be somewhat reduced due to genetic drift (García-Dorado 2012). On the contrary, d estimates obtained for simulated purged lines maintained with N=50 are larger than the actual d value, while δ is simultaneously underestimated. In all cases, using jointly the δ and d estimates obtained in the same analysis gives appropriate predictions for the evolution of mean fitness (Figure 2).

The software also allows including additional factors, both in the linear and the nonlinear models. However, the addition of factors with a strong association with g, as maternal inbreeding or year of birth, often causes a slight overfitting, again due to the approximate nature of the program. The overfitted model gives spurious significant effects for such factors as well as some distortion in the estimates of $b(g)$ and d (results not shown) due to confounded effects. Therefore, results obtained by incorporating additional factors should better be used when those factors are uncorrelated to g, so that including them just reduces sampling error. Additional factors should also be tentatively

included when there is external evidence that they have a highly relevant effect, so that including them cause an important improvement of the fitting statistics. However, when these additional factors are correlated to *g*, these results should be interpreted with caution and those obtained including no additional factors should also be considered.

It is interesting to note that, using in Equation 12 the estimates of δ and *d* obtained by the software produces predictions that adequately fit the evolution of mean fitness through generations in the simulated lines, both in the absence and in the presence of purging (Figure 2).

Summarizing, we present a version of the IP model that analyzes individual fitness data for pedigreed individuals and is able to detect purging and to estimate genetic parameter that are useful to predict the joint consequences of inbreeding and purging. However, it is necessary to explore the properties of this approach more extensively through the analysis of simulated data with different rates of inbreeding and with different distributions of the *h* and *s* values of deleterious mutations. Furthermore, it would be useful to compare its performance with that of previous methods based on ancestral inbreeding, and to characterize the possible biases of our method regarding the estimates of *d* and δ caused by the approximate nature of our IP model, as well as their predictive implications. This exploration needs to analyze a wide range of simulated situations, including different population sizes, generation numbers and distributions of the deleterious effects, and will be addressed in a different paper.

Acknowledgements

We thank Armando Caballero Rúa for his useful comments on the optimization of the software and to Diego Bersabé for providing simulated lines to analyze. This work was funded by grant CGL2015-53274-P and Formación del Personal Investigador (FPI) research fellowship (BES-2012-055006) from Ministerio de Economía y Competitividad MINECO.

Literature Cited

Ballou, J. D. 1997. Ancestral inbreeding only minimally affects inbreeding depression in mammalian populations. *Journal of Heredity 88*(3): 169-178.

Bersabé, D. and A. García-Dorado. 2013. On the genetic parameter determining the efficiency of purging: an estimate for Drosophila egg-to-pupae viability. *Journal of evolutionary biology 26(2): 375-385.*

Bersabé, D., A. Caballero, A. Pérez-Figueroa, and A. García-Dorado. 2016. On the consequences of purging and linkage on fitness and genetic diversity. G3 6(1): 171-181.

Boakes, E. H. and J. Wang. 2005. A simulation study on detecting purging of inbreeding depression in captive populations. *Genetical Research 86*(02): 139-158.

Boakes, E. H., J. Wang, and W. Amos. 2007. An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity 98*(3): 172-182.

Ceballos, F. C., and G. Álvarez, 2013. Royal dynasties as human inbreeding laboratories: the Habsburgs. *Heredity 111*(2): 114-121.

Falconer, D. S., and T. Mackay. 1996. Introduction to quantitative genetics. Ed. 4th. Longman, Essex, England.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation 126*: 131-150.

García-Dorado, A. 2012. Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics 190*(4): 1561-1576.

García-Dorado, A. 2015. On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. *Heredity 115*, 185-187.

Hedrick, P. W. 1994. Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity* 73: 363-372.

Hedrick, P. W. and S. T. Kalinowski. 2000. Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics* 31: 139-162.

Hedrick, P. W., U. Hellsten, and D. Grattapaglia. 2016. Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. *New Phytologist 209*(2): 600-611.

Karaboga, D. and B. Basturk. 2007. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of Global Optimization* 39: 459-471.

Karaboga, D., B. Gorkemli, C. Ozturk, and N. Karaboga. 2014. A comprehensive survey : artificial bee colony (ABC) algorithm and applications. *Artificial Intelligence Review* 42 (1): 21-57.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. *Trends in Ecology and Evolution 17*(5): 230-241.

Kennedy, E. S., C. E. Grueber, R. P. Duncan, and I. G. Jamieson. 2014. Severe inbreeding depression and no evidence of purging in an extremely inbred wild species—the Chatham Island black robin. *Evolution 68*(4): 987-995.

Kruuk, L. E., B. C. Sheldon, and J. Merilä. 2002. Severe inbreeding depression in collared flycatchers (*Ficedula albicollis*). *Proceedings of the Royal Society of London. Series B: Biological Sciences 269*(1500): 1581-1589.

Lacy R. C., and J. D. Ballou. 1998. Effectiveness of selection in reducing the genetic load in populations of *Peromyscus polionotus* during generations of inbreeding. *Evolution* 52: 900-909.

López-Cortegano, E., A. Vilas, A. Caballero, and A. García-Dorado. 2016. Estimation of genetic purging under competitive conditions. Evolution 70: 1856–1870.

Liberg, O. H., Andrén, H.-C., Pedersen, H. Sand, D. Sejberg *et al*. 2005. Severe inbreeding depression in a wild wolf *Canis lupus* population. Biology letters *1*(1): 17-20.

Malécot, G. 1948. Les mathématiques de l'hérédité. Paris: Masson & Cie., Paris.

Morton, N. E., J. F. Crow, H. J. Muller. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. *PNAS USA* 42: 855-63.

O'Grady, J. J., B. W. Brook, D. H. Reed, J. D. Ballou, D. W. Tonkyn, and R. Frankham. 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation 133*(1): 42-51.

Press, W. H. *et al*. 1992. Minimization or maximization of functions in Numerical recipes in C. Cambridge University Press, Cambridge, pp. 394-455.

Saccheri I, M. Kuussaari, M. Kankare, P. Vikman, W. Fortelius, and I. Hanski. 1998. Inbreeding and extinction in a butterfly metapopulation. *Nature* 392: 491-494.

Swindell, W. R., and J. L. Bouzat. 2006. Ancestral inbreeding reduces the magnitude of inbreeding depression in *Drosophila melanogaster*. *Evolution 60*(4): 762-767.

Templeton, A. R. and B. Read. 1984. The elimination of inbreeding depression in a captive herd of Speke's gazelle. Zoo Biology 3: 177-199.

Walling, C. A., D. H. Nussey, A. Morris, T. H. Clutton-Brock, L. E. Kruuk, and J. M. Pemberton. 2011. Inbreeding depression in red deer calves. *BMC Evolutionary Biology 11*(1): 318.

CAPÍTULO TERCERO

ON THE PREDICTIVE VALUE OF THE EFFECTIVE PURGING COEFFICIENT ESTIMATED FROM PEDIGREED POPULATIONS

Eugenio López-Cortegano[1], Diego Bersabé[1], Jinliang Wang[2] and Aurora García-Dorado[1]

[1] Departamento de Genética. Facultad de Biología. Universidad Complutense. Madrid 28040. Spain.

[2] Institute of Zoology. Zoological Society of London. London NW1 4RY. United Kingdom.

Running title: Purging analysis of pedigreed data

Key words: Inbreeding depression; Inbreeding-Purging model; ancestral inbreeding; PURGd software; simulation.

Corresponding author:

Aurora García-Dorado

Departamento de Genética, Facultad de Biología, Universidad Complutense, Ciudad Universitaria, Madrid 28040, Spain.

Phone number: 0034 913 944 975

Email address: augardo@ucm.es

Abstract

The consequences of inbreeding on fitness are of main importance in evolutionary and conservation biology, but can critically depend on genetic purging. However, estimating purging has proven elusive. We assay the performance of the Inbreeding-Purging (IP) model and of models based on ancestral inbreeding to detect and estimate purging from fitness data in simulated pedigreed populations, and explore the reliability of the predictions obtained using these estimates. First, we compare several estimation methodologies and conclude that numerical non-linear regression is to be preferred to linear regression of log-fitness data or to logistic regression for binary fitness data. Using this numerical estimation method, we find that both the IP and Ballou's ancestral inbreeding models have similar power to detect purging from slow inbreeding data, but Ballou's model produces many false positives when based on few generations of quick inbreeding, as it uses to happen for many endangered populations. Both models produce reliable estimates of the rate of inbreeding depression from short-term data, while they give biased estimates from data of long lasting inbreeding processes. However, IP estimates have smaller standard errors. Under the IP model, data from long lasting inbreeding processes gives downwardly biased estimates for both the rate of inbreeding depression and the purging parameter, but these biases cancel each other so that the joint estimates produce quite reliable predictions for the evolution of mean fitness. Thus, using the estimates obtained into the corresponding model, we find that Ballou's model produces quite erratic predictions, while IP predictions are accurate as far as the population size is not too small.

Introduction

Inbreeding depression is a major threat to the survival of small endangered populations. It is mainly due to the increase in the frequency of homozygous genotypes for recessive deleterious alleles, which leads to fitness decay and can boost the risk of extinction (Lande 1994, Hedrick & Kalinowski 2000, O'Grady *et al*. 2006, Charlesworth & Willis 2009). However, under inbreeding, selection can be more efficient, as deleterious recessives that normally escape selection can be purged when exposed in homozygosis, resulting in a reduction of fitness depression and, potentially, in some fitness recovery (García-Dorado 2012, García-Dorado 2015).

While important inbreeding depression has been often demonstrated (Crnokrak & Roff 1999, O'Grady *et al.* 2006), there is less evidence on the effect of genetic purging, and many studies have failed to detect it, both in wild or captive populations, or it's magnitude has been found to be small (Ballou 1997, Bryant *et al.* 1999, Boakes *et al*. 2007, Kennedy *et al.* 2014). Nonetheless, failure to detect purging does not mean purging is irrelevant in actual populations, as it may pass undetected in many situations (Hedrick & García-Dorado2016, López-Cortegano *et al.* 2016). Developing methods and tools to detect and evaluate purging is of critical importance in conservation, as it may help to improve management policies.

The first models aimed to detect purging on fitness from pedigree data used an ancestral inbreeding coefficient (F_a) in the purging term (Ballou 1997, Boakes *et al*. 2006). This coefficient, first described by Ballou (1997), represents the average proportion of an individual's genome that has been in homozygosis by descent at least once in one ancestor. Its interest in purging analysis is that recessive deleterious alleles can be purged in inbred ancestors, so that individuals with higher F_a are expected to carry

fewer such alleles than those with the same level of inbreeding and lower F_a values, and should therefore have higher fitness.

More recently, an Inbreeding-Purging (IP) model has been developed that produces good predictions for the joint consequences of inbreeding and purging. This model predicts the evolution of mean fitness under inbreeding as a function of the "purged inbreeding coefficient" (g), which represents Wright's inbreeding coefficient adjusted for the reduction in frequency of the deleterious alleles ascribed to purging, so that it can be used to predict the increase in homozygosis for these alleles. This purged inbreeding coefficient g can be predicted using a purging coefficient (d) that represents the enhancement of selection under inbreeding (García-Dorado 2012). The model was derived under the assumption of a constant d value across loci, so that d represented the recessive component of the deleterious effect. However, using extensive simulation it was shown that, when d varies across loci, reliable predictions can also be obtained by using an empirically defined effective purging coefficient. This IP model predicts the evolution of mean fitness and of inbreeding load (B) in the population. Thus, using this model, the effective purging coefficient has been estimated from the evolution of mean fitness in experiments carried out with the fruit fly *Drosophila melanogaster*, the IP model giving a much better explanation of such evolution than a model without purging (Bersabé & García-Dorado 2013, López-Cortegano *et al*. 2016). Furthermore, equations have been derived in order to apply the IP model to predict the fitness of pedigreed individuals and, using this approach, a free software package PURGd has been developed that analyzes such pedigreed data to estimate IP parameters, *i.e.*, the rate of inbreeding depression δ (which represents the rate at which fitness would decline with increasing inbreeding in the absence of purging) and the effective purging coefficient d (García-Dorado 2012, García-Dorado *et al.* 2016).

However, it should be pointed out that the IP model produces very good but not exact predictions, particularly when *d* varies across loci so that predictions are based on the empirical effective purging coefficient. This implies that maximizing the fit of IP predictions to the available data can lead to some bias for the estimates of IP parameters (δ and *d*). The reliability of the method depends on whether these estimates, despite their possible bias, produce reasonable IP predictions for scenarios that are different from those characterizing the data used in the estimation process.

Here, we first analyze fitness data of simulated pedigreed individuals undergoing inbreeding and purging, in order to investigate how often the IP and F_a-based approaches allow detecting purging and the extension to which the estimates of the model's parameter depend upon the rate of increase of inbreeding (here determined by the population size *N*) and the length of the inbreeding period (*t*) characterizing the data. Then, we explore how reliable are both IP and F_a-based predictions for (*N*, *t*) scenarios different from those used to estimate the model´s parameters.

Material and Methods

THE SIMULATED POPULATIONS

Simulations were performed under a mutation-selection-drift (MSD) scenario, where a population of size $N = 10^3$ is simulated over 10^4 generations to obtain a base population that can be assumed to be at the MSD balance. Mutations occur at a rate λ per genome and generation, and have selection coefficient s and degree of dominance h, so that fitness is reduced by $h \cdot s$ and s when the mutant allele is in heterozygosis and homozygosis, respectively. Fitness is multiplicative across loci. Details on the program

are described in Bersabé *et al*. 2016.

Two main cases are considered that roughly account for the larger inbreeding load detected in the wild, compared to that of captive populations (Ralls *et al*. 1988, O'Grady *et al.* 2006, Hedrick and García-Dorado 2016). In both cases, a variable selection coefficient is sampled from a gamma distribution with shape parameter $\alpha = 3^{-1}$ and rate parameter $\beta = \alpha / E(s)$, where E(*s*) stands for the expected *s* value. Then a variable degree of dominance is sampled from a uniform distribution ranging between 0 and $e^{-7.5s}$, as in Pérez-Figueroa *et al*. 2009. In both cases, *s* values larger than 1 were assigned *s* = 1. For the CAPTIVE case, both the mutation rate and the average selection coefficient are lower than in the WILD case, giving a larger average degree of dominance (E(*h*)). These mutational parameters are summarized in Table 1.

	E(*s*)	E(*h*)	λ
CAPTIVE	0.1	0.337	0.1
WILD	0.2	0.283	0.2

Table 1: Genetic parameters used in simulations corresponding to the two different cases studied: CAPTIVE and WILD. These parameters include the expected (E) values of the selection coefficient (*s*, gamma distributed with shape parameter 1/3) and of the degree of dominance (*h.* uniformly distributed between 0 and $e^{-7.5s}$), and the mutation rate (λ).

For each case considered, 10 base populations were simulated. Lines of reduced size *N=10*, *N=25* and *N=50* were generated from the corresponding base populations at the MSD equilibrium (250, 100 and 50 replicates, respectively, the 10 base population contributing equal numbers of replicates for each size). All lines were simulated during *2N* generations following the same protocol as for the base populations (*i.e.*, under mutation, selection and drift), and pedigrees and individual fitness were recorded.

ESTIMATION OF INBREEDING DEPRESSION AND PURGING

IP Model: This model predicts fitness as a function of a purged inbreeding coefficient *g* that is defined as Wright' *F* inbreeding coefficient corrected for the reduction in frequency of deleterious alleles expected from purging. This *g* coefficient is computed as a function of a purging coefficient *d* (García-Dorado 2012). For a model with constant effects across loci, $d=s(1/2-h)$. For models where deleterious effects vary across loci, an effective purging coefficient, here referred to just as purging coefficient *d* for simplicity, is defined empirically as that producing the best predictions when used into the IP equations. The model can predict either the average fitness expected at generation *t* (W_t) or the expected fitness for individual *i* with pedigree records (W_i). García-Dorado *et al*. (2016) give general equations to compute g_i in pedigrees with overlapping generations. In the case of individual fitness,

$$W_i = W_0 \cdot e^{-\delta \cdot g_i} \qquad (1)$$

where δ is the rate of inbreeding depression, g_i is the purged inbreeding coefficient calculated using *d*, and W_0 is the expected fitness in the non-inbred population.

Note that, if natural selection is relaxed during the inbreeding period, *g* can be replaced with *F* and δ equals the inbreeding load B in the base population, computed as the sum over loci of $2s(1/2-h)pq$, as defined by Morton *et al.* (1956), where *p* and *q* are the frequency of the wild and deleterious alleles, respectively. Thus, the inbreeding load B can be interpreted as the rate of inbreeding depression expected in the absence of selection.

However, in the presence of natural selection, purging is taken into account by using *g* instead of *F*. Furthermore, under slow inbreeding, non-purging selection should be

taken into account by using the Full Model approach (FM) in García-Dorado (2012). This approach considers that, in any population, non-purging selection can continuously cancel the inbreeding depression ascribed to the inbreeding load expected at the Mutation-Selection-Drift (MSD) balance. Thus, after a reduction in size, while the population transits to the new MSD balance with equilibrium inbreeding load B*, standard selection cancels the inbreeding depression ascribed to B*. Therefore, to account for the consequences of mutation and non-purging selection during the transition from the original MSD balance in the base population (with inbreeding load B) to the new MSD balance for the population size of the lines (with inbreeding load B*), we use the FM rate of inbreeding depression δ_{FM}= B – B*. Thus, if the size of the lines equates that of the base population, then B = B*, and the Full Model predicts no inbreeding depression (δ_{FM} =0), as expected. For small lines, B* can be neglected and we can use δ = B, as usually assumed. We will compute the inbreeding loads (B or B*) expected at the MSD balance by averaging predictions obtained from Equations 10 and 13 in García-Dorado 2007 over 10^6 *s,h* values sampled from the corresponding joint distribution, where *s* values larger than 1 where assigned *s*=1 as in the simulation process.

For each pedigree generated, we estimated the purging coefficient d and the inbreeding depression rate δ using the PURGd software package (García-Dorado *et al*. 2016). These estimates were obtained running the two methods available in PURGd.

First, we used the linear regression method for log-fitness (LR), which uses log-fitness data to fit the logarithmic transformation of Equation 1:

$$\ln(Wi) = a + b\, g_i\,,$$

where a = E[ln(W_0)] (E stands for expected value). It should be noted that the slope *b* of individual log-fitness on *g* is larger than the rate of inbreeding depression for the expected

fitness (δ). For effects constant across loci and known *d*, an estimate of δ can be computed from *b* as

$$\delta = 2\, d\, b / \ln(1-2d), \qquad (2)$$

(García-Dorado *et al*., 2016). In practice, since *d* varies across loci, we inferred δ using our estimates *d* of the effective purging coefficient in Equation 2.

Second, we used the numerical non-linear regression method (NNLR) with untransformed fitness data to fit predictions from Equation 1 by numerically searching for estimates that minimize the residual sums of squares (RSS). To check for the quality of the numerical algorithm, we estimated the genetic parameters for each pedigree as the result of a single run (x1) and as the average results of five (x5) and ten (x10) independent runs.

The expected value in the non-inbred population ($E(W_0)$ or $E[\ln(W_0)]$, depending on the estimation method) was obtained in a previous step as the mean fitness of non-inbred individuals with non-inbred ancestors ($F = F_a = 0$), as explained in García-Dorado *et al*. 2016).

For each replicate, the statistical significance of the NNLR estimate for *d* was tested against the null hypothesis *d*=0 using bootstrap as follows. The squared residual error was computed for each individual *i* as $e_i^2 = [W_i - E(W_i)]^2$, where W_i is the fitness of individual *i* and $E(W_i)$ is its expected value computed using the IP approach. Two e_i^2 values were obtained. One predicting $E(W_i)$ by using in Equation 1 the estimates of δ and *d* obtained from the same replicate ($e_i^2{}_d$). The other one ($e_i^2{}_0$), using into Equation 1 *d*=0 and the corresponding NNLR estimate of δ (obtained in the same replicate by assuming *d*=0 as a known parameter). Then, we computed the variable $D_i = e_i^2{}_0 - e_i^2{}_d$, with mean MD, that

measures how much the prediction of fitness for individual *i* improves by including purging. Then, in order to infer the distribution of D_i under the null hypothesis we define $Y_i = D_i$-MD, with mean 0, and we obtained 10^4 bootstrap samples for Yi of the same size as the replicate. For each replicate, we decide that the estimate of purging ($d > 0$) significantly improves the fitting to the data compared to a non-purging null hypothesis (*i.e.*, compared to $d = 0$) when the mean for Y was larger than MD in at most 5% if the bootstrap samples. This bootstrap method has been incorporated to PURGd.

Ancestral Inbreeding models: Ballou (1997) defined the ancestral inbreeding coefficient (F_a) as the fraction of an individual's genome that has been in homozygosis by descent in at least one ancestor, calculated in terms of the inbreeding coefficient (F) and the ancestral inbreeding coefficient of the individual's parents (sire S, and dame D) as

$$F_a = \frac{1}{2}\left\{F_{a\,(D)} + \left[1 - F_{a\,(D)}\right] \cdot F_{(D)} + F_{a\,(S)} + \left[1 - F_{a\,(S)}\right] \cdot F_{(S)}\right\} \qquad (3)$$

Thus, F_a is connected with the purging opportunities in the ancestors of an individual. Ballou proposed a linear model to fit the joint effect of inbreeding and purging on individual fitness, given by

$$W_i = W_0 - b_F F + b_{F.Fa} F. F_a ,$$

where b_F is the partial regression coefficient that gives the decline of fitness with increasing inbreeding (F) for any constant value of the product $F{\cdot}Fa$ and, according to Ballou, it represents the rate of inbreeding depression, while the coefficient b_{FFa} measures the increase of fitness in inbred individuals due to reduced inbreeding depression, caused by purging in their ancestors.

Since we use a multiplicative fitness model, we write Ballou´s model as

$$W_i = W_0 \cdot e^{-b_F \cdot F_{(i)} + b_{FF_a} F_{(i)} \cdot F_{a(i)}} \qquad (4)$$

Two additional linear models have been proposed by Boakes and Wang (2005) to analyze purging using ancestral inbreeding. The first of these two models (BW) considers that the effect of purging on fitness does not depend on the level of inbreeding but just on previous purging opportunities. For multiplicative fitness, this model is written as

$$W_i = W_0 \cdot e^{-b_F \cdot F_{(i)} + b_{F_a} F_{a(i)}} \quad , \qquad (5)$$

where the coefficient of the purging term b_{F_a} is the average rate of increase of individual fitness due to the opportunities of purging in the ancestors.

Finally, Boakes and Wang also proposed a mixed "Ballou-Boakes & and Wang" model (here B-BW) (2005) where the purging term is the sum of those in Ballou and BW models, giving

$$W_i = W_0 \cdot e^{-b_F \cdot F_{(i)} + b_{F_a} F_{a(i)} + b_{FF_a} F_{(i)} \cdot F_{a(i)}}. \qquad (6)$$

Both Ballou and Boakes and Wang tested their models fitting dichotomical (0, 1) fitness data, and used logistic regression in their analysis. However, in order to compare ancestral inbreeding and IP approaches under similarly optimum conditions, we simulate and analyze fitness data as a continuous variable defined in the interval (0, 1). When linearizing F_a-based models by taking logarithms, δ cannot be inferred from the b_F estimate obtained by fitting Equation 2, as F_a -based approaches do not give an estimate for *d*. Therefore, we use the NNLR method to directly fit the predictions of the above exponential equations to untransformed fitness data. We have also analyzed dichotomical fitness data using Ballou´s model both with the NNLR and the Logistic methods (see Supplementary Material).

A bootstrap contrast analogous to that performed in the NNLR IP analysis was used

in each replicate to test the significance of purging in Ballou's analysis. Thus, squared residuals ($e_i^2 = [W_i - E(W_i)]^2$) were obtained computing $E(W_i)$ using into Equation 4 the estimates of b_F and b_{FFa} obtained in the replicate ($e_i^2{}_b$), or using $b_{FFa}=0$ and the corresponding estimate of b_F ($e_i^2{}_0$). Then bootstrap was performed for the mean of the variable $D_i = e_i^2{}_0 - e_i^2{}_b$, as in the IP model. Significant purging was detected in the replicates with $b_{FFa} > 0$ where at most 5% of the bootstrap samples had mean values larger than the mean of $D_i = e_i^2{}_0 - e_i^2{}_b$ obtained in the replicate.

Non-Linear Regression coefficients for F_a-based models, as well as bootstrap errors, were computed using an update of PURGd. As in the case of the IP model, the intercept was obtained in a previous step as the mean fitness for non-inbred individuals with non-inbred ancestors ($F = F_a = 0$).

ANALYSIS OF THE PREDICTIVE VALUE OF THE ESTIMATES

In order to evaluate the predictive value of the parameters estimated in the previous section, we use estimates obtained from different numbers of generations (t) in lines of different size (N), to predict the evolution of average fitness in lines maintained with different population sizes (crossed predictions). We check how these predictions fit the corresponding simulated data by graphically comparing the observed and predicted evolution of mean fitness.

In the case of the IP model, predictions of the expected fitness at generation t (W_t) are computed using the equation for the evolution of mean fitness, obtained by replacing W_i and g_i in Equation 1 with their expected values at generation t (W_t and g_t). For this purpose, g_t is computed as a function of N using the expression provided in García-

Dorado (2012) and, when using LR estimates, δ is inferred using the estimates of *b* and *d* into Equation 2. A neutral prediction is also obtained by replacing g_t with the standard inbreeding coefficient ($F(t)$) into Equation 1 and using the inbreeding load computed in the simulated population ($\delta = B_{SIM}$).

In the case of models based on ancestral inbreeding, predictions for mean fitness are obtained replacing F_i and F_{ai} in Equations 4-6 with their expected value through generations, $F(_t)$ and $F_{a(t)}$. Below we derive an expression for the expected evolution of ancestral inbreeding through generations in a panmictic populations maintained with effective size *N*.

From equation 3, since the expected F_a values (or *F* values) are the same for sires as for dams, the average ancestral inbreeding at generation *t* can be computed by iterating the expression

$$F_{a\,(t)} = F_{a\,(t-1)} + \left[1 - F_{a\,(t-1)}\right] \cdot F_{(t-1)}\ ,$$

which, noting that $F_t = 1 - \left(1 - \frac{1}{2N}\right)^t$ and rearranging, gives

$$F_{a\,(t)} = 1 - \left(1 - \frac{1}{2N}\right)^{t-1} \cdot \left[1 - F_{a\,(t-1)}\right] \qquad (7)$$

In addition, an expression giving the expected ancestral inbreeding after *t* generations can be derived. For simplicity, we define $x_t = 1 - F_{a\,(t)}$ and $k = \left(1 - \frac{1}{2N}\right)$, so that Equation 7 can be written as $x_t = x_{t-1} \cdot k^{t-1}$. Therefore, since x_0=1, the expected value of x_t can be computed as

$$x_t = x_o \prod_{i=0}^{t-1} k^i \ = k^{\sum_{i=0}^{t-1} t} = k^{\,t\,(t-1)/2}$$

and, replacing x_t and *k* into this expression and rearranging, we obtain

$$F_{a\,(t)} = 1 - \left(1 - \frac{1}{2N}\right)^{\frac{1}{2}t(t-1)} \qquad (8)$$

Results

IP ESTIMATES OF THE RATE OF INBREEDING DEPRESSION AND THE PURGING COEFFICIENT

The inbreeding loads in the simulated base populations (B_{SIM}) were close to the corresponding expectations (B) (Tables 2 and 3). For N=10, we obtain $\delta_{FM} \approx B$, as usually assumed. However, δ_{FM} declines when larger sizes are considered and, in agreement with this prediction, the estimates of δ in Tables 2 and 3 also show a reduction for larger lines. In general, the estimates of δ are close to their expected values (δ_{FM}) when based in N/2 generations, standard errors being smaller for NNLR than for LR estimates, but δ estimates based in longer periods become downwardly biased.

Both the LR and the NNLR methods produce large estimates of *d* indicating substantial purging (Tables 2 and 3 for the CAPTIVE and WILD cases, respectively). There is a trend for a reduction of *d* when estimated from longer inbreeding periods, which is associated to a parallel reduction in the estimate of δ. As expected, the estimates

CAPTIVE		B	B_{sim}	δ_{FM}	δ_{LR}	δ_{NNLR}	d_{LR}	d_{NNLR}	$d_{(\delta_{FM})}$
	t=N/2				0.5388 ± 0.0282	0.5667 ± 0.0185	0.2563 ± 0.0145	0.2572 ± 0.0136	0. 2856 ± 0.0144
N=10	t=N	0.6266	0.5828 ± 0.0144	0.5540	0.5090 ± 0.0234	0.5511 ± 0.0166	0.2629 ± 0.0127	0.2568 ± 0.0121	0. 2860 ± 0.0119
	t=2N				0.4941 ± 0.0209	0.4955 ± 0.0149	0.2226 ± 0.0107	0.1981 ± 0.0099	0. 2492 ± 0.0103
	t=N/2				0.4311 ± 0.0315	0.5152 ± 0.0205	0.3132 ± 0.0199	0.3065 ± 0.0193	0.3018 ± 0.0212
N=25	t=N	0.6266	0.5828 ± 0.0144	0.5006	0.4332 ± 0.0290	0.4784 ± 0.0212	0.2753 ± 0.0171	0.2553 ± 0.0172	0. 2956 ± 0.0173
	t=2N				0.4155 ± 0.0274	0.4046 ± 0.0187	0.2048 ± 0.0174	0.1902 ± 0.0167	0. 2551 ± 0.0168
	t=N/2				0.4048 ± 0.0302	0.5004 ± 0.0266	0.3296 ± 0.0245	0.2915 ± 0.0247	0. 2781 ± 0.0281
N=50	t=N	0.6266	0.5828 ± 0.0144	0.4448	0.4261 ± 0.0280	0.4352 ± 0.0234	0.2339 ± 0.0234	0.2018 ± 0.0216	0. 2371 ± 0.0221
	t=2N				0.3972 ± 0.0271	0.3745 0.0195	0.1495 ± 0.0190	0.1499 0.0199	0. 1958 ± 0.0201

Table 2. Estimates of rates of inbreeding depression and purging coefficients in the CAPTIVE case from lines of different sizes (N) and different numbers of generations (*t*). Estimates are averaged over replicates, and are given with their empirical standard errors. This table gives the expected (B) and observed (B_{SIM}) inbreeding load in the base population, and the Full-Model rate of inbreeding depression expected in the lines (δ_{FM}) together with the corresponding PURGd estimates obtained from the LR or NNLR methods (δ_{LR} and δ_{NNLR}). It also gives the corresponding estimates of the purging coefficient (d_{LR} and d_{NNLR}). A NNLR estimate of *d* is also obtained by forcing PURGd to use δ_{FM} as the known rate of inbreeding depression ($d(\delta_{FM})$).

WILD		B	B_{sim}	δ_{FM}	δ_{LR}	δ_{NNLR}	d_{LR}	d_{NNLR}	$d_{(\delta_{FM})}$
	t=N/2				1.8004 ± 0.0461	2.2899 ± 0.0541	0.2976 ± 0.0145	0.3233 ± 0.0131	0.3476 ± 0.0130
N=10	t=N	2.5511	2.5370 ± 0.0460	2.2846	1.6368 ± 0.0653	2.1213 ± 0.0464	0.3130 ± 0.0112	0.3099 ± 0.0099	0.3650 ± 0.0092
	t=2N				1.6547 ± 0.0600	1.8043 ± 0.0392	0.2459 ± 0.0092	0.2196 ± 0.0076	0.3015 ± 0.0082
	t=N/2				1.3330 ± 0.0916	2.0721 ± 0.0574	0.3932 ± 0.0146	0.4108 ± 0.0111	0.4239 ± 0.0110
N=25	t=N	2.5511	2.5370 ± 0.0460	2.0926	1.4348 ± 0.0812	1.8381 ± 0.0519	0.3409 ± 0.0129	0.3191 ± 0.0122	0.3867 ± 0.0102
	t=2N				1.4333 ± 0.0752	1.4282 ± 0.0461	0.2513 ± 0.0140	0.2050 ± 0.0116	0.3221 ± 0.0113
	t=N/2				1.1489 ± 0.0830	1.8686 ± 0.0626	0.4022 ± 0.0165	0.3954 ± 0.0159	0.4036 ± 0.0152
N=50	t=N	2.5511	2.5370 ± 0.0460	1.8861	1.2683 ± 0.0797	1.6301 ± 0.0527	0.3283 ± 0.0200	0.3116 ± 0.0179	0.3675 ± 0.0158
	t=2N				1.2512 ± 0.0863	1.4010 ± 0.0632	0.2684 ± 0.0215	0.2539 ± 0.0218	0.3389 ± 0.0177

Table 3. Estimates of rates of inbreeding depression and purging coefficients in the WILD case from lines of different sizes (N) and different numbers of generations (*t*), Estimates are averaged over replicates, and are given with their empirical standard errors .This table gives the expected (B) and observed (B_{SIM}) inbreeding load in the base population, and the Full-Model rate of inbreeding depression expected in the lines (δ_{FM}) together with the corresponding PURGd estimates obtained from the LR or NNLR methods (δ_{LR} and δ_{NNLR}). It also gives the corresponding estimates of the purging coefficient (d_{LR} and d_{NNLR}). A NNLR estimate of *d* is also obtained by forcing PURGd to use δ_{FM} as the known rate of inbreeding depression ($d(\delta_{FM})$)

of this purging parameter are always larger in the WILD case than in the CAPTIVE one. In the NNLR method, estimates are very similar regardless of the number of runs averaged (results not shown). Thus, no more than one run should be needed to estimate purging parameters, though this may change if additional factors were included (*i.e.* environmental factors) adding dimensions and complexity to the model. The estimates presented here were obtained from just one run.

In addition we have also estimated the purging coefficient by forcing PURGd to use δ_{FM} as the known rate of inbreeding depression (also shown in Tables 2 and 3). These estimates are only obtained using the NNLR method, as the expected value of the coefficient *b* for individual log-fitness on *g* is larger than δ_{FM}. It is interesting that these NNLR estimates of *d* obtained from lines of different sizes or from different numbers of generations are more consistent than when both *d* and δ are jointly estimated from the data, and are more similar to those obtained by jointly estimating *d* and δ using data from *t=N/2* generations. This suggests that bias due to overfitting can be reduced if an unbiased estimate of δ can be obtained independently.

ESTIMATES OF THE COEFFICIENTS IN ANCESTRAL INBREEDING MODELS

Tables 4 and 5 show the estimates of non-linear regression coefficients for F_a-based models, obtained using the NNLR method. In both Ballou´s and B-BW models, $-b_F$ estimated from short term data for different population sizes (*N*) gives reasonable estimates of the expected rate of inbreeding depression (δ_{FM}), although standard errors

CAPTIVE		Ballou		BW		B-BW		
		b_F	b_{FF_a}	b_F	b_{F_a}	b_F	b_{FF_a}	b_{F_a}
	t=N/2	-0.5529 ± 0.0217	0.1529 ± 0.0842	-0.5396 ± 0.0185	0.0410 ± 0.0119	-0.5556 ± 0.0219	0.0562 ± 0.1167	0.0325 ± 0.0159
N=10	t=N	-0.5687 ± 0.0202	0.2888± 0.0279	-0.5151 ± 0.0157	0.0777 ± 0.0074	-0.5847 ± 0.0198	0.2114 ± 0.0327	0.0381 ± 0.0082
	t=2N	-0.6247 ± 0.0214	0.4040 ± 0.0222	-0.3536 ± 0.0113	0.0565 ± 0.0064	-0.6163 ± 0.0212	0.3921 ± 0.0250	0.0010 ± 0.0073
	t=N/2	-0.5757 ± 0.0246	0.4655 ± 0.0398	-0.5000 ± 0.0173	0.0517 ± 0.0072	-0.5774 ± 0.0238	0.2832 ± 0.0500	0.0212 ± 0.0097
N=25	t=N	-0.6006 ± 0.0282	0.4139 ± 0.0261	-0.3982 ± 0.0138	0.0601 ± 0.0055	-0.6232 ± 0.0268	0.3678 ± 0.0319	0.0237 ± 0.0075
	t=2N	-0.6885 ± 0.0411	0.5644 ± 0.0397	-0.1393 ± 0.0112	0.0000 ± 0.0059	-0.6248 ± 0.0324	0.5319 ± 0.0356	-0.0160 ± 0.0076
	t=N/2	-0.5506 ± 0.0361	0.3265 ± 0.0434	-0.4965 ± 0.0206	0.0504 ± 0.0057	-0.6096 ± 0.0392	0.2434 ± 0.0569	0.0319 ± 0.0080
N=50	t=N	-0.6534 ± 0.0458	0.5009 ± 0.0440	-0.2430 ± 0.0149	0.0212 ± 0.0052	-0.7498 ± 0.0563	0.5476 ± 0.0575	0.0167 ± 0.0061
	t=2N	-0.7228 ± 0.0515	0.6377 ± 0.0523	-0.0575 ± 0.0105	-0.0176 ± 0.0054	-0.6363 ± 0.0608	0.5961 ± 0.0595	-0.0222 ± 0.0077

Table 4: Non-linear regression coefficients estimated in the CAPTIVE case for Ballou's model (B), BW model and B-BW model using the NNLR method in pedigrees of different populations sizes (*N=10, N=25* and *N=50*) and numbers of generations (**$t = N/2$, $t = N$** and **$t = 2N$**). Estimates are averaged over replicates, and are given with their empirical standard errors.

WILD		Ballou		BW		B-BW		
		b_F	b_{FF_a}	b_F	b_{F_a}	b_F	b_{FF_a}	b_{F_a}
	t=N/2	-2.4140 ± 0.0657	2.0244 ± 0.2515	-2.2974 ± 0.0581	0.3210 ± 0.0438	-2.4481 ± 0.0678	1.3741 ± 0.2922	0.1763 ± 0.0531
N=10	t=N	-2.2623 ± 0.0554	1.4595 ± 0.0658	-2.0940 ± 0.0468	0.4605 ± 0.0222	-2.4021 ± 0.0579	0.9985 ± 0.0936	0.2456 ± 0.0297
	t=2N	-2.5070 ± 0.0663	1.9002 ± 0.0648	-1.2819 ± 0.0301	0.3079 ± 0.0154	-2.5667 ± 0.0637	1.8801 ± 0.0702	0.0465 ± 0.0232
	t=N/2	-2.2362 ± 0.0660	1.7462 ± 0.0914	-2.0815 ± 0.0542	0.3174 ± 0.0150	-2.3632 ± 0.0701	1.0976 ± 0.1221	0.1932 ± 0.0217
N=25	t=N	-2.2581 ± 0.0731	1.7819 ± 0.0732	-1.4159 ± 0.0336	0.2771 ± 0.0124	-2.6589 ± 0.0781	1.7863 ± 0.0782	0.1499 ± 0.0165
	t=2N	-2.5705 ± 0.0806	2.2831 ± 0.0799	-0.3195 ± 0.0193	-0.0195 ± 0.0099	-2.4440 ± 0.0931	2.2108 ± 0.0930	-0.0330 ± 0.0147
	t=N/2	-2.1444 ± 0.0805	1.6447 ± 0.0899	-1.9312 ± 0.0525	0.2697 ± 0.0133	-2.5151 ± 0.0929	1.1384 ± 0.1004	0.1994 ± 0.0184
N=50	t=N	-2.4045 ± 0.0966	2.1059 ± 0.0941	-0.6156 ± 0.0271	0.0709 ± 0.0102	-2.6501 ± 0.1028	2.1502 ± 0.1070	0.0678 ± 0.0149
	t=2N	-2.6496 ± 0.1065	2.4997 ± 0.1066	-0.0908 ± 0.0144	-0.0448 ± 0.0089	-2.4896 ± 0.1217	2.4214 ± 0.1323	-0.0421 ± 0.0128

Table 5: Non-linear regression coefficients estimated in the WILD case for Ballou's model (B), BW model and B-BW model using the NNLR method in pedigrees of different populations sizes (*N=10*, *N=25* and *N=50*) and numbers of generations ($t = N/2$, $t = N$ and $t = 2N$). Estimates are averaged over replicates, and are given with their empirical standard errors.

are larger than in the IP model. These estimates tend to increase when based in more generations of inbreeding, leading to values well above δ_{FM} in the WILD case.

The estimates of the coefficients for terms including F_a usually take positive values indicating purging, but vary depending on N and t in an unpredictable way, particularly for BW and B-BW models where b_{Fa} takes even negative values in some cases.

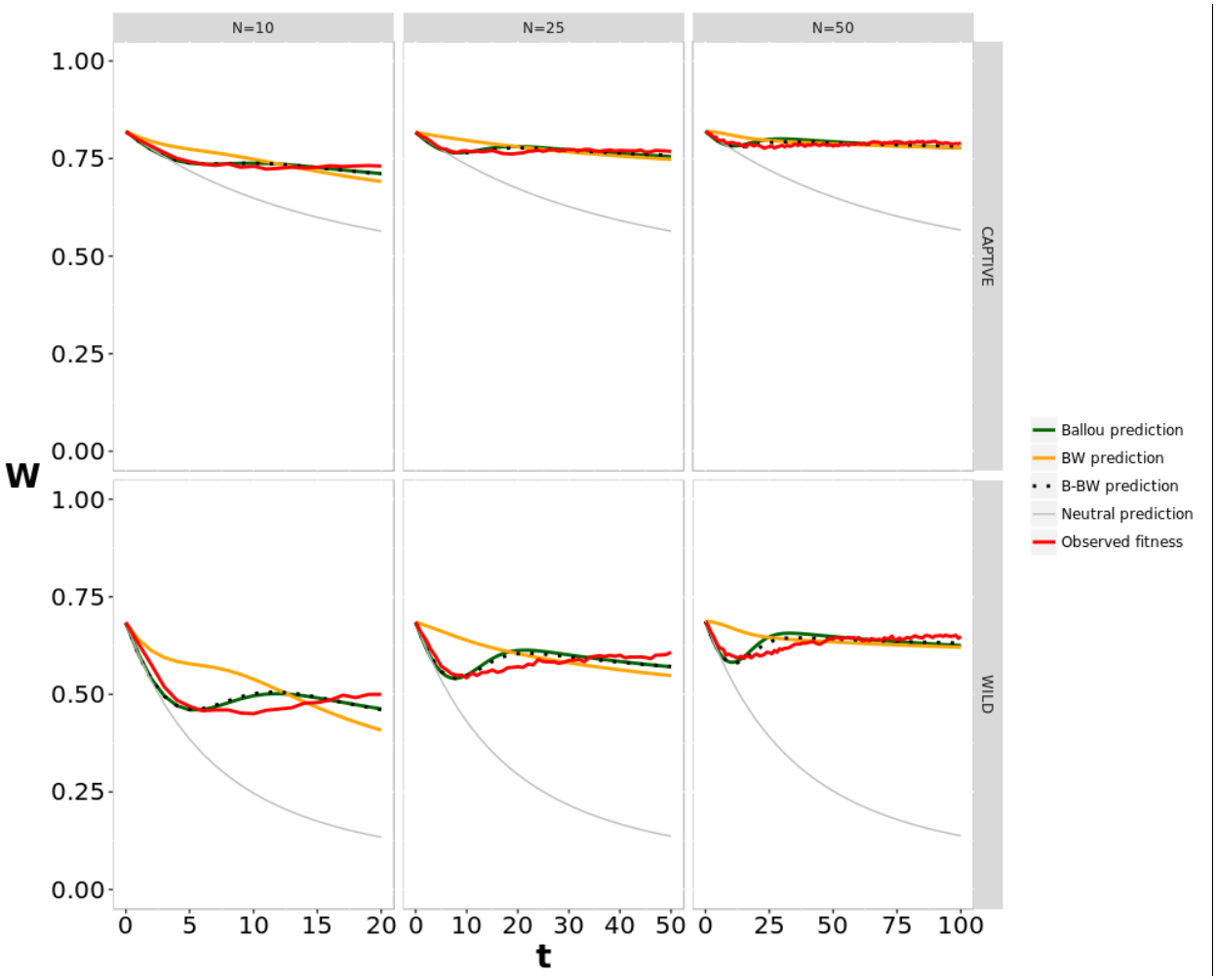


Figure 1: Evolution of mean fitness in simulated lines (red) and the corresponding predictions obtained using F_a-based models. Predictions are computed for two different cases, CAPTIVE and WILD, and three different population sizes (10, 25 and 50) over 2N generations using the coefficients estimated from the same lines and number of generations. Three models based on ancestral inbreeding are used: Ballou's (green), BW (yellow) and B-BW model (black dotted), as well as a prediction without selection (grey).

Figure 1 illustrates how different F_a -based models fit the data for lines of different sizes, by showing the observed evolution of fitness during *2N* generations together with the corresponding predictions computed using coefficients estimated from the same data. BW model fits the data poorly, showing a systematic overestimation of fitness during the first *N* generations and an increasing underestimation later, while Ballou's model fitting is remarkably good. B-BW model does not improve fitting over Ballou's one, which is not surprising as b_{Fa} estimates are usually small. Therefore, hereafter we will use Ballou's model to evaluate the predictive value of F_a-based methods.

THE EFFICIENCY OF IP AND BALLOU'S MODELS TO DETECT PURGING

Table 6 gives the percent of replicates were a model including purging fitted the data significantly better than a non-purging model, both for the IP or Ballou approaches. Results are from bootstrap contrasts on NNLR estimates, as this estimation method gives lower standard errors than LR analysis and produces estimates for the rate of inbreeding depression (δ). For both models, purging detection is more likely in larger lines and when larger periods are available, as expected from more efficient purging and a larger sample sizes. Detection is also more likely for the WILD than for the CAPTIVE case, as there are more mutations with large *d* values.

Under both IP and Ballou's models, the proportion of detected cases in the more adverse situation (*N*=10, *t=N/2*, CAPTIVE) is smaller than 5%. This suggests that the test is conservative. It also implies that, although both approach detect purging when estimates are averaged over replicates in that adverse situation, they are not able to do so when replicates are separately considered. In the more favorable cases, both IP and Ballou models give substantial detection rates, usually somewhat larger for the former model.

CAPTIVE	IP	BALLOU
N=10 t=N/2	2.4	1.6
N=10 t=N	3.6	4.4
N=10 t=2N	28.8	19.6
N=25 t=N/2	14	8
N=25 t=N	49	29
N=25 t=2N	68	65
N=50 t=N/2	36	26
N=50 t=N	74	58
N=50 t=2N	80	74

WILD	IP	BALLOU
N=10 t=N/2	13.6	5.2
N=10 t=N	31.2	24.8
N=10 t=2N	81.6	74
N=25 t=N/2	52	36
N=25 t=N	95	90
N=25 t=2N	97	99
N=50 t=N/2	94	94
N=50 t=N	100	100
N=50 t=2N	96	100

Table 6: Purging detection. Percent of replicates were a model including purging fitted the data significantly better than a non-purging model. Results are for NNLR analysis under the IP or Ballou approaches, both for CAPTIVE and WILD mutational models (bootstrap contrasts with α=0.05).

THE RELIABILITY OF PREDICTIONS OBTAINED USING ESTIMATES IN IP AND BALLOU'S MODELS

We evaluate the reliability of IP predictions for the evolution of fitness for each set of lines (N=10, 25 or 50) during $2N$ generations using δ and d NNLR estimated from lines maintained with different sizes during different numbers of generations. Figure 2 gives these IP predictions, both for the CAPTIVE and WILD cases, together with the prediction obtained assuming no selection and using the inbreeding load of the base population ($\delta = B_{SIM}$), and with the observed evolution of mean fitness.

IP predictions remain quite accurate up to the first N generations. Although they tend to overestimate fitness in the long term, this bias is usually small, with the exception of N=10 lines in the WILD case. In general, there is a slight trend for long-term fitness

being better predicted using (δ, *d*) estimates from long term data. Furthermore, predictions computed using (δ, *d*) estimates obtained from small lines, where purging tends to be overwhelmed by genetic drift, tend to underrate fitness.

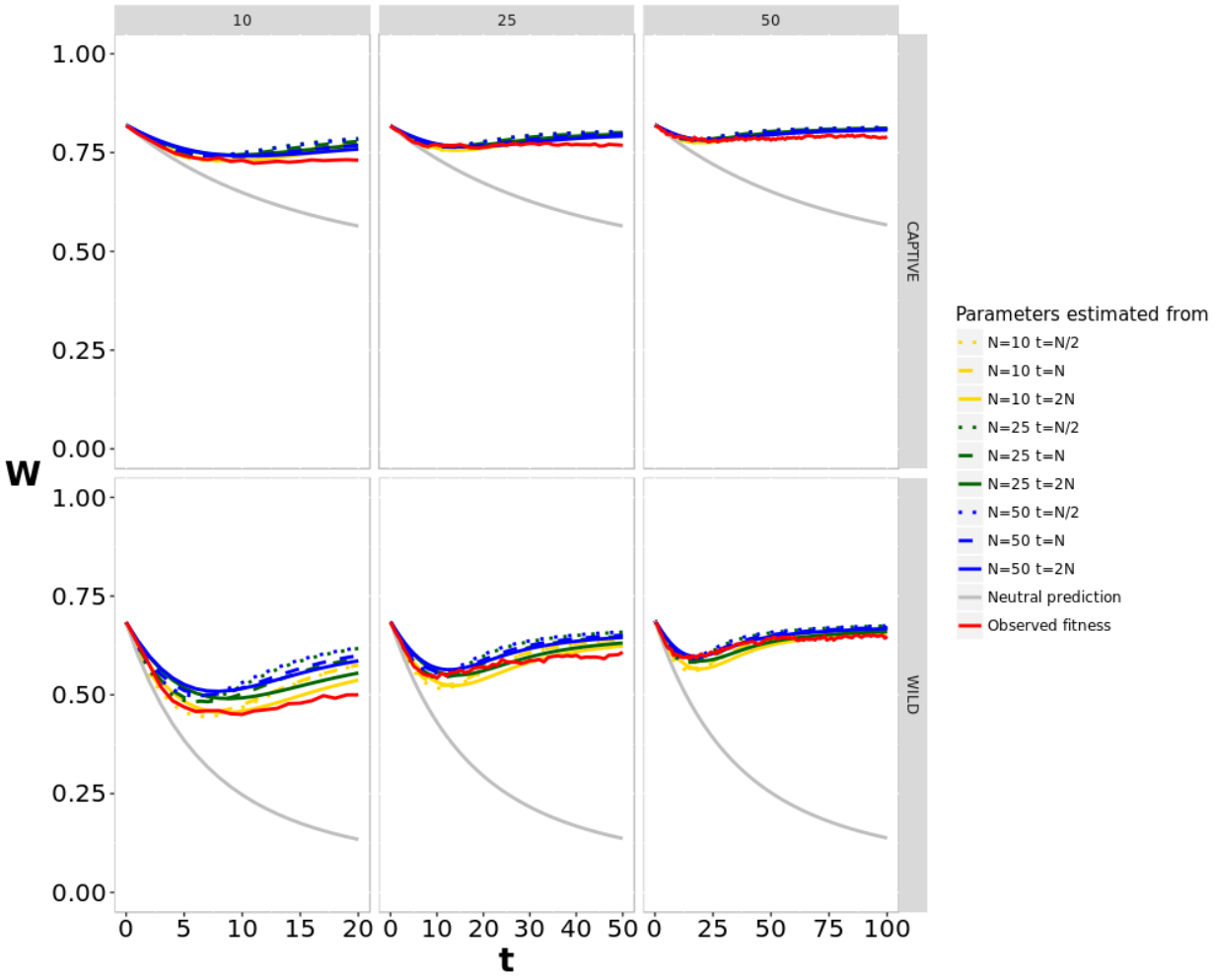


Figure 2: Observed fitness for the CAPTIVE (up) and WILD (down) cases, and the corresponding prediction obtained using NNLR estimates in IP model. In each panel, observed and predicted values over t=2N generations correspond to the population size indicated in the column (N=10, N=25 and N=50), and different predictions are plotted using estimates obtained from different data sets, denoted by different colors and strokes as shown in the lateral panel. Neutral predictions, computed assuming no selection and using the inbreeding load observed in the simulated base population (B_{SIM}) are also shown.

In any case, despite the variability observed between (δ, *d*) estimates obtained from different data sets (Tables 2 and 3), IP predictions remain quite accurate and fit the data

much better than a model assuming no selection, as reductions in the estimate of δ obtained from longer periods are compensated by reductions in the corresponding estimate of *d*.

Similar results obtained using LR estimates are given in the Supplementary Material (Figure S1), showing that NNLR estimates give slightly better fitness predictions than LR ones. Taking into account this result, as well as the smaller standard errors of NNLR IP estimates compared to LR ones, and the fact that the LR analysis of F_a-based models do not give an estimate for δ, only NNLR results will be used to compare IP and F_a based models.

Figure 3 shows a similar evaluation of the reliability of Ballou's predictions computed using NNLR estimates of the coefficients obtained from different data sets (Tables 4 and 5). As expected, the best fitting is obtained when mean fitness is observed in the same data set where the coefficients used to obtain predictions had been estimated. Fitting also improves when estimates are based in longer inbreeding periods. Predictions are generally reliable during the first few generations, where purging is irrelevant. However, they become unreliable later on. Predictions that used parameters estimated in smaller lines underestimate long-term fitness, while those obtained from larger lines tend to overestimate medium-term fitness but can still underestimate fitness later on.

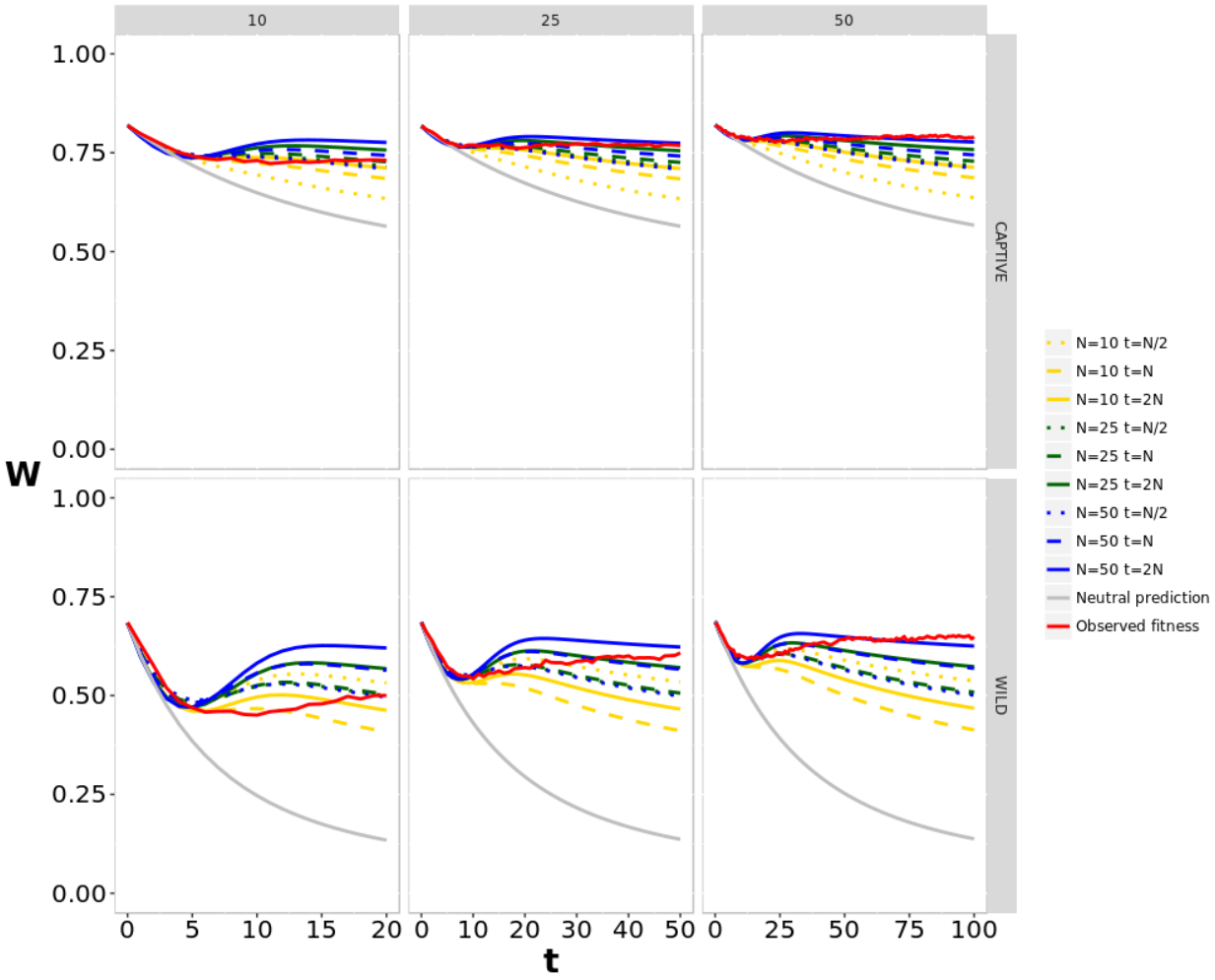


Figure 3: Observed fitness for the CAPTIVE (up) and WILD (down) cases, and the corresponding prediction obtained using NNLR estimates in Ballou's model. In each panel, observed and predicted values over *t=2N* generations correspond to the population size indicated in the column (*N*=10, *N*=25 and *N*=50), and different predictions are plotted using estimates obtained from different data sets, denoted by different colors and strokes as shown in the lateral panel. Neutral predictions, computed assuming no selection and using the inbreeding load observed in the simulated base population (B_{SIM}) are also shown.

Thus, Ballou's predictions are highly dependent on the conditions used to estimate the coefficients of the model, fitness predictions at generation *t=2N* being very erratic. The same analysis was performed for the BW model, giving even less reliable predictions (data not shown).

Comparison of figures 2 and 3 show that IP predictions are more accurate than those

of Ballou's F_a -based model, the IP model being able to reasonably predict the evolution of fitness using parameters estimated under different conditions.

Discussion

We have analyzed the performance of the Inbreeding-Purging model (IP) and of models based on ancestral inbreeding (F_a) to detect, measure and predict purging using simulated fitness data of pedigreed individuals. Simulated populations were maintained with size N=1000 until they reached the MSD balance and then replicated lines were drawn and maintained at smaller size (N=10, 25 or 50) during $2N$ generations. The IP model is based on the expected effect of selection against the recessive component of deleterious effects (d) exposed in homozygotes due to inbreeding, while the F_a approach is based on the statistical fitting of models including inbreeding (F) and ancestral inbreeding (F_a) terms.

THE ESTIMATION METHODS

Since we assume that fitness is multiplicative across loci, both IP and F_a models produce exponential predictive equations for fitness. In order to estimate the parameters of the models we have used PURGd software (García-Dorado *et al.* 2016), which offers two alternative methods. The first one uses linear regression (LR) to fit the linear model obtained by log-transforming the exponential predictive equations. The second one consists on numerically fitting the non-linear exponential equations to untransformed fitness data. This numerical non-linear regression method (NNLR) has the advantage of handling zero values, so that it can analyze binary (0, 1) fitness data. Furthermore, it has the advantage of producing estimates of the rate of inbreeding depression (δ) that can be directly used to predict the evolution of average fitness, while the LR produces estimates

of the slope of log individual fitness on inbreeding (*b*) that are expected to be larger than δ.

The IP approach gives an opportunity to infer δ from LR estimates by adjusting *b* using the estimate of *d*, so we have used this IP approach to compare the LR and NNLR methods. We have found that both methods produce good estimates for δ when based on data from the early phase of the inbreeding process (*t=N/2*), but estimates obtained using NNLR have smaller standard errors (Tables 2 and 3). Obtaining δ estimates from LR analysis by separately adjusting *b* in each replicate, instead of adjusting average *b* with the average *d* estimate, would produce additional downwards bias and larger standard errors. Furthermore, NNLR estimates produce more accurate predictions for the evolution of mean fitness than LR ones (Figures 2 and S1).

Previous investigations of F_a models for individual fitness handled binary (0, 1) data that were analyzed using logistic regression (Ballou 1997, Boakes *et al*. 2007, Ceballos and Álvarez 2013, Kennedy *et al*. 2014). In the Supplementary Material we report an analysis of binary fitness data using Ballou's model, illustrating that, as expected, the coefficient b_F in the logistic model does not estimate the rate of inbreeding depression but gives values much larger than δ. Furthermore, NNLR estimates give slightly better predictions for mean fitness than logistic ones (Figure S2). The analysis also illustrates that binary fitness data leads to less accurate estimates and predictions than the underlying continuous fitness variable, as they imply an important random error in the observation of fitness. Using NNLR to fit the genetic exponential model to untransformed fitness data has the advantage of avoiding the arbitrary (0, 1) codification of fitness that has in occasions been used to allow logistic analysis and that can imply important loss of information.

Therefore, we encourage the use of the NNLR method and, hereafter, we discuss the properties of both IP and F_a models using NNLR estimates obtained from untransformed fitness data.

THE MUTATIONAL MODELS

We have analyzed fitness under two mutational models intended to explore the consequences of purging against the inbreeding load expressed in wild or captive populations. The rational for these mutational models is that, according to available estimates, the inbreeding load in the wild seems to be up to four times larger than in captive populations. Thus, we used a mutational model producing a small inbreeding load at the MSD balance (CAPTIVE case) and another one producing larger inbreeding load (WILD case). The inbreeding loads computed in the simulated base populations (B_{SIM}, Tables 2 and 3) were close to the corresponding expectations (B), and were smaller than the average values reported in the literature (on the average B≈6 and B≈1.5, respectively), particularly in the WILD case (Ralls *et al.* 1988, O'Grady *et al.* 2006, Yun & Agrawal 2014, Hedrick and García-Dorado 2016). It should be noted, however, that our base populations have an effective population size (N=1000) that is relatively small for many wild unthreatened population, and that the mutational parameters of our WILD case predict B≈6 for $N=10^4$. In any case, our purpose is not to obtain realistic predictions of the inbreeding load in the wild, but to evaluate the properties of purging models under different plausible distributions of the deleterious effects. Our estimates of the purging coefficient d in the CAPTIVE case are larger than those estimated in non-competitive conditions for Drosophila (Bersabé and García-Dorado 2013), but the estimate obtained in our WILD case is similar to that obtained in competitive conditions (López-Cortegano *et al.* 2016). In any case, our CAPTIVE and WILD cases parallel the non-competitive and

competitive conditions of those experiment as, both in our two mutational models and in the two Drosophila experiments, the WILD case gives larger inbreeding load but also larger purging coefficient so that, under slow inbreeding, long term inbreeding depression is small in both instances.

PERFORMANCE OF IP AND F_a MODELS

The inbreeding load in the base population (B) represents the rate of inbreeding depression that would be expected in the absence of selection and it also represents the rate of inbreeding depression with increasing *g* in the IP model ($\delta = B$), as this model accounts for purging selection but does not account for non-purging selection. However, in relatively large lines, non-purging selection prevents the expression of a fraction of B. This fraction equates the inbreeding load expected at the MSD balance for the new reduced population size *N*. Thus, according to the Full Model (FM), that accounts for non-purging selection, the expected value for δ in Equation 1 is $\delta_{FM} = B - B^*$, where B* is the inbreeding load at the MSD balance corresponding to the size of the lines (García-Dorado, 2012). Thus, if the size *N* of the lines was close to the size of the base population, δ_{FM} and, therefore, the FM prediction of fitness depression, would approach zero, as it should be expected. For the smaller lines (N=10), we expect $\delta_{FM} \approx B$, as usually assumed. However, δ_{FM} declines when larger lines are considered, due to their larger B*. In agreement with these predictions, the IP estimates of δ obtained using early data from small lines produce good estimates of δ_{FM} under both estimation methods (δ_{LR} and δ_{NNLR}), *i.e.*, δ is close to δFM using up to $t=N$ generations in the CAPTIVE case or $t=N/2$ in the WILD case; Tables 2 and 3). Nevertheless, these estimates show some reduction when larger lines are considered

However, the estimates of δ become downwardly biased when based on longer periods, which is associated to a reduction of the estimates of *d*. The reason is that, for *t* = *2N*, most purging has already occurred during a large proportion of the period considered, and the model overfits long-term data by giving low δ and *d* estimates. More stable estimates of *d* can be obtained by introducing into the model the expected rate of inbreeding depression (δ_{FM}) as a known value for -*b*, as this reduces overfitting (in fact, using the inbreeding load B_{SIM} observed in the base population instead of δ_{FM} makes little difference; results not shown). In practice, δ_{FM} is unknown, but overfitting could be reduced by using an estimate of δ obtained by analyzing early generations in a previous step (about *t=N/2*).

Despite the overfitting described above, each IP joint estimate of δ and *d* produces good predictions for the evolution of mean fitness over the whole range of line sizes, with the exception of the smallest ones (*N*=10) for the WILD case where IP predictions overrate fitness unless (δ, *d*) were also estimated from the same data (*N*=10 lines). Furthermore, (δ, *d*) estimates obtained from *N*=10 lines slightly underrate medium term fitness in larger lines. The reason is that, in the WILD case, there are more mutations with effects small enough to escape selection under important drift. It has been found that drift roughly overwhelms purging for $Nd < 1$ (García-Dorado 2012), so that mutations with $d < 0.1$ will hardly be purged in the $N = 10$ lines, and this mutation class contributes twice the inbreeding load in the WILD case than in the CAPTIVE one (0.36 vs. 0.18). However, even in this *N*=10 case, IP predictions are much more accurate than those computed ignoring purging.

However, IP predictions do not account for the fitness decline caused by the continuous accumulation of newly arisen mutations, which explains why IP predictions

tend to overestimate long-term fitness. This bias, although can be corrected in theoretical situations (see Full Model approach in García-Dorado 2012), is unknown in standard practice. In our data, this mutational fitness decline is usually small for the periods considered, again with the exception of the WILD case for N=10 lines. However, it should be noted that this decline from new deleterious mutation continuously accumulates in the long term, so that it can be dramatic for periods longer than considered here. This result warns that, as well as inbreeding load seem to be larger in wild than in captive populations, mutational decline in the wild could also be more threatening than inferred from Mutation Accumulation experiments under laboratory non-competitive conditions (García-Dorado *et al.* 1999, Ávila & García-Dorado 2002, García-Dorado 2003, Caballero *et al.* 2002, Halligan and Keightley 2009). However, excluding cases where drift becomes relevant to the evolution of fitness in the period analyzed, our IP estimates are reliable for predictive purposes (Figure 2).

In addition to the IP model, we used three different models estimating the dependence of individual fitness on F and F_a, where this latter parameter (the ancestral inbreeding) is used as an indirect measure of the purging opportunities in the individual´s ancestors. According to Ballou (1997), when F_a is included into the model, the regression coefficient of fitness on F (b_F) represents the rate of inbreeding depression δ. This can be illustrated by considering the particular case F_a=0, where b_F estimates the rate of inbreeding depression for fitness in non-purged individuals. In agreement with this interpretation, b_F gives reasonable NNLR estimates of δ_{FM} when based on short-term data, where F_a is small. The meaning of b_F is less clear for $F_a > 0$ since, as shown in the IP approach, the dependence of fitness on F among purged individuals depends on how fast inbreeding has been produced and, therefore, it also depends on F_a. This explains

why b_F becomes a poor estimator of δ_{FM} when based on longer periods, showing important bias of different sign depending on the model used.

In Ballou's model, purging is measured by the coefficient (b_{FFa}) of the interaction factor ($F \cdot F_a$). Thus, this model considers that the role of purging is reducing inbreeding depression and, therefore, only affects inbred individuals. Thus, b_{FFa} measures the rate of reduction of inbreeding depression with increasing F_a. Due to this interaction term, a common feature of this model and the IP one is that the effect of purging increases as inbreeding accumulates, so that both models predict an initial fitness decline that is latter reversed to some extent, in agreement with the pattern observed in simulated lines.

On the contrary, in the BW model purging is measured by the coefficient b_{Fa}, which represents the rate of increase in fitness with F_a averaged over all F values, and does not account for the reversal of initial depression. However, this BW model accounts for the increase of fitness in outbred individuals that is expected as purging reduces the expressed load (*i.e.*, the non-inbreeding genetic load included in the A term of the seminal Muller's *et al.* model (1956)). Boakes and Wang (2005) found that this model was more efficient detecting purging against mildly deleterious alleles, probably because these mutations were assumed to occur at a larger mutational rate than more severely deleterious ones, causing larger expressed load. Furthermore, they measured the efficiency of the model as the ability to detect cases where purging had reduced the overall load, including inbreeding and expressed genetic loads. However, here we evaluate the ability of the model to detect the reduction in inbreeding depression, so that Ballou's model is more appropriate. Regarding the B-BW model, it did not outperform Ballou's nor B-BW models in Boakes & Wang study (2005), nor in the present analysis. Therefore, in order to compare the performance of IP and F_a-based models to detect and

predict the consequences of purging on inbreeding depression we concentrate in Ballou's F_a-based model.

The estimates for the interaction term in Ballou's model (b_{FFa}) are very dependent on both the size of the lines and the number of generations. Consequently, different pairs of joint estimates (b_F, b_{FFa}) produce different predictions for the evolution of fitness. Although those predictions are always better than the ones computed ignoring selection, their erratic behavior compromises the reliability of Ballou´s method for predictive purposes. It is interesting to note that, as F_a approaches 1, ($b_F \cdot F + b_{FFa} \cdot F \cdot F_a$) approaches ($b_F + b_{FFa}$)·F. Thus, after the early fitness recovery ascribed to purging, this model predicts a continuous rate of decline of fitness with increasing *F*. Since that late decline is not expected as a general consequence of inbreeding and purging, this prediction can be considered a flaw of the model. However, during some medium term period, it can account for the fitness decline ascribed to the fraction of the inbreeding load caused by deleterious alleles that are not being successfully purged (those with $Nd < 1$), or can mimicry the decline from continuous fixation of new deleterious mutations. On the overall, the erratic nature of Ballou's model predictions, ascribed to the inconsistency of the corresponding estimates, makes IP the model to be preferred in order to estimate parameters and to predict the evolution of fitness under inbreeding.

The methods analyzed here are often intended merely with the purpose of detecting purging, with no immediate interest in prediction. The bootstrap contrasts performed to detect purging at each replicate seem to be conservative for both IP and F_a models, and still both models give high rates of purging detection when the size of the lines and the number of generations analyzed are not too small. However, the rate of detection is somewhat larger for IP than for Ballou's analysis. This bootstrap test detects cases where

a model including purging fits data significantly better than a model that predicts the evolution of fitness expected under inbreeding in the absence of selection. However, it is worthwhile to mention that the evolution of each individual replicate is to some extent truly different than expected, due to genetic drift. The difference can be particularly relevant when purging is week ($d \approx 0$) so that genetic drift may become the leading factor governing gene frequencies. Thus, due to drift, some lines can show a true late fitness recovery (no just a large estimate of mean fitness due to within line sampling error) that can be to some extents confounded with a significant effect of purging. In theory, in order to avoid confounding drift and genetic purging it is necessary to average estimates over replicated experiments or to perform meta-analysis including many different estimates.

Our results encourage the use of the IP approach for detection purposes. However, they show that the detection and measurement of purging is very demanding, even in cases where important purging is expected. In the present study, purging can hardly be detected in single pedigreed lines of effective size N=10 recorded for five inbreeding generations, despite high rates of initial inbreeding load and intense purging. Furthermore, the detection rates reported here should not be interpreted as a guide for experimental design with actual data, where detection will depend on the accuracy of the fitness measures, as well as on other concurrent environmental processes. Thus, purging detection is expected to be much more difficult in practice than for these simulation data. These reasons, together with other methodological issues (García-Dorado 2015, López-Cortegano *et al*. 2016), explains why purging detection has been experimentally elusive, particularly in the wild where replicates are not available (Hedrick & García-Dorado 2016). In practice, sampling error in the evaluation of fitness is usually large and genetic drift for fitness is usually small unless population size is very small. Therefore, lack of power is usually a more relevant concern than confounding the consequences of genetic

drift with those of genetic purging in the analysis of actual fitness data. Hopefully, pedigree reconstruction based on massive molecular markers applied to individuals assayed for fitness traits, will allow to obtain large samples of data useful to detect and measuring purging in the future (Fernández & Toro 2006, Wang 2011, Jiménez-Mena *et al*. 2016).

It should be noted that in the IP analysis reported here, the NNLR method searches the *d* estimate only for positive *d* values ($0 \leq d \leq 0.5$). Therefore, if the true *d* value is close to zero, some positive bias is expected for the *d* estimate. When investigating purging in a particular population, including negative *d* values in the NNLR search is not appropriated because g_t need not to converge with increasing *t* for negative *d* values. However, the positive bias expected from searching only positive *d* values should not be a problem, as large bias is expected to be associated to estimates that are not significantly larger than 0 in the bootstrap test. Therefore, here we have presented analysis where NNLR searches only the range of *d* values that makes sense into the IP model ($0 \leq d \leq 0.5$). However, when analyzing a pool of replicates, it could be useful to run additional analysis including negative *d* values in the search in order to check for possible bias in the average of the *d* estimates. In the cases analyzed here, the results reported are very similar to those obtained searching in the interval $-0.5 \leq d \leq +0.5$, except for some underestimation of purging in the case *N*=10 *t=N/2* data (see Supplementary Material). However, the difference can be larger when purging is weaker. An alternative solution to reduce bias in a meta-analyisis is to run NNLR searching only positive *d* values but exclude *d* estimates that do not fit the data significantly better than *d*=0 in the bootstrap analysis. The argument discussed above also applies to Ballou's model regarding the estimate of b_{FFa}. Here we have explored an interval around b_{FFa} =0 for continuity with the original Ballou's approach ($-10 < b_{FFa} < 10$). Searching in the interval ($0 < b_{FFa} < 10$)

gives results quite similar to those reported here, but the estimate of b_{FFa} was larger in the smaller lines (see Supplementary Material), and can lead to important upwards bias in the predictions of fitness (results not shown).

We have noted above that detecting purging may require using data from longer periods of inbreeding, but that this can lead to downwardly biased IP estimates due to overfitting of long-term fitness. Although the bias in joint (δ, d) estimates compensate each other so that they have minor consequences for predictive purposes, it can be relevant when each parameter is considered separately and, in some cases, can induce considerable underestimation of the inbreeding load in the base population. As suggested above, in IP analysis of long-lasting inbreeding processes, it can be convenient to estimate first the rate of inbreeding depression using short term data ($t \leq N/2$) and use this estimate as a known parameter, in order to increase detection rate and reduce bias in the estimate of purging when analysing full-term data. Alternatively, δ could be estimated as the rate of inbreeding depression using data from individuals with no ancestral inbreeding (F_a=0). Furthermore, in some occasions, B could be estimated in the base population from the fitness decline after a single generation of inbreeding. Then, IP estimates of d can be obtained by analyzing full-term data using these estimates of δ (or, in the case of small lines, also using the estimate of B) as known parameters. In any case, Figure 2 clearly illustrates that, even when (δ, d) were jointly estimated from data including long term fitness, the IP approach is a useful predictive model for the range of population sizes and generations considered here and for both CAPTIVE and WILD mutational cases.

Acknowledgements

We are grateful to Carlos López-Fanjul for helpful comments. This work was funded by grant CGL2014-53274-P and FPI research fellowship (BES-2012-055006) from MINECO (Spanish Government).

Literature Cited

Ávila, V., & García-Dorado, A. (2002). The effects of spontaneous mutation on competitive fitness in Drosophila melanogaster. Journal of Evolutionary Biology, 15: 561-566.

Ballou, J.D. (1997). Ancestral inbreeding only minimally affects inbreeding depression in mammalian populations. Journal of Heredity, 88: 169-178.

Bersabé, D. and García-Dorado, A. (2013). On the genetic parameter determining the efficiency of purging: an estimate for Drosophila egg-to-pupae viability. J. Evol. Biol. 26: 375-385.

Bersabé, D., Caballero, A., Pérez-Figueroa, A. and García-Dorado, A. (2015). On the consequences of purging and linkage on fitness and genetic diversity. *G3: Genes*| *Genomes*| *Genetics* , 6: 171-181.

Boakes, E. H. and Wang, J. (2005). A simulation study on detecting purging of inbreeding depression in captive populations. Genetic Research, 86: 139-148.

Boakes, E.H., Wang, J. and Amos, W. (2007). An investigation of inbreeding depression and purging in captive pedigreed populations. Heredity, 98: 172-182.

Bryant, E. H., Backus, V. L., Clark, M. E., and Reed, D. H. (1999). Experimental tests of captive breeding for endangered species. Conservation Biology, 13: 1487-1496.

Caballero, A., Cusi, E., Garcia, C., & Garcia-Dorado, A. (2002). Accumulation of deleterious mutations: additional *Drosophila melanogaster* estimates and a simulation of the effects of selection. Evolution, 56: 1150-1159.

Ceballos, F. C., & Álvarez, G. (2013). Royal dynasties as human inbreeding laboratories: the Habsburgs. Heredity, 111: 114-121.

Conde, D.A., Resness, N., Colchero, F., Jones, O.R. & Scheuerlein, A. An emerging role of zoos to conserve biodiversity. Science 331: 1390-1391.

Charlesworth, D. and Willis, J.H. (2009). The genetics of inbreeding depression. Nat. Rev. Genet. 10: 783-796.

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

Falconer, D.S. and Mackay, T. (1996). Introduction to quantitative genetics. Ed. 4. Longman, Essex, England.

Fernández, J. & Toro M.A. (2006) A new method to estimate relatedness from molecular markers. Molecular Ecology, 15: 1657-1667.

Frankham, R., Bradshaw, C.J.A. & Brook, B.W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biological Conservation, 170: 56-63.

Avila, V., & García-Dorado, A. (2002). The effects of spontaneous mutation on competitive fitness in *Drosophila melanogaster*. Journal of Evolutionary Biology, 15: 561-566.

García-Dorado, A. (2003). Tolerant versus sensitive genomes: the impact of deleterious mutation on fitness and conservation. Conservation Genetics, 4: 311-324.

García-Dorado, A. (2007). Shortcut predictions for fitness properties at the mutation–selection–drift balance and for its buildup after size reduction under different management strategies. Genetics, 176: 983-997.

García-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. Genetics, 190: 1561-1576.

García-Dorado, A., López-Fanjul, C., & Caballero, A. (2004). Rates and effects of deleterious mutations and their evolutionary consequences. Evolution: From Molecules to Ecosystems, 20-32.

García-Dorado, A. (2015). On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. Heredity, 115: 185-187.

García-Dorado, A., López-Fanjul, C. and Caballero, A. (1999). Properties of spontaneous mutations affecting quantitative traits. Genetics Research, 74, 341-350.

García-Dorado, A., Wang, J. and Lopez-Cortegano, E. (2016). Predictive model and software for inbreeding-purging analysis of pedigreed populations. Genes Genomes and 6: 3593-3601.

Halligan, D. L., & Keightley, P. D. (2009). Spontaneous mutation accumulation studies in evolutionary genetics. Annual Review of Ecology, Evolution, and Systematics, 40: 151-172.

Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. Trends in Ecology & Evolution, 31: 940-952.

Hedrick, P. W., & Kalinowsky, S. T. (2000). Inbreeding depression in conservation biology. Annual Review of Ecology and Systematics, 31: 139-162.

Jiménez-Mena, B., Schad, K., Hanna, N. & Lacy, R.C. (2016). Pedigree analysis for the genetic management of group-living species. Ecology and Evolution, 6: 3067-3078.

Kennedy, E.S., Grueber, C.E., Duncan, R.P. and Jamieson, I.G. (2014). Severe inbreeding depression and no evidence of purging in an extremely inbred wild species – The Chatham Island black robin. Evolution, 68: 987-995.

Lande, R. (1994). Risk of population extinction from fixation of new deleterious mutations. Evolution, 48: 1460-1469.

López-Cortegano, E., Vilas, A., Caballero, A., and García-Dorado, A. (2016). Estimation of genetic purging under competitive conditions. Evolution 70: 1856-1870.

Morton, N.E., Crow, J.F. and Muller, H.J. (1956). An estimate of the mutational damage in man from data on consanguineous marriages. PNAS USA, 42: 855-63.

O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W. and Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biol Conserv., 133: 42-51.

Pérez-Figueroa, A., Caballero, A., Garc´ıa-Dorado, A. and Lopez-Fanjul, C. (2009). The action of purifying selection, mutation and drift on fitness epistatic systems. Genetics, 183: 299-313.

Ralls, K., Ballou, J. D., & Templeton, A. (1988). Estimates of lethal equivalents and the cost of inbreeding in mammals. Conservation biology, 2: 185-193.

Toro, M.A., Silió, L., Rodrigañez, J., Rodríguez, C. & Fernández, J. (1998). Optimal use of genetic markers in conservation programmes. Genet. Sel. Evol., 31: 255-261.

Wang, J. (2011). COANCESTRY. A program for simulating, estimating and analyzing relatedness and inbreeding coefficients. Molecular Ecology Resources, 11: 141-145.

Wang, J., El-Kassaby, Y.A., Ritland, K. (2011). Estimating selfing rates from reconstructed pedigrees using multilocus genotype data. Molecular Ecology, 21: 100-116.

Yun, L. & Agrawal, A.F. (2014). Variation in the strength of inbreeding depression across environments: Effects of stress and density dependence. Evolution, 68: 3599-3606.

DISCUSIÓN

ESTIMAS DE LOS PARÁMETROS DE LA PURGA GENÉTICA

El presente trabajo aborda, a través de sus tres capítulos, diversos aspectos de la purga genética utilizando el modelo IP (*Inbreeding-Purging*) (García-Dorado 2012). Este modelo ha demostrado proporcionar buenas predicciones de los efectos conjuntos de la consanguinidad y la purga utilizando datos de simulación. Sin embargo, su uso práctico requiere disponer de estimas del coeficiente de purga *d*, un parámetro relacionado con la magnitud del componente recesivo de los efectos deletéreos ($d = s(1/2 - h)$) o, al menos, de cierto conocimiento de su magnitud esperada. Por este motivo, la estimación de dicho coeficiente juega un papel central en este trabajo.

Hasta la realización de esta tesis, la única estima de *d* disponible correspondía a un experimento con *Drosophila melanogaster,* que proporcionó una primera validación empírica del modelo IP. Sin embargo, este experimento adolecía de algunas limitaciones en lo que refiere a su uso práctico, debido a que se había obtenido en líneas mantenidas en condiciones de competitividad mínima y con censos efectivos pequeños (6 y 12), muy inferiores a los de la mayoría de las poblaciones amenazadas. Esta estima de *d* podría haber subestimado el coeficiente de purga de las poblaciones silvestres por dos razones. En primer lugar, porque censos efectivos del orden de la decena no permitirán detectar toda la purga atribuible a deletéreos con valores de *d* inferiores a 0.1, cuya evolución estará gobernada en gran medida por la deriva. En segundo lugar, porque existen evidencias de que los efectos deletéreos causantes de la depresión consanguínea son en

promedio mayores cuando se expresan en las condiciones competitivas que caracterizan a las poblaciones naturales que cuando se expresan en un medio benigno menos competitivo (Yun & Agrawal 2014), como el que caracteriza la cría en cautividad o las condiciones del experimento de laboratorio de Bersabé y García-Dorado.

Por este motivo, el primer objetivo abordado en esta tesis es la obtención de nuevas estimas de *d*, también para la misma especie pero en líneas de censo más elevado mantenidas en condiciones de alta competitividad, de modo que pudiesen proporcionar una mejor orientación del valores de *d* esperable en poblaciones naturales que sufren una reducción rápida en su censo efectivo.

Así pues, el primer capítulo presenta estimas de *d* obtenidas para dos medidas de eficacia en dos poblaciones de laboratorio. En ambos experimentos las poblaciones se mantenían desde su captura con censo elevado en condiciones muy competitivas. De estas poblaciones se obtenían líneas de censo efectivo reducido (N=52 y N=43, respectivamente) que se mantenían en las mismas condiciones de alta competitividad. Las dos medidas de eficacias eran evaluaciones de la productividad de hijos adultos, y la fundamental diferencia entre ambas es que una se obtenía en condiciones muy competitivas y la otra en condiciones no competitivas. Aunque ambas poblaciones corresponden a experimentos llevados a cabo en diferentes laboratorios (uno en la Universidad de Vigo y otro ejecutado por mí en la Universidad Complutense de Madrid), su diseño es básicamente análogo. Los dos experimentos proporcionan resultados coherentes, a pesar de que las estimas del experimento de Vigo se basan en la observación de la reducción del lastre de consanguinidad, que tanto en el modelo IP como en el de Morton y colaboradores equivale a la tasa de depresión consanguínea δ, mientras que en el experimento de Madrid se basan en la reducción de la eficacia media. A largo plazo, la

caída de la eficacia en las líneas fue imperceptible en el experimento de Vigo y muy pequeña en el de Madrid, y en ambos casos muy inferior al esperado solo por depresión consanguínea.

Por una parte, nuestros resultados indican que el lastre de consanguinidad en la población silvestre que dio origen al experimento de Madrid debió ser aproximadamente B=6 para la eficacia medida en condiciones competitivas, el doble de la estima obtenida cuando la eficacia se evaluaba en condiciones no competitivas. El lastre de consanguinidad de la población silvestre que originó el experimento de Vigo pudo ser algo menor, lo cual no resulta extraño pues las poblaciones de esta especie en el noroeste peninsular son probablemente más pequeñas que las de la zona del Penedés dedicada al cultivo de la vid. Así pues, el valor de B para eficacia competitiva en la población silvestre del experimento de Madrid coincide con la estima media obtenida en el meta-análisis de O'Grady (2006). De todas formas esta coincidencia no debe sobrevalorarse, pues es sabido que existen grandes diferencias entre los valores de B de diferentes poblaciones incluso dentro de un mismo grupo taxonómico, debido en gran medida a diferencias en sus respectivas historias demográficas, habiéndose documentado incluso estimas de $B > 12$ en poblaciones de mamíferos (Jiménez *et al* 1994). De hecho, las estimas incluidas en el meta-análisis de O'Grady refieren a especies de mamíferos y aves, cuyos tamaños poblacionales y estructuras demográficas pueden ser bien distintos de los de un díptero. En todo caso, cabe señalar que las especies de aves están sobrerrepresentadas en dicho meta-análisis. Por tanto, dado que la dispersión (y, por tanto, el censo efectivo) es en general mayor en las especies de aves que en las de mamíferos, el valor B=6 publicado por O'Grady podría sobreestimar el lastre de consanguinidad medio de las poblaciones de mamíferos y, tal vez, ser relativamente parecido al de dípteros.

Por otra parte, las estimas del coeficiente de purga obtenidas en nuestros experimentos son del orden de d=0.3 en lo que refiere al lastre total de consanguinidad, y de d=0.2 en lo que refiere al atribuible a deletéreos no letales. Estas estimas eran muy similares para ambas medidas de eficacia. Es decir, el coeficiente de purga depende fundamentalmente de las condiciones de competitividad en que los deletéreos se exponen a la purga durante el mantenimiento de la línea, más que del grado de competitividad durante la evaluación de la eficacia. Este resultado sugiere que la mayoría de los alelos que son deletéreos en condiciones no competitivas lo son también y con mayor efecto en condiciones más competitivas. Ello explica que, en nuestras condiciones de mantenimiento con alta competitividad, el coeficiente de purga sea muy superior a la estima de Bersabé y García-Dorado (0.09 y 0.02 para el conjunto del lastre de consanguinidad y para el de origen no letal, respectivamente).

Las elevadas estimas del coeficiente de purga obtenidas en nuestros experimentos implican que la purga puede jugar un papel decisivo en la supervivencia de las poblaciones naturales amenazadas, si bien las elevadas estimas del lastre de consanguinidad también implican que la depresión consanguínea puede llegar a ser importante en el corto plazo, antes de que se manifiesten los efectos de la purga. Por tanto, el censo efectivo mínimo requerido para la supervivencia de una población dependerá de si el potencial reproductivo de la misma es suficiente para tolerar dicha depresión transitoria (García-Dorado 2015). Así por ejemplo, utilizando las ecuaciones IP se obtiene que una población con B=6 y d=0.2 y con censo efectivo 50 experimentará una reducción de más del 40% en su eficacia media esperada en menos de 20 generaciones, pero recuperará después paulatinamente un valor próximo a la inicial. Por tanto, en estas condiciones, gracias a la purga un censo efectivo de 50 puede permitir la supervivencia de una población, pero solo si su potencial reproductivo inicial está bien por encima de,

digamos, cuatro hijos por pareja. En el mismo sentido, Caballero *et al* (2017), teniendo en cuenta la acción de la purga, recomendaron un censo efectivo de 70 para evitar un riesgo de extinción atribuible a la depresión a corto plazo en una población simulada con $B = 6$ bajo un modelo mutacional diferente, si bien N=50 podía ser suficiente si el potencial reproductivo inicial correspondía a un máximo de 7 hijos por pareja. Aunque el debate sobre el tamaño mínimo de una población viable seguirá activo, queda claro que la purga juega un papel decisivo en su determinación.

Evidentemente estas consideraciones basadas en nuestras estimas experimentales son solo indicaciones preliminares. Para poder hacer recomendaciones útiles, es necesario obtener estimas conjuntas de B y de *d* en las poblaciones interés o, al menos, en poblaciones de los mismos grupos taxonómicos. Por este motivo los siguientes capítulos de esta tesis abordan el desarrollo y validación de métodos de estima de los coeficientes de purga aplicables a datos de poblaciones naturales o cautivas en que no es posible llevar a cabo experimentos específicamente diseñados para este fin, pero en que se dispone de medidas de eficacia o sus componentes para individuos con registros genealógicos.

Así pues, en el artículo que constituye el segundo capítulo de esta tesis se desarrolla la metodología necesaria para analizar este tipo de información. En primer lugar, se obtienen ecuaciones genealógicas sencillas que permiten el cálculo del coeficiente de consanguinidad purgado (*g*) en función del coeficiente de parentesco purgado y que guardan paralelismo con las ecuaciones clásicas que calculan la consanguinidad de Wright en función de los coeficientes de parentesco de Malécot. A diferencia de las ecuaciones genealógicas para *g* publicadas previamente (García-Dorado 2012), éstas son aplicables a cualquier genealogía, incluyendo la posibilidad de solapamiento de generaciones. En segundo lugar, se estudia el modo de analizar el modelo IP de predicción

de la eficacia teniendo en cuenta su naturaleza exponencial. Por una parte, se contempla el método clásico basado en linealizar el modelo trabajando en escala logarítmica. Se muestra que, en concordancia con los planteamientos de Morton y colaboradores, este método es una buena aproximación cuando los logaritmos se toman sobre la eficacia media como ha venido siendo habitual. No obstante, se demuestra que la pendiente del logaritmo de la eficacia individual sobre el coeficiente de consanguinidad es mayor, en valor absoluto, que la correspondiente al logaritmo de la eficacia media, y se deduce el correspondiente factor de corrección que permite obtener estimas insesgadas de la tasa de depresión consanguínea (δ) a partir de la pendiente del logaritmo de la eficacia individual en la consanguinidad. Este resultado es de interés pues alerta contra el sesgo en que se puede incurrir por interpretar dicha pendiente como una estima directa de δ, y porque permite analizar la fiabilidad de las estimas obtenidas en los análisis de nuestros datos simulados. No obstante, cuando se usen datos reales solo se podrán obtener aproximaciones de ese factor de corrección cuya validez será difícil de evaluar. Además, la transformación logarítmica impide el uso de la información contenida en los individuos en que la eficacia observada es cero. Esta limitación cobra una importancia decisiva en el análisis de datos dicotómicos, como los datos (0, 1) que se general al evaluar la supervivencia hasta un determinado estadio del ciclo biológico. El problema se ha solventado habitualmente mediante regresión logística atendiendo a consideraciones estadísticas, pero este tipo de regresión ajusta un modelo que no coincide con el modelo genético, de naturaleza exponencial. Por este motivo proponemos como método alternativo ajustar directamente el modelo exponencial a los datos de eficacia por procedimientos numéricos.

Así pues, la consecución del segundo objetivo concluye con el desarrollo de una herramienta informática (PURGd) que ofrece dos métodos alternativos para analizar

datos de eficacia de individuos con registros genealógicos utilizando el modelo IP, un ajuste de regresión lineal sobre el logaritmo de la eficacia individual (método LR), y un ajuste de regresión numérica no lineal (método NNLR). En el caso de datos simulados se dispone de información para calcular el factor de corrección que permite inferir δ a partir de la pendiente de regresión del logaritmo de la eficacia en el coeficiente de consanguinidad estimado en el método LR, pero solo el método NNLR ofrece directamente una estima de δ. No obstante, ambos métodos proporcionan estimas bastante similares del coeficiente efectivo de purga. En ambos casos, la exploración de datos obtenidos mediante simulación suponiendo efectos mutacionales constantes en poblaciones mantenidas con dos censos diferentes fue suficiente para ilustrar la alta calidad de los ajustes al modelo IP, y la pobreza de ajuste a un modelo sin purga como el de Morton y colaboradores.

Sin embargo, el modelo IP, aún proporcionando predicciones muy aceptables, es en esencia un modelo aproximado. Es decir, el verdadero valor esperado de la eficacia no es exactamente el valor predicho por el modelo, aun cuando en la predicción se utilicen los verdaderos valores de *d* y δ. Por tanto, cualquier método que estime los parámetros del modelo (*d* y δ) maximizando el ajuste del modelo a los datos puede producir cierto sobreajuste acompañado de cierto sesgo en las estimas. Este problema puede ser más importante en situaciones en que, como ocurre en la naturaleza, el efecto deletéreo de las mutaciones y su grado de dominancia varíe de mutación a mutación. En estos casos, no existe una relación analítica explícita entre el valor de *d* que proporciona predicciones mejores y la distribución del valor de *d* en los loci individuales. Así pues, *d* se define empíricamente como el valor que, al utilizarse en las ecuaciones IP, predice mejor la evolución de la eficacia. Utilizando esta definición, el modelo IP proporciona un ajuste muy razonable. Aun así, el ajuste es algo peor que en un modelo de efectos constantes y,

por tanto, el sobreajuste y los correspondientes sesgos pueden ser mayores.

Para valorar las consecuencias de los sobreajustes que acabamos de mencionar, el último capítulo de esta tesis, correspondiente a un manuscrito aún no publicado, utiliza modelos de efectos mutacionales variables para explorar la validez del modelo IP como método de detección y evaluación de la purga, y valora la utilidad predictiva de las estimas de δ y de *d* obtenidas mediante el análisis IP de datos genealógicos. Además, para evaluar la capacidad de detección de estos análisis, esta versión de PURGd incorpora un método bootstrap para contrastar la significación de la purga en cada réplica simulada. Hemos analizado resultados correspondientes a la eficacia de los individuos con registros genealógicos de poblaciones simuladas mantenidas con distintos censos efectivos (N=10, N=25 y N=50) durante periodos de diferente duración (t=N/2, t=N, t=2N para cada N). Además, todo el proceso se ha repetido utilizando dos juegos diferentes de parámetros mutacionales, analizándose múltiples réplicas en cada caso. En ambos modelos mutacionales, el efecto en homocigosis de las nuevas mutaciones se obtenía de una distribución gamma con la misma forma, y con la misma relación exponencial negativa entre el valor esperado del grado de dominancia (h) y el efecto en homocigosis (s). Sin embargo, en el caso denominado WILD, la tasa de mutación deletérea y el efecto deletéreo medio en homocigosis son el doble que el otro caso, denominado CAPTIVE. Estos parámetros mutacionales están pensados para intentar dar cuenta de la acción de la purga en condiciones silvestres y en cautividad, teniendo en cuenta que en las primeras se observa una tasa de depresión consanguínea que en promedio podría ser la cuarta parte que en las segundas (Frankham *et al*. 2014), debido a que en las poblaciones silvestres segregan más alelos deletéreos y a que su efecto es mayor en las condiciones naturales que en las de cautividad. De hecho, utilizando el modelo mutacional WILD se espera un lastre de consanguinidad B≈6 en el equilibrio mutación-selección-deriva para un censo

efectivo 10^4, similar a la media del meta-análisis de O'Grady y al valor inicial inferido para eficacia competitiva en la población silvestre de *Drosophila* que originó el experimento de Madrid. No obstante, el lastre de esta población después de un periodo largo de mantenimiento en el laboratorio con censo efectivo del orden de 10^3 es B≈2.9, más parecido al de nuestras poblaciones base simuladas en equilibrio mutación-selección-deriva para censo efectivo 10^3 (B≈2.3). En el modelo CATIVE por su parte, el valor de B de nuestras poblaciones base simuladas de censos $N=10^3$ era también del mismo orden que la estima de B para productividad no competitiva de la población grande del experimento de Vigo al final de su mantenimiento en el laboratorio. En definitiva, los modelos mutacionales utilizados para simular las poblaciones analizadas en el tercer capítulo parecen apropiados para ilustrar las propiedades de los distintos métodos de análisis de la purga.

En primer lugar, los resultados muestran que el mejor método de análisis fue el método numérico de regresión no lineal (NNLR) ya que, además de proporcionar estimas directas del lastre de consanguinidad y permitir incorporar valores nulos de la eficacia, produce estimas con errores típicos menores que el método LR, y con las cuales se obtienen predicciones ligeramente más ajustadas a los observados que con el método LR.

Cuando se analizaron datos de las primeras N/2 generaciones, los promedios sobre réplicas de las estimas de δ fueron muy similares a sus valores esperados y las estimas de *d* fueron elevadas. Además, las predicciones IP calculadas usando ambas estimas proporcionaban generalmente buenas predicciones de la evolución de la media de las líneas de los diferente tamaños durante todo el periodo simular, es decir, en cada caso hasta t=2N. Solo en el caso WILD y para líneas mantenidas con censo 10, donde la deriva impide una purga eficiente sobre una buena parte de los alelos deletéreos, se obtiene una

subestima de *d* que afecte de modo perceptible las predicciones de la evolución de la eficacia media de la líneas. Aun así, el ajuste de todas las predicciones es siempre muy bueno, exceptuando, de nuevo, las líneas con censo 10 del caso WILD en que la depresión observada es mayor que las predicciones calculadas con estimas obtenidas en líneas de censo mayor que 10.

Sin embargo, cuando se analizan más de N/2 generaciones, se produce un sobreajuste del modelo que induce subestimas en los dos parámetros. La razón de este sesgo probablemente sea la acumulación progresiva, con las generaciones, de una cola de valores de eficacias altas para individuos con consanguinidad elevada. De todas formas, aunque tanto δ como *d* resultan subestimadas, los dos sesgos se compensan de tal modo que apenas afecta su poder predictivo. Así pues, las estimas obtenidas del análisis de hasta 2N generaciones pueden considerarse fiables en el sentido de que cada pareja (δ, *d*) permite obtener buenas predicciones de la evolución esperada de la eficacia para líneas de diferentes tamaños durante diferentes periodos. No obstante, utilizando el valor teórico de la tasa de depresión consanguínea (δ_{FM}), comprobamos que, si se dispone de una estima externa del lastre de consanguinidad, es posible estimar valores del coeficiente de purga mucho más estables, próximos a los estimados en el periodo inicial. Así pues, para que tanto el lastre inicial de consanguinidad (aproximadamente equivalente a δ en líneas de censo no muy elevado) como *d* se puedan considerar individualmente bien estimados, es preferible analizar solo datos del periodo inicial (t=N/2). Alternativamente, para obtener una buena estima de *d*, podemos introducir en el análisis una estima externa fiable del lastre de consanguinidad (el valor teórico en nuestro caso), o utilizar una estima de *d* previamente obtenida analizando solo las primeras generaciones disponibles. En ambas situaciones utilizando líneas de censo mayor que 10, obtenemos estimas de *d* cercanas a 0.3 en nuestras líneas simuladas bajo el modelo mutacional CAPTIVE, y estimas en torno

a 0.4 en el caso WILD. Para interpretar estos valores, sirva como referencia que en un modelo con $\delta>0$ en que todas las mutaciones tengan los mismos efectos deletéreos y el mismo grado de dominancia, d=0 implicaría que se ha relajado la selección natural y por tanto no hay purga y d = 0.5 corresponde a la situación en que la selección no se ha relajado y toda la depresión se debe a letales recesivos. Nuestras estimas indican por tanto que la purga ha sido muy importante, en concordancia con la observación de depresiones consanguíneas que, a largo plazo, son inapreciables o muy inferiores a las esperadas por simple depresión consanguínea.

Dado que, como se ha discutido más arriba, los lastres de consanguinidad de nuestras poblaciones base CAPTIVE y WILD son del orden de las estimas obtenidas en el capítulo 1 para eficacia no competitiva en una población de laboratorio y para eficacia competitiva en una población silvestre, respectivamente, resulta interesante notar que las estimas de d son también del mismo orden que las correspondientes estimadas obtenidas en *Drosophila*, cuyos intervalos de confianza incluyen valores entre 0.28 y 0.5. Esta observación, si bien no garantiza que nuestros casos CAPTIVE y WILD permitan modelizar fielmente la situación de las poblaciones naturales, sí que avala su adecuación para investigar las propiedades de los métodos de estima y predicción.

Además de analizar la fiabilidad de las estimas de los parámetros IP promediadas sobre réplicas obtenidas mediante simulación, resulta interesante considerar la capacidad del método para detectar la purga ocurrida en las líneas consideradas individualmente, como será el caso habitual al tratar con poblaciones reales. Por una parte, observamos que en el caso WILD hay muchas más réplicas en que esta detección es significativa ($\alpha < 0.05$) que en el caso CAPTIVE, como cabría esperar debido al mayor número de deletéreos segregando, y al mayor valor medio de d. Además, ese porcentaje de detección se

incrementa con el tamaño poblacional, y especialmente con el número de generaciones registrado, por una parte debido al incremento en la información disponible, y por otra a la mayor eficiencia de la purga en poblaciones más grandes y a la manifestación retardada de la sus efectos. De hecho, en el caso CAPTIVE este porcentaje de detección solo superaba el 50% cuando $N \geq 25$ y $t \geq 50$, condición que cumplen pocos conjuntos de datos de poblaciones mantenidas en cautividad. Así pues nuestros resultados sugieren que la detección de la purga puede ser difícil en poblaciones de zoológicos o en programas de conservación *ex situ*. Por el contrario, la detección siempre fue superior al 50% en las líneas simuladas del caso WILD, salvo en las líneas de menor censo (N=10) evaluadas durante solo 5 o 10 generaciones. No obstante, es posible que la detección en poblaciones silvestres se vea entorpecida por la dificultad de evaluación de la eficacia con precisión y por los numerosos factores no genéticos, incluyendo tendencias temporales de origen ambiental, que la afectan.

Por otra parte, las circunstancias óptimas para la detección de la purga, que incluyen la observación de periodos prolongados de consanguinidad en poblaciones no demasiado pequeñas, no coinciden con las circunstancias que proporcionan estimas de *d* consistentes con estimas insesgadas de δ y que implican analizar solo datos del periodo inicial (t=N/2). Una solución para obtener una buena estima de *d* analizando un periodo lo bastante prolongado para tener buenas oportunidades de detección, es introducir en el análisis una estima externa fiable del lastre de consanguinidad (el valor teórico en nuestro caso). Sin embargo es generalmente más factible utilizar solo las primeras generaciones para estimar δ, y usar los datos de un periodo prolongado para detectar la purga y para obtener estimas de *d* condicionadas a dicha estima de δ.

EL MODELO IP FRENTE A MODELOS DE PURGA BASADOS EN CONSANGUINIDAD ANCESTRAL

Seguidamente, hemos comparado la utilidad del modelo IP con la de los métodos que venían utilizándose hasta ahora para detectar y evaluar la purga genética, basados en modelos que utilizan como variables regresoras el coeficiente de consanguinidad (F) y el coeficiente de consanguinidad ancestral (F_a). Para ello, se ha desarrollado una actualización del software PURGd que incorpora la posibilidad de analizar varias alternativas de este método. De este modo, hemos estimado en nuestras líneas simuladas los parámetros de modelos basados en este coeficiente, como son el del propio Ballou (1997), pero también el modelo de Boakes y Wang (BW) (2005) y el modelo mixto resultante de combinar ambos (B-BW). Dado que el modelo de Ballou era el único que ajustaba bien a los datos, nos centramos en comparar la calidad de las estimas obtenidas para los parámetros del modelo IP y del modelo de Ballou. En este modelo de Ballou, F_a aparece en un término de interacción con F, de modo que el coeficiente asociado a este factor representa el efecto de la purga sobre la depresión de la eficacia en individuos consanguíneos con ancestros consanguíneos. Nuestro objetivo era comprobar qué estimas, al ser utilizadas en sus respectivos modelos (IP *vs*. Ballou), predecían mejor la evolución de la eficacia en líneas de distinto censo o durante periodos de diferente duración. Para facilitar este objetivo, desarrollamos una expresión sencilla que predice la evolución de la consanguinidad ancestral media de una población panmíctica de censo efectivo constante.

El resultado principal es que las estimas de los parámetros del modelo de Ballou proporcionan malas predicciones de la evolución de la eficacia, a menos que las predicciones refieran a líneas del mismo censo y al mismo periodo que los datos utilizados

para la estimación. Esto se debe en parte a que el modelo tiene algunas propiedades que no son coherentes con la teoría. Por ejemplo, dado que F_a se aproxima a su asíntota mucho más deprisa que F, en pocas generaciones la reducción predicha de la eficacia dependerá solo de *F*. Además, en el modelo de Ballou, el coeficiente del término en *F* solo estima el lastre de consanguinidad cuando F_a es aproximadamente cero. Por último, las tasas de detección de purga en las líneas individuales también fueron un poco más bajas en los análisis que utilizan el modelo de Ballou que en los que utilizan el modelo IP. Para completar el análisis, también hemos comprobado que NNLR proporciona mejor ajuste a los datos que el análisis logístico previamente utilizado para analizar datos dicotómicos con el modelo de Ballou. Así pues, el modelo IP representa una alternativa más sensible para detectar y medir la purga, y más adecuada para predecir la evolución de la eficacia que los modelos basados en consanguinidad ancestral.

En definitiva, los resultados de los experimentos realizados con *Drosophila* muestran que la purga es eficiente incluso en poblaciones de solo unas pocas decenas de individuos y contra los deletéreos responsables de prácticamente todo el lastre de consanguinidad inicial, incluyendo los deletéreos no severos (parcialmente) recesivos que pudieran segregar en la población base. Además nuestros análisis de poblaciones simuladas muestran que, en líneas con registros genealógicos, se pueden obtener estimas de los parámetros del modelo IP con un buen valor predictivo. Debe notarse que las tasas de detección de purga observadas son modestas en muchos de los casos, a pesar de que los datos son óptimos y el efecto esperado de la purga es importante, lo cual pone de manifiesto la dificultad de detección de la purga en poblaciones concretas en que no existe replicación. Aun así, esta metodología es prometedora en el campo de la conservación, pues los registros genealógicos son habituales en los programas de cría en cautividad y en muchas poblaciones naturales monitorizadas intensivamente debido a su estatus de

amenaza, y afortunadamente es posible inferir o completar genealogías utilizando información molecular (Fernández & Toro 2006, Wang 2011, Jiménez-Mena *et al* 2016). Así pues, la purga debe ser tenida en cuenta en los programas de conservación, y el modelo IP proporciona una metodología adecuada que permite analizar satisfactoriamente datos de eficacia de individuos con genealogía conocida y proporciona estimas de los parámetros del modelo que tienen utilidad predictiva.

CONCLUSIONES

- **La magnitud de la tasa de depresión consanguínea para medidas de eficacia en *Drosophila melanogaster* es mayor en ambientes más competitivos que en condiciones más benignas**, **pero ello se acompaña de un también mayor coeficiente de purga**. Además, cuando la purga actúa en condiciones competitivas es capaz de purgar los deletéreos responsables de la depresión consanguínea para eficacia no competitiva.

- **En poblaciones con registros genealógicos, las estimas empíricas conjuntas de los parámetros del modelo IP (la tasa de depresión δ y el coeficiente de purga *d*) poseen un elevado poder predictivo.** Cuando se utiliza la regresión numérica no lineal (NNLR) propuesta en este trabajo para analizar datos del periodo inicial de consanguinidad, el ajuste a este modelo proporciona buenas estimas del lastre de consanguinidad de la población base y estimas estables de *d*. Las posibilidades de detectar purga significativa se incrementan cuando se analizan datos correspondientes a periodos prolongados de consanguinidad, pero las estimas correspondientes de δ y *d* pueden subestimar los verdaderos valores, si bien ambas subestimas se compensan de tal modo que su uso en el modelo proporciona buenas predicciones de la evolución de la eficacia.

-

- **El modelo IP (García-Dorado 2012) ofrece un marco de trabajo adecuado para predecir las principales consecuencias de la purga**: la reducción del lastre de consanguinidad y la reducción de la depresión consanguínea, tanto en situaciones benignas como competitivas, en poblaciones panmícticas o genealógicas, y bajo diferentes modelos mutacionales.

- **La interacción entre el coeficiente de consanguinidad y la consanguinidad ancestral es un buen indicador de la existencia de purga pero un mal predictor de sus consecuencias cuantitativas**. El modelo de Ballou, cuando se analiza utilizando la metodología NNLR propuesta en este trabajo, presenta tasas de detección de la purga solo ligeramente inferiores a las del modelo IP. No obstante, las estimas de los parámetros del modelo de Ballou tienen malas propiedades predictivas. De nuevo, la estima de la tasa de depresión consanguínea solo es fiable cuando se obtiene del periodo inicial en que la consanguinidad ancestral es reducida.

REFERENCIAS BIBLIOGRÁFICAS

Ager, S. L. 2005. Familiarity breeds: Incest and the Ptolemaic dynasty. *The Journal of Hellenic Studies* 125, 1-34.

Agrawal, A. F. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature 411 (6838)*, 692-695.

Agrawal, A. F. 2010. Ecological determinants of mutation load and inbreeding depression in subdivided populations. *The American Naturalist 176* (2), 111-122.

Agrawal, A. F., and J. R. Chasnov. 2001. Recessive mutations and the maintenance of sex in structured populations. *Genetics 158* (2), 913-917.

Agrawal, A. A., J. K. Conner, and S. Rasmann. 2010. Tradeoffs and negative correlations in evolutionary ecology. *Evolution since Darwin: the first* 150, 243-268.

Álvarez G., F. C. Ceballos and C. Quintero. 2009. The role of inbreeding in the extinction of a European royal dynasty. *PLoS ONE* 4 (4), e5174. doi: 10.1371/journal.pone.0005174.

Amador, C., A. García-Dorado, D. Bersabé, and C. López-Fanjul. 2010. Regeneration of the variance of metric traits by spontaneous mutation in a *Drosophila* population. *Genetics Research 92* (2), 91-102.

Ávila, V., and A. García-Dorado. 2002. The effects of spontaneous mutation on competitive fitness in *Drosophila melanogaster*. *Journal of Evolutionary Biology 15* (4), 561-566.

Ávila, V., C. Amador, and A. García-Dorado. 2010. The purge of genetic load through restricted panmixia in a *Drosophila* experiment. *Journal of Evolutionary biology 23* (9), 1937-1946.

Ávila, V., J. Fernández, H. Quesada, and A. Caballero. 2011. An experimental evaluation with *Drosophila melanogaster* of a novel dynamic system for the management of subdivided populations in conservation programs. *Heredity* 106 (5), 765-774.

Bachtrog, D., M. Agis, M. Imhof, and C. Schlötterer. 2000. Microsatellite variability differs between dinucleotide repeat motifs—evidence from *Drosophila melanogaster*. *Molecular Biology and Evolution* 17(9), 1277-1285.

Ballou, J. 1997. Ancestral inbreeding only minimally affects inbreeding depression in mammalian populations. *Journal of Heredity 88* (3), 169-178.

Berra, T. M., G. Álvarez and F. C. Ceballos. 2010. Was the Darwin/Wedgwood dynasty adversely affected by consanguinity? *BioScience* 60, 376-383.

Bersabé, D., and A. García-Dorado. 2013. On the genetic parameter determining the efficiency of purging: an estimate for *Drosophila* egg-to-pupae viability. *Journal of Evolutionary Biology 26* (2), 375-385.

Bersabé, D., A. Caballero, A. Pérez-Figueroa, and A. García-Dorado. 2016. On the consequences of purging and linkage on fitness and genetic diversity. *G3* 6 (1), 171-181.

Bittles, A. H. and J. V. Neel. 1994. The costs of human inbreeding and their implications for variations at the DNA level. *Nature Genetics* 8, 117-121.

Boakes, E. H. and J. Wang. 2005. A simulation study on detecting purging of inbreeding depression in captive populations. *Genetical Research* 86 (2), 139-158.

Boakes, E., J. Wang, and W. Amos. 2007. An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity* 98 (3), 172-182.

Burnham, K. P. and D. R. Anderson. 2002. *Information and likelihood theory: A basis for model selection and inference* in *Model selection and multimodel inference*, (pp. 49-96). Springer.

Byers, D. L. and D. M. Waller.1999. Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. *Annual Review of Ecology and Systematics* 30, 479-513.

Bryant, E. H., V. L. Backus, M. E. Clark, and D. H. Reed. 1999. Experimental tests of captive breeding for endangered species. *Conservation Biology* 13 (6), 1487-1496.

Caballero, A., E. Cusi, C. García, and A. García-Dorado. 2002. Accumulation of deleterious mutations: additional *Drosophila melanogaster* estimates and a simulation of the effects of selection. *Evolution* 56, 1150-1159.

Caballero, A., I. Bravo and J. Wang. 2017. Inbreeding load and purging: implications for the short-term survival and the conservation management of small populations. *Heredity* 118 (2), 177-185.

Casella, G. and R. L. Berger. 2002. *Statistical Inference*. Second Ed. Pacific Grove, CA: Duxbury.

Ceballos, F. C., and G. Álvarez, 2013. Royal dynasties as human inbreeding laboratories: the Habsburgs. *Heredity* 111 (2), 114-121.

Charlesworth, D., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics* 18, 237-268.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. *Nature Reviews Genetics* 10 (11), 783-796.

Conde, D. A., N. Resness, F. Colchero, O. R. Jones and A. Scheuerlein. 2011. An emerging role of zoos to conserve biodiversity. *Science* 331, 1390-1391.

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

Crnokrak, P., and S. C. Barrett. 2002. Perspective: purging the genetic load: a review of the experimental evidence. *Evolution 56* (12), 2347-2358.

Deák, P., M. M. Omar, R. D. C. Saunders, M. Pál, O. Komonyi, J. Szidonya, P. Maróy, Y. Zhang, M. Ashburner, P. Benos, C. Savakis, I. Siden-Kiamos, C. Louis, V. N. Bolshakov, F. C. Kafatos, E. Madueno, J. Modolell, and D. M. Glover. 1997. P-element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: correlation of physical and cytogenetic maps in chromosomal region 86E-87F. *Genetics* 147 (4), 1697-1722.

Falconer, D. S., and T. Mackay. 1996. *Introduction to quantitative genetics*. Ed. 4th. Longman, Essex, England.

Fernández, J., S. T. Rodríguez-Ramilo, A. Pérez-Figueroa, C. López-Fanjul, and A. Caballero. 2003. Lack of nonadditive genetic effects on early fecundity in *Drosophila melanogaster*. *Evolution* 57 (3), 558-565.

Fernández, J. and M. A. Toro. 2006. A new method to estimate relatedness from molecular markers. *Molecular Ecology* 15, 1657-1667.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation* 126, 131-150.

Frankham, R. 2008. Genetic adaptation to captivity in species conservation programs. *Molecular Ecology* 17 (1), 325-333.

Frankham, R., C. J. Bradshaw, and B. W. Brook. 2014. Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation* 170, 56-63.

Frankham, R., C. M. Gilligan, D. Morris, and D. A. Briscoe. 2001. Inbreeding and extinction: effects of purging. *Conservation Genetics* 2 (3), 279-284.

Franklin, I. R. 1980. Evolutionary change in small populations. Pp. 135-149 M. E. Soule and B. A. Wilcox (eds.) in *Conservation Biology: an Evolutionary-Ecological Perspective. Sinauer Associates, Sunderland, Massachusetts*.

Franklin, I. R., F. W. Allendorf, and I. G. Jamieson. 2014. The 50/500 rule is still valid–Reply to Frankham *et al*. *Biological Conservation* 176, 284-285.

García-Dorado, A. 2003. Tolerant versus sensitive genomes: the impact of deleterious mutation on fitness and conservation. *Conservation Genetics* 4, 311-324.

García-Dorado, A. 2007. Shortcut predictions for fitness properties at the mutation–selection–drift balance and for its buildup after size reduction under different management strategies. *Genetics* 176, 983-997.

García-Dorado, A. 2012. Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics* 190 (4), 1461-1476.

García-Dorado, A. 2015. On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. *Heredity* 115, 185-187.

García-Dorado, A. 2017. An explicit model for the inbreeding load in the evolutionary analysis of selfing. *Evolution*, doi: 10.111/evo.13231.

García-Dorado, A., López-Fanjul, C. and Caballero, A. 1999. Properties of spontaneous mutations affecting quantitative traits. *Genetics Research* 74, 341-350.

García-Dorado, A., C. López-Fanjul and A. Caballero. 2004 Rates and effects of deleterious mutations and their evolutionary consequences, pp. 20–32 in *Evolution: From Molecules to Ecosystems*, edited by A. Moya and E. Font. Oxford University Press, Oxford.

García-Dorado, A., and A. Caballero. 2000. On the average coefficient of dominance of deleterious spontaneous mutations. *Genetics* 155, 1991-2001.

Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* 57 (12), 2678-2687.

Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* 86 (6): 485-486.

Gulisija, D. and J. F. Crow. 2007. Inferring purging from pedigree data. *Evolution* 61 (5), 1043-1051.

Haag, C. R., and D. Roze. 2007. Genetic load in sexual and asexual diploids: segregation, dominance and genetic drift. *Genetics* 176(3), 1663-1678.

Halligan, D. L., and P. D. Keightley. 2009. Spontaneous mutation accumulation studies in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics* 40, 151-172.

Harr, B. and C. Schlötterer. 2000. Long microsatellite alleles in *Drosophila melanogaster* have a downward mutation bias and short persistence times, which cause their genome-wide underrepresentation. *Genetics* 155 (3): 1213-1220.

Hedrick, P. W. 1994. Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity* 73 (4), 363-372.

Hedrick, P. W. and. S. T. Kalinowsky. 2000. Inbreeding Depression in Conservation Biology. *Annual Review of Ecology and Systematics* 31, 139-162.

Hedrick, P. W., U. Hellsten, and D. Grattapaglia. 2016. Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. *New Phytologist 209* (2), 600-611.

Hedrick, P. W., and A. Garcia-Dorado. 2016. Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology & Evolution* 31, 940-952.

IUCN. 2001. *IUCN Red List Categories and Criteria: Version 3.1*. IUCN Species Survival Comission. IUCN, Gland, Switzerland and Cambridge, UK.

Jamieson, I. G., and F. W. Allendorf. 2012. How does the 50/500 rule apply to MVPs? *Trends in Ecology and Evolution* 27 (10), 578-584.

Jiménez, J. A., K. A. Hughes, G. Alaks, L. Graham and R. C. Lacy. 1994. An experimental study of inbreeding depression in a natural habitat. *Science* 266, 271-273.

Jiménez-Mena, B., K. Schad, N. Hanna and R. C. Lacy. 2016. Pedigree analysis for the genetic management of group-living species. *Ecology and Evolution* 6, 3067-3078.

Karaboga, D. and B. Basturk. 2007. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of Global Optimization* 39, 459-471.

Karaboga, D., B. Gorkemli, C. Ozturk, and N. Karaboga. 2014. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. *Artificial Intelligence Review* 42 (1), 21-57.

Kauer, M., and C. Schlötterer. 2004. An analysis of genetic differentiation among assortatively mating *Drosophila melanogaster* in Zimbabwe. *Journal of Evolutionary Biology* 17 (3): 493-500.

Keller, L.F., and D. M. Waller. 2002. Inbreeding effects in wild populations. *Trends in Ecology and Evolution* 17, 230-241.

Kennedy, E. S., C. E. Grueber, R. P. Duncan, and I. G. Jamieson. 2014. Severe inbreeding depression and no evidence of purging in an extremely inbred wild species—the chatham island black robin. *Evolution* 68 (4), 987-995.

Kruuk, L. E., B. C. Sheldon, and J. Merilä. 2002. Severe inbreeding depression in collared flycatchers (*Ficedula albicollis*). *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269 (1500), 1581-1589.

Lacy R. C., and J. D. Ballou. 1998. Effectiveness of selection in reducing the genetic load in populations of *Peromyscus polionotus* during generations of inbreeding. *Evolution* 52, 900-909.

Lande, R. 1988. Genetics and demography in biological conservation. *Science* 241, 1455-1460.

Lande, R. 1994. Risk of population extinction from fixation of new deleterious mutations. *Evolution* 48, 1460-1469.

Latter, B., J. Mulley, D. Reid, and L. Pascoe. 1995. Reduced genetic load revealed by slow inbreeding in *Drosophila melanogaster*. *Genetics* 139 (1), 287-297.

Laws, R. and I. Jamieson. 2011. Is lack of evidence of inbreeding depression in a threatened New Zealand robin indicative of reduced genetic load*? Animal Conservation* 14 (1), 47-55.

Leberg, P. L., and B. D. Firmin. 2008. Role of inbreeding depression and purging in captive breeding and restoration programmes. *Molecular Ecology* 17 (1), 334-343.

Lee, J. K., M. Lascoux and E. V. Nordheim. 1996. Number of lethal equivalents in human populations: how good are the previous estimates?, *Heredity* 77, 209-216.

Liberg, O., H. Andrén, H.-C. Pedersen, H. Sand, D. Sejberg, P. wabakken, M. Kesson, and S. Bensch. 2005. Severe inbreeding depression in a wild wolf *Canis lupus* population. *Biology Letters* 1 (1), 17-20.

Lumley, A. J., L. Michalczyk, J. J. Kitson, L. G. Spurgin, C. A. Morrison, J. L. Godwin, M. E. Dickinson, O. Y. Martin, B. C. Emerson, T. Chapman *et al*. 2015. Sexual selection protects against extinction. *Nature* 522, 470-473.

Lynch, M. and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152 (4), 1753-1766.

Malécot, G. 1948. *Les mathématiques de l'hérédité*. Paris: Masson & Cie., Paris.

Morton, N. E., J. F. Crow, and H. Muller. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. *Proceedings of the National Academy of Sciences of the United States of America 42* (11), 855-863.

Norman, A. J. and G. Spong. 2015. Single nucleotide polymorphism-based dispersal estimates using noninvasive sampling. *Ecology and Evolution* 5 (15), 3056-3065.

O'Grady, J. J., B. W. Brook, D. H. Reed, J. D. Ballou, D. W. Tonkyn, and R. Frankham. 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation 133* (1), 42-51.

Otto, S. P., and A. C. Gerstein. 2008. The evolution of haploidy and diploidy. *Current Biology 18* (24), R1121-R1124.

Pedersen, K. S., T. N. Kristensen, and V. Loeschcke. 2005. Effects of inbreeding and rate of inbreeding in *Drosophila melanogaster*–Hsp70 expression and fitness. *Journal of Evolutionary Biology 18* (4): 756-762.

Pekkala, N., K. Emily Knott, J. S. Kotiaho, and M. Puurtinen. 2012. Inbreeding rate modifies the dynamics of genetic load in small populations. *Ecology and Evolution 2* (8), 1791-1804.

Pemberton, J. M. 2008. Wild pedigrees: the way forward. *Proc. R. Soc. B.* 275, 613-621.

Pemberton, J. M., D. W. Coltman, J. A. Smith and J. G. Pilkington. 1999. Molecular analysis of a promiscuous, fluctuating mating system. *Biological Journal of the Linnean Society* 68, 289-301.

Pérez-Figueroa, A., A. Caballero, A. García-Dorado and C. López-Fanjul. 2009. The action of purifying selection, mutation and drift on fitness epistatic systems. *Genetics* 183, 299-313.

Perrimon, N., L. Engstrom, and A. Mahowald. 1989. Zygotic lethals with specific maternal effect phenotypes in *Drosophila melanogaster*. I. Loci on the X chromosome. *Genetics*,121 (2): 333-352.

Press, W. H. *et al*. 1992. Minimization or maximization of functions in *Numerical recipes in C*. Cambridge University Press, Cambridge, pp. 394-455.

Ralls, K., J. D. Ballou, and A. Templeton. 1988. Estimates of lethal equivalents and the cost of inbreeding in mammals. *Conservation Biology* 2 (2), 185-193.

Reed, D. H., E. H. Lowe, D. A. Briscoe, and R. Frankham. 2003. Inbreeding and extinction: effects of rate of inbreeding. *Conservation Genetics* 4 (3), 405-410.

Ripoll, P. 1977. Behavior of somatic cells homozygous for zygotic lethals in *Drosophila melanogaster*. *Genetics* 86(2): 357-376.

Roze, D., and F. Rousset. 2004. Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. *Genetics* 167 (2), 1001-1015.

Saccheri I, M. Kuussaari, M. Kankare, P. Vikman, W. Fortelius, and I. Hanski. 1998. Inbreeding and extinction in a butterfly metapopulation. *Nature* 392, 491-494.

Sánchez-Molano, E., and A. García-Dorado. 2011. The consequences on fitness of equating family contributions: inferences from a *Drosophila* experiment. *Conservation genetics* 12 (2), 343-353.

Shaffer, M. L. 1981. Minimum population sizes for species conservation. *BioScience* 31 (2), 131-134.

Shearn, A., T. Rice, A. Garen, and W. Gehring. 1971. Imaginal disc abnormalities in lethal mutants of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* 68 (10): 2594-2598.

Siller, S. 2001. Sexual selection and the maintenance of sex. *Nature* 411 (6838), 689-692.

Simmons, M. J., and J. F. Crow. 1977. Mutations affecting fitness in *Drosophila* populations. *Annual Review of Genetics* 11 (1), 49-78.

Soulé, M. E. 1980.Tresholds for survival: maintaining fitness and evolutionary potential. In Soulé, M. E., Wilcox, B. A. (Eds.), *Conservation Biology: An Evolutionary-Ecological perspective*. Sinauer, Sunderland, MA, pp. 151-169.

Swindell, W., and J. Bouzat. 2006. Reduced inbreeding depression due to historical inbreeding in *Drosophila melanogaster*: evidence for purging. *Journal of Evolutionary Biology* 19 (4), 1257-1264.

Templeton, A. R. and B. Read. 1984. The elimination of inbreeding depression in a captive herd of Speke's gazelle. *Zoo Biology* 3, 177-199.

Toro, M.A., L. Silió, J. Rodrigañez, C. Rodríguez. And J. Fernández. 1998. Optimal use of genetic markers in conservation programmes. *Genet. Sel. Evol*. 31 255-261.

Török, T., G. Tick, M. Alvarado, and I. Kiss. 1993. P-lacW insertional mutagenesis on the second chromosome of *Drosophila melanogaster*: isolation of lethals with different overgrowth phenotypes. *Genetics*, 13 (1): 71-80.

Vilas, A., A. Pérez-Figueroa, H. Quesada, and A. Caballero. 2015. Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. *Molecular Ecology* 24 (17): 4419-4432.

Walling, C. A., D. H. Nussey, A. Morris, T. H. Clutton-Brock, L. E. Kruuk, and J. M. Pemberton. 2011. Inbreeding depression in red deer calves. *BMC Evolutionary Biology* 11 (1), 318.

Wang, J. 2011. COANCESTRY. A program for simulating, estimating and analyzing relatedness and inbreeding coefficients. *Molecular Ecology Resources* 11, 141-145.

Wang, J., W. G. Hill, D. Charlesworth, and B. Charlesworth. 1999. Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genetical Research* 74 (2), 165-178.

Wang, J., Y. A. El-Kassaby and K. Ritland. 2011. Estimating selfing rates from reconstructed pedigrees using multilocus genotype data. *Molecular Ecology* 21, 100-116.

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

Weisberg, S. 2005. *Multiple regression* in *Applied linear regression* (pp. 47-68). Wiley Interscience.

Woodworth, L. M., M. E. Montgomery, D. A. Briscoe, and R. Frankham. 2002. Rapid genetic deterioration in captive populations: causes and conservation implications. *Conservation Genetics* 3 (3), 277-288.

Yun, L., and A. F. Agrawal. 2014. Variation in the strength of inbreeding depression across environments: Effects of stress and density dependence. *Evolution* 68 (12): 3599-3606.

APÉNDICE PRIMERO

SUPPORTING INFORMATION

ESTIMATION OF GENETIC PURGING UNDER COMPETITIVE CONDITIONS

Eugenio López-Cortegano[1], Ana Vilas[2], Armando Caballero[2] and Aurora García-Dorado[1]

[1] Departamento de Genética. Facultad de Biología. Universidad Complutense. Madrid 28040. Spain.

[2] Departamento de Bioquímica, Genética e Inmunología. Facultad de Biología. Universidad de Vigo. Vigo 36310 (Pontevedra). Spain.

THE MODEL

According to the Inbreeding-Purging (IP) model (García-Dorado 2012), the mean for fitness ω (or its component traits, here *P* or *W*, as described below) expected in a population after *t* generations since its reduction in size is

$$\omega_t = \omega_0 \exp[-\delta g_t] , \quad \text{(S1)}$$

In this expression, δ is the inbreeding depression rate expected in the absence of selection that, according to classical theory, equals the inbreeding load in the base population, *i.e.*

in the large population at the time when the size is reduced. (Table 1 in the main text gives a glossary of the main terms and subscripts used in the manuscript.) Thus, ω_0 represents mean fitness at the base population or at a synchronous control non-inbred population. Finally, g_t is the purged inbreeding coefficient, which is an analogous to Wright's inbreeding coefficient F_t but is corrected for the reduction of the deleterious allele frequencies expected from purging. It can be approximated as

$$g_t \approx \{\ 1/(2N_e)\ +\ [1 - 1/(2N_e)\ ]\ g_{t\text{-}1}\ \}\ (1 - 2d\ F_{t\text{-}1})\ , \qquad \text{(S2)}$$

where N_e is the reduced effective population size, $F_t = 1 - [1 - 1/(2N_e)]^t$ and d is the purging coefficient, which represents the recessive component of the deleterious effect. In other words, it is the per-copy deleterious effect on relative fitness that is only expressed in homozygosis but is concealed in the heterozygotes. Thus, if the disadvantage of the homozygote is s and that of the heterozygote is hs, we get $d = (s - 2hs)/2$. For example, for a completely recessive lethal allele $d = 0.5$, so that in the lethal homozygote, the two copies account for the corresponding lethal effect ($s = 1$). For neutral and/or additive alleles, $d = 0$, and g_t reduces to F_t. Although the purging coefficient d is defined in the context of a single locus model, it has been shown that an approximated g_t , computed using a single effective d value applying to the pool of deleterious alleles responsible for inbreeding depression through the genome, gives good predictions for the evolution of mean fitness (García-Dorado 2012). Here we use d to denote the effective purging coefficient that accounts for the realised joint consequences of inbreeding and purging. However, improved accuracy can be obtained by taking separately into account the purging upon the inbreeding load ascribed to lethal and non-lethal deleterious alleles (δ_L and δ_{NL}, respectively), *i. e.*,

$$\omega_t = \omega_0 \exp[-\delta_{NL}\ g_{NLt} - \delta_L\ g_{Lt}]\ , \qquad \text{(S3)}$$

where both the lethal and non-lethal purged inbreeding coefficients (g_{Lt} and g_{NLt}, respectively) are computed from Eq. S2 using the corresponding lethal and effective non-lethal purging coefficients (*i.e.*, replacing d in S2 with $d_L = 0.5$ or with d_{NL}, respectively).

In parallel, the inbreeding load expected after t generations since population shrinkage can be approximated as

$$\delta_t = \delta\, g_t\, (1 - F_t) / F_t\,, \tag{S4}$$

which, in the absence of selection, reduces to $\delta_t = \delta\,(1 - F_t)$. Again, accuracy can be gained by separately considering the evolution of non-lethal and lethal inbreeding loads:

$$\delta_t = [\,\delta_{NL}\, g_{NL\,t} + \delta_L\, g_{L\,t}\,]\,(1 - F_t) / F_t\,, \tag{S5}$$

Alternatively, predictions can be computed using the more comprehensive Full-Model (**FM**) approach, which takes into account new mutation and standard non-purging selections. As explained in the main text, it is convenient to use **FM** predictions to take into account the consequences on δ_t of the continuous mutational input of deleterious alleles and the also continuous erosion of the inbreeding load by non-purging selection in our large Vigo population (García-Dorado, 2012). Under this **FM** approach, δ_t can be computed using the expression

$$\delta_t \approx \delta^* + (\delta - \delta^*)\, g_t\, (1 - F_t) / F_t\,, \tag{S6}$$

where δ^* is the inbreeding load expected in the new Mutation-Selection-Drift (MSD) balance for the new reduced population size N_e. Or, considering separately the lethal and non-lethal inbreeding load, as

$$\delta_t \approx \delta_L^* + (\delta_L - \delta_L^*)\, g_{Lt}\,(1 - F_t) / F_t + \delta_{NL}^* + (\delta_{NL} - \delta_{NL}^*)\, g_{NLt}\,(1 - F_t) / F_t\,, \quad \text{(S7)}$$

Thus, these predictions rely on estimates of δ^* that are usually not available.

ORIGIN AND MAINTENANCE OF POPULATIONS AND LINES

Vigo experiment

A laboratory population was founded from about 1000 females captured in a wine cellar close to Vigo (Northwest Spain) in November 2006 (Ávila *et al*. 2011), and maintained at 25ºC under continuous lighting in 7 cm ∅ bottles filled with about 2 cm of standard agar-yeast-sugar medium. Each generation (every two weeks), 30 bottles were established, each containing about 50 males and 50 females sampled from two bottles from the previous generation following a circular scheme.

At generation 86, 1000 males and 1000 females were sampled from the large population to establish 20 lines, each maintained thereafter in a single bottle with exactly 50 male and 50 female parents during 42 generations synchronously to the large population.

Madrid experiment

A population was founded from 276 females captured in Segura Viudas cellar in 2009 (Sant Sadurni d'Anoia, Penedés, Spain), and maintained in the same conditions as in Vigo experiment in 32 (5 cm ∅) bottles filled with about 2 cm of standard agar-yeast-sugar medium, with 40 males and 40 females per bottle. At generation 83, 40 males and 40 females were sampled from each bottle of the large population to found a line. The next generation, each line was split into two similar lines. Therefore, 64 lines were founded, and each was thereafter maintained in the same way in a single bottle,

synchronously to the large population, by transferring 40 males and 40 females to a new bottle during 40 generations.

FITNESS TRAITS ASSAYED

Non-competitive pupae productivity P

In both experiments, non-competitive pupae productivity (*P*) was measured for single 4-days old females placed in individual vials after being mated for 4 days to single males. It was assayed as the number of pupae produced in the vial 11 days after mating. This trait includes egg to pupae viability and female fecundity fitness components, assayed under relatively low density (one mating pair per vial) and in the absence of competition regarding fecundity.

Competitive productivity W

Competitive productivity (*W*) was measured by reference to that of a curly (*Cy/If*) strain (see Bersabé and García-Dorado 2013 for details) that was maintained in the lab in bottles, following a circular scheme similar to that used to maintain the large population. Each competitive productivity value corresponds to a set of 20 mated females from a population (line), competing with 20 mated curly females, all mixed in a single evaluation bottle. In order to minimize environmental effects ascribed to differences in culture density, as well as the corresponding maternal effects, tested females were obtained after two generations of maintenance in single mating vials. The mating protocol to obtain the 20 females to be tested for each evaluation bottle depended on the purpose of the estimate, as explained below. All the 32 bottles of the curly strain contributed similar numbers of females to evaluation. Competitive productivity was estimated in the progeny of each evaluation bottle as the ratio of the number of offspring contributed by the assayed

population or line (wild progeny) to the number of offspring contributed by the marker strain plus 1 (number of curly offspring + 1). The addition of 1 is intended to reduce the estimation bias caused by random sampling when it induces too small numbers in the denominator (Haldane, 1956). Bottles with no curly progeny were excluded. *W* includes egg to adult viability and female fecundity fitness components, both assayed in crowded competitive conditions.

ESTIMATES OF THE EFFECTIVE POPULATION SIZE OF THE LINES FROM MICROSATELLITE DATA

In Vigo experiment, samples of 10, 10 and 50 males from each of the 20 lines were taken at generations 5, 10 and 25, respectively, and characterized for 9 microsatellites: AC002446, AC004641 (Harr and Schlötterer, 2000), Dm1639-TC (Bachtrog *et al*. 2000), 3L9222187ca, 3R1302339ga, 3R16177365gt, 3R22473342gt, 3R24298455ca and 3R11178343ga (located at the right arm of chromosome III (Kauer & Schlötterer, 2004). The protocol is described by Ávila *et al.* (2011) and Vilas *et al*. (2015).

Accordingly, in Madrid experiment, 10 random males per line (from a random sample of 24 lines out of the 64 available) were also chosen at generations 10 and 20 and analyzed for the same microsatellites.

At each generation, F_{ST} (θ from Weir and Cockerham 1984) and their bootstrap intervals were estimated with the software FSTAT (Goudet 1995). Effective population sizes for the lines were then inferred as $N_e = (1/2)[1 - \exp[\ln(1 - F_{ST}) / t]]^{-1}$.

EVALUATION OF THE OVERALL INBREEDING LOAD (δ_t) FOR *P* IN THE LARGE POPULATION IN VIGO EXPERIMENT

The inbreeding load for non-competitive pupae productivity (*P*) was estimated in the large population of this experiment at generations 22, 50, 103 and 111. In each of these generations, a number of males and virgin females were sampled and mated at random in single mating vials. The progeny of these vials was randomly mated (outbred group) or full-sib mated (inbred group) during two or three generations. Non-competitive pupae productivity was assayed under inbreeding (P_I) for females obtained after one generation of full-sib mating, mated to their brothers (generations 22, 50 and 111), or after two generations of full-sib mating, also mated to their brothers (generation 103). Thus, in generations 22, 50 and 111, it was assayed from the number of pupae with inbreeding coefficient 0.375 produced by females with inbreeding coefficient 0.25, while in generation 103 it was assayed from the number of pupae with inbreeding coefficient 0.5 produced by females with inbreeding coefficient 0.375. In all cases, in order to estimate δ, productivity was considered to be equally controlled by the genotype of the mother and the offspring. Therefore, the average of the inbreeding of the mother and the offspring was used as the inbreeding coefficient *F* of the inbred group. In each evaluation, productivity was synchronously assayed as the number of outbred pupae produced by outbred females in the outbred group (P_O). The number of outbred and inbred females evaluated were, respectively: 79 outbred females and 97 inbred females in generation 22; 150 and 149 in generation 50; 44 and 47 in generation 103; and 227 and 251 in generation 111. In each case, the inbreeding load was estimated as $\delta_t = \ln(P_O / P_I) / F$.

EVALUATION OF THE OVERALL INBREEDING LOAD FOR *P* IN THE LARGE POPULATION IN MADRID EXPERIMENT

In Madrid experiment, the inbreeding depression rate for *P* in the large laboratory population was assayed at generation 112. Four males and four virgin females were sampled from each bottle of the large population at that generation, and were randomly mated in individual vials. Three male and three virgin females were obtained from the progeny of each vial *i*. One male was mated to a virgin female from vial *i* + 1, to obtain a set of random mating vials. In parallel, two males from each vial *i* were individually mated to females born in the same vial, to produce two full-sib mating vials (*iA* and *iB*; inbred set). The next generation, a male from each vial *i* from the random mating set was mated to a virgin female from vial *i* + 2, while a male of each *iA* vial from the inbred set was mated to a female from the *iB* vial. Thus, two sets of vials were produced; an outbred set, and an inbred set where offspring of full-sib parents were mated to their double cousins. Therefore, in the inbred set, both parents and offspring had inbreeding coefficient F = 0.25. Productivity was assayed from the number of pupae produced in these vials. A total of 108 and 200 females were analyzed in the outbred and inbred assays, respectively to estimate P_O and P_I . The inbreeding load was estimated as $\delta_t = \ln(P_O / P_I) / F$.

EVALUATION OF THE OVERALL INBREEDING LOAD FOR *W* IN THE LARGE POPULATION IN VIGO EXPERIMENT

The inbreeding depression rate for *W* was estimated in the large population at generation 83 (*i.e*., in the base population of the lines) using a design aimed to obtain an estimate of δ ascribed to inbreeding in the mother, the offspring or both (δ_m, δ_o or δ_{mo}, respectively) as well as the pooled estimate (δ) that we used for predictive purposes.

About five males and five virgin females were randomly sampled from each of the 32 bottles from the large population and were randomly assigned to single mating vials. From each vial, two male and two female offspring were sampled. One male was mated to its sister, the other male was mated to a female from a different randomly sampled vial. Therefore, we obtained a set of outbred single mating vials and another set of brother-sister mating vials. For each set, we sampled 8 males and 8 virgin females from each vial. When four-days old, four of these males were mated to four sisters ìn a single vial for three days. Similarly, four males were mated to four females from a different randomly sampled vial from the same set. Thus we had the following four sets of vials (with four males and four females per vial):

-Outbred-outbred set (OO): mating between unrelated non-inbred individuals.

- Outbred-inbred set (OI): mating between non-inbred full sibs.

- Inbred-outbred set (IO): mating between unrelated individuals that were inbred, as they were offspring of full sibs.

- Inbred-Inbred set (II): mating between full sibs that were inbred, as they were themselves offspring of full sibs.

For each set, the four females of five vials were mixed in a bottle with 20 mated females of the marker strain. After 14 days, the number of wild and Cy/If progeny was recorded for each bottle. Therefore, competitive productivity was assayed in 25-26 bottles for each OO, OI, IO and II scheme, assaying a total of 520, 500, 520 and 520 wild females, respectively. From this design, three different estimates can be obtained for the inbreeding load as explained below.

The inbreeding depression rate ascribed to fitness components that are expressed in the mother, including fecundity or maternal components of viability, was estimated as

$$\delta_m = \ln(W_{IO} / W_{OO}) / 0.25 , \quad \text{(S8)}$$

since in the IO scheme competitive productivity was assayed from the relative number of outbred offspring produced by mothers with inbreeding $F_m = 0.25$.

Analogously, the inbreeding depression rate ascribed to fitness components that are expressed in the offspring, including non-maternal viability components, was estimated as

$$\delta_o = \ln(W_{OI} / W_{OO}) / 0.25 , \quad \text{(S9)}$$

since in the OI scheme competitive productivity was assayed from the relative number of offspring with inbreeding $F_o = 0.25$ produced by non-inbred mothers.

Thus, the total inbreeding depression rate expected when both mother and offspring have similar inbreeding, as in our evaluations of non-competitive pupae productivity in the lines, could be estimated as

$$\delta_{m+o} = \delta_m + \delta_o , \quad \text{(S10)}$$

In addition, the overall inbreeding depression rate for competitive productivity can be estimated from the inbreeding depression observed when both the mother and the offspring are inbred (*i. e.*, using the set II). As in the estimate for non-competitive productivity in Vigo experiment, we computed this estimate by assigning the average inbreeding of mothers and offspring (*i.e.*, [0.25 + 0.375] / 2 = 0.3125) to the average competitive productivity of the II set. This gives

$$\delta_{mo} = \ln(W_{II} / W_{OO}) / 0.3125 , \quad \text{(S11)}$$

The results obtained are summarized in Table S1. The two estimates δ_{mo} and δ_{m+o} were non significantly different from each other ($p < 0.37$). In order to combine all this information, we estimated the overall inbreeding depression rate for competitive productivity by averaging both estimates:

$$\delta = (\delta_{mo} + \delta_{m+o}) / 2 , \qquad \text{(S12)}$$

δ_m	δ_o	δ_{m+o}	δ_{mo}	δ
1.772	1.735	3.507	2.261	2.884
(0.870)	(0.854)	(1.220)	(0.671)	(0.696)

Table S1. Inbreeding depression rates estimated for *W* in the base population in Madrid experiment (bootstrap errors are given in parenthesis).

ESTIMATION OF THE LETHAL INBREEDING LOAD

Chromosomes II were sampled and tested for lethality using a classical design with the Cy/If marker strain, in order to estimate the proportion of lethal chromosomes II (Q_{II}). The protocol can be found in Bersabé and García-Dorado (2013). Since chromosome II is expected to account for 46% of the inbreeding load due to nuclear genes, excluding X chromosome (see *D. melanogaster* Release 5 genome annotation), the overall inbreeding load caused by lethal alleles was estimated as

$$\delta_{Lt} = -\ln(1 - Q_{II}) / 0.46 , \qquad \text{(S13)}$$

Vigo experiment

At generation 128, 549 chromosomes II were sampled from the large population.

Synchronously, three randomly selected lines were assayed in a similar way (149, 169 and 166 chromosomes, respectively).

In addition, during this test we registered the ratio of the number of wild (+/+) to Curly (*Cy*/+) offspring in the vials corresponding to non-lethal chromosomes, which measures the mean fitness of non-lethal wild chromosomes II in homozygosis relative to that of *Cy*/+ heterozygous individuals.

Madrid experiment

At generation 57 after the capture of the laboratory population, 447 chromosomes II were sampled from the large population and tested for lethality using the same protocol as in Vigo experiment. The lethal inbreeding load estimated at generation 57 was used as a proxy for that of generation 83, *i.e.*, those corresponding to the base population of the lines ($\delta_{L(P)}$).

LETHAL COMPONENTS OF THE INBREEDING LOADS FOR *P* AND *W*

The lethal component of the inbreeding load for competitive productivity (*W*), which includes the expression of lethality across the whole life cycle (not just egg to pupae viability), is the overall lethal inbreeding load estimated using Eq. S13.

However, since only 75.6% of all recessive alleles that are lethal for egg to adult viability, affect egg to pupae viability (Deák *et al*. 1997; Perrimon *et al*. 1989; Ripoll, 1977; Shearn *et al*. 1971; Török *et al*. 1993), the lethal component for the inbreeding depression load

of P was estimated as

$$\delta_{Lt} = -0.756 \ln(1 - Q_{II}) / 0.46 , \quad \text{(S14)}$$

EVALUATION OF THE DECLINE FOR THE MEAN OF NON-COMPETITIVE PUPAE PRODUCTIVITY (*P*) IN VIGO LINES

This trait was assayed 25 generations after Vigo lines were founded. Fifteen single pairs were sampled per bottle, both from the large population and from the lines. Pairs from the same population or line were randomly mated for three generations, and non-competitive pupae productivity was synchronously assayed at the second and third generations (assays 1 and 2). For the lines, 256 outbred females and 261 inbred females were evaluated in assay 1, and 239 and 230, respectively, in assay 2. For the large population the corresponding numbers were 264 and 274 in assay 1 and 227 and 251 in assay 2.

EVALUATION OF THE DECLINE FOR THE MEAN OF NON-COMPETITIVE PUPAE PRODUCTIVITY (*P*) IN MADRID LINES

At generation 30 for the lines and, synchronously, at generation 113 for the large population, mean productivity was assayed from the progeny of individuals sampled from the base population and lines, and randomly mated in single mating vials. In both cases, individual matings were done as described for the outbred OO set mentioned above. On the average, 3.5 productivity measures were obtained per bottle for a total of 159 females analyzed for the large population, and 171 females for the lines.

EVALUATION OF THE DECLINE FOR THE MEAN OF COMPETITIVE PRODUCTIVITY (*W*) IN MADRID LINES

This trait was assayed in each of the 64 lines at generations 10, 20, 30 and 40. In each case, it was synchronously assayed at the large population. The design was similar to that of the outbred set (OO) used for the estimate of δ. Arrangements were made to obtain an evaluation bottle per maintenance bottle. The total numbers of wild females evaluated at generations 10, 20, 30 and 40 were 560, 700, 680 and 640 for the large population, and 1000, 1200, 1180 and 1220 for the lines, respectively. At each of these four assays, the mean for competitive productivity W is given as the ratio of the average of the lines to the synchronous estimate of the average of its large population.

INFERENCE OF THE INBREEDING PURGING COEFFICIENT

In Vigo experiment, δ was estimated for P at different times in the large population, but average P in the lines was assayed only once. The opposite was done in Madrid experiment for W. Thus, the inference of the purging coefficient (d) for P was obtained from the evolution of the inbreeding load in the large Vigo population, whereas that for W was obtained from the evolution of the average in Madrid lines.

Inference of d for P from the evolution of δ_t in the large Vigo population

We used Inbreeding-Purging predictions (Eq. S4) to obtain the values of the initial inbreeding load (δ) and of the overall purging coefficient (d) that better fit the observed evolution of the inbreeding load according to a Least Square (LS) criterion. Estimates producing the least square deviations from observed to predicted values were numerically searched over a grid for δ and d ($0 \leq d \leq 0.5$) with steps 0.02 and 0.01, respectively. No local minima were detected.

The procedure was repeated by accounting separately for lethal and non-lethal depression (Eq. S5), in order to obtain estimates of the non-lethal effective purging coefficient d_{NL}. In both instances, the LS estimate of δ was 1.85.

We also computed approximate LS estimates of d and d_{NL} using the Full-Model (**FM**) approximations (Eqs. S6 and S7). For this purpose, we extrapolated δ^* from a laboratory population at the MSD balance maintained with a roughly similar protocol (Amador *et al*. 2010), where homozygosis for lethal and non-lethal chromosomes II produced a viability decline equal to 0.112 and 0.090, respectively. Using expressions analogous to our Eq. S13 (where Q_{II} is replaced with viability decline in homozygosis), these figures allowed us to obtain estimates for the non-lethal and lethal components of the inbreeding load at the MSD balance ($\delta_{NL}^* = 0.15$ and $\delta_L^* = 0.18$), which give $\delta^* = 0.33$. In order to avoid biases in the estimate of the initial inbreeding load associated to the extrapolation of δ_{NL}^* and δ_L^*, we used the initial inbreeding load obtained from the above **IP** estimates. These **FM** estimates, and the corresponding predictions, must be taken as rough approximations.

FM predictions (Eq. S6) were also used to infer the lethal component of the inbreeding load for trait P in the large Vigo population at generation 86, which estimates the corresponding lethal inbreeding load in the base population of Vigo lines. This was obtained using the inbreeding load corresponding to lethal alleles into Eq. S6 (*i.e.*, assuming $\delta = 0.9$ and $\delta^* = 0.15$), where g_t for lethal alleles is computed assuming $d = 0.5$. Results are reported in Table 4 of the main text. The same procedure was used to predict the lethal inbreeding load at generation 128, which gives $\delta_{L128} = 0.28$, a value quite close to the corresponding estimate obtained from lethal analysis (0.316).

Inference of d for W from the evolution of the trait's mean in Madrid lines

A Least Square estimate (LS) was obtained for the overall inbreeding purging coefficient d by fitting the mean relative W observed in the four evaluations at generations 10, 20, 30 and 40 to the corresponding **IP** predictions (Eq. S1), where, for relative competitive productivity, W_0 equals 1. The LS estimate was numerically searched, as in the previous case. Similarly, a LS estimate was obtained for the non-lethal inbreeding purging coefficient d_{NL} by fitting observed values of relative W to predictions separately accounting for purging against lethal and non-lethal inbreeding load (Eq. S3).

APPROXIMATE STATISTICAL TESTS AND CONFIDENCE INTERVALS

Means were compared using *t* or *z* tests.

To obtain an approximate idea on the precision and significance of our estimates of the purging coefficients, we considered the asymptotic χ^2 distribution of the likelihood ratio statistic (Casella and Berger, 2002), in order to derive a statistic

$$F = [\mathrm{SSE}(\theta) - \mathrm{SSE}(\hat{\theta})] / [\mathrm{SSE}(\hat{\theta}) / (n - 1)] , \qquad (15)$$

where n is the sample size, $\mathrm{SSE}(\theta)$ is the sum of square deviations from observed values (either δ_t or W_t) to the corresponding predictions computed using the true value of parameter θ (which stands for d or d_{NL}), and $\mathrm{SSE}(\hat{\theta})$ is the sum of square deviations between the observed values and the predictions computed using the maximum likelihood estimate of the parameter $\hat{\theta}$. Assuming normality for these residual errors, LS estimates are also maximum likelihood estimates, and this statistic has Snedecor $F_{1,\,n-1}$ distribution.

F is computed over a range of hypothetical θ values, in order to look for the interval where p-values were large enough that $\hat{\theta}$ could be considered non significantly different from θ.

REFERENCES

Amador, C., A. García-Dorado, D. Bersabé, and C. López-Fanjul. 2010. Regeneration of the variance of metric traits by spontaneous mutation in a Drosophila population. *Genetics Research* 92(02): 91-102.

Ávila, V., J. Fernández, H. Quesada, and A. Caballero. 2011. An experimental evaluation with Drosophila melanogaster of a novel dynamic system for the management of subdivided populations in conservation programs. *Heredity* 106(5): 765-774

Bachtrog, D., M. Agis, M. Imhof, and C. Schlötterer. 2000. Microsatellite variability differs between dinucleotide repeat motifs—evidence from Drosophila melanogaster. *Molecular Biology and Evolution* 17(9): 1277-1285.

Bersabé, D. and A. Garcia-Dorado. 2013. On the genetic parameter determining the efficiency of purging: an estimate for Drosophila egg-to-pupae viability. *Journal of Evolutionary Biology* 26(2): 375-385.

Casella, G. and R. L. Berger. 2002. *Statistical Inference* (Ed. 2nd) Pacific Grove, CA: Duxbury.

Deák, P., M. M. Omar, R. D. C. Saunders, M. Pál, O. Komonyi, J. Szidonya, P. Maróy, Y. Zhang, M. Ashburner, P. Benos, C. Savakis, I. Siden-Kiamos, C. Louis, V. N. Bolshakov, F. C. Kafatos, E. Madueno, J. Modolell, and D. M. Glover. 1997. P-element insertion alleles of essential genes on the third chromosome of Drosophila melanogaster: correlation of physical and cytogenetic maps in chromosomal region 86E-87F. *Genetics* 147(4): 1697-1722.

García-Dorado, A. 2012. Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics* 190(4): 1461-1476.

Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* 86(6): 485-486.

Harr, B. and C. Schlötterer. 2000. Long microsatellite alleles in Drosophila melanogaster have a downward mutation bias and short persistence times, which cause their genome-wide underrepresentation. *Genetics* 155(3): 1213-1220.

Kauer, M., and C. Schlötterer. 2004. An analysis of genetic differentiation among assortatively mating Drosophila melanogaster in Zimbabwe. *Journal of Evolutionary Biology* 17(3): 493-500.

Perrimon, N., L. Engstrom, and A. Mahowald. 1989. Zygotic lethals with specific maternal effect phenotypes in Drosophila melanogaster. I. Loci on the X chromosome. *Genetics*,121(2): 333-352.

Ripoll, P. 1977. Behavior of somatic cells homozygous for zygotic lethals in Drosophila melanogaster. *Genetics* 86(2): 357-376.

Shearn, A., T. Rice, A. Garen, and W. Gehring. 1971. Imaginal disc abnormalities in lethal mutants of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* 68(10): 2594-2598.

Török, T., G. Tick, M. Alvarado, and I. Kiss. 1993. P-lacW insertional mutagenesis on the second chromosome of Drosophila melanogaster: isolation of lethals with different overgrowth phenotypes. *Genetics*, 13(1): 71-80.

Vilas, A., A. Pérez-Figueroa, H. Quesada, and A. Caballero. 2015. Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. *Molecular Ecology* 24(17): 4419-4432.

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

APÉNDICE SEGUNDO

PURGd 1.1.0 User's Guide

1. PURGd 1.1.0

PURGd is a software developed to detect purging and to estimate inbreeding-purging (IP) genetic parameters in pedigreed populations. The models and methods used in this software are described in García-Dorado *et al*. 2016 [5].

The main objective of this program is to estimate the effective purging coefficient (de , hereafter d for simplicity) described by García-Dorado [4], which is an overall genomic measure of the component of the deleterious effects that is only expressed in homozygosis and is therefore responsible for purging under inbreeding. Furthermore, the program also estimates regression coefficients on the purged inbreeding coefficient (g) and on additional regressors, such as environmental factors or maternal inbreeding. This software also includes options to purging parameters for purging models based on ancestral inbreeding, developed by as Ballou [1] and by Boakes & Wang [2].

Two alternative approaches are implemented:

Linear regression method (LR): A range of d values is numerically searched for a Least Square (LS) estimate. In this process, for each d value considered, a linear regression model is fitted for log-transformed fitness. When using this method, the regression

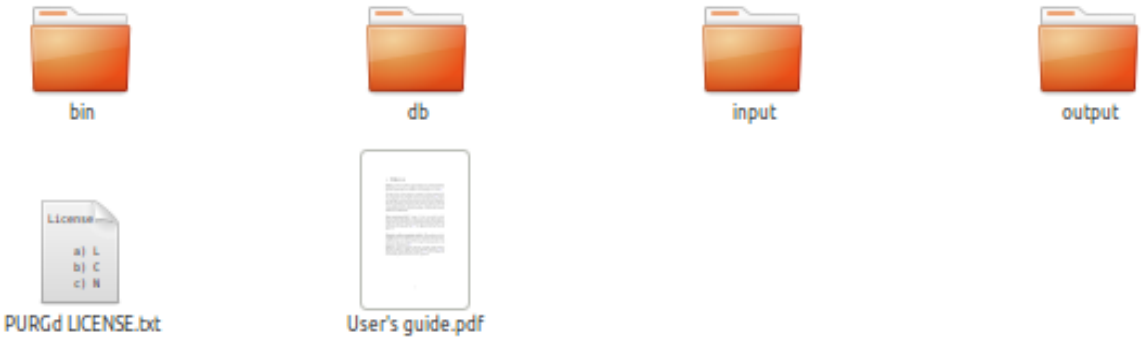
coefficient (b1) on purged inbreeding coefficient (g) overestimates the inbreeding load [5]. This approach cannot use data with fitness ≤ 0.

Numerical non-linear regression method (NNLR): The non-linear model for untransformed fitness is numerically explored searching for the joint numerical LS estimates of d and of the non-linear regression coefficients. In this method, under the IP model, the regression coefficient (b1) on purged inbreeding coefficient (g) estimates the inbreeding load [5].

PURGd also calculates inbreeding coefficients the standard Wright's inbreeding coefficient F, Ballou's ancestral inbreeding coefficient Fa [1] and García-Dorado's purged inbreeding coefficient g [4], as well as the effect of other genetic and environmental factors of interest introduced in the model.

2. PROGRAM FOLDERS

This software is distributed in a package that includes several folders:



bin: executable binaries and setting files.

db: databases, and output files with inbreeding-related coefficients.

input: pedigree files and setfiles.

output: output files with estimated parameters.

Moreover, it contains this user's guide in pdf format, and a copy of the License of the software PURGd as a text file.

3. INSTALLATION

The present software has been written in C++ language, using geany 1.29, and it has been compiled with GNU g++ 6.3.1 under a GNU/Linux (Arch Linux) environment with kernel version 4.9.8-1-ARCH. It is also compiled under Windows 10.

PURGd is a command line software, and it should be used from the terminal.

GNU/Linux: An executable binary file (PURGd) of the program can be found in the bin folder. No installation is needed.

Windows: An executable binary file (PURGd.exe) for Windows can also be found in the bin folder. No installation is needed.

4. INPUT FILES

All input files must be located in the input folder. This program works with two kinds of input files: single pedigree files and setfiles including a set of pedigrees to be analyzed. All input files must be in comma-separated values (.csv) format.

4.1. Pedigree files

Files containing pedigree information must have at least four columns, with the following precise order: individual identity (ID), mother ID, father ID and fitness, as in the examples provided. Fitness values must be numeric, but IDs may be numeric or entered as strings of characters (excluding comma).

Individuals must be ordered in the file, from older to younger. Missing values for fitness or for additional factors must be coded as unknown (NA). In the columns B and C, parents of founder individuals or of individuals from the non-inbred base population can be named using the same ID, like 0 for instance (Table 4.1).

Additional columns can be added containing additional causal factors to be fitted in the model.

Qilin.csv

```
ID,Dam,Sire,Longevity,YOB
1,0,0,1942,273
2,0,0,2106,273
3,0,0,2781,273
4,2,1,2051,275
5,3,1,2593,275
6,0,0,2399,273
7,2,1,4717,276
8,2,1,757,276
9,3,1,919,276
10,4,1,2655,277
11,4,1,2,277
12,5,1,518,277
13,5,1,422,277
```

	A	B	C	D	E
1	ID	Dam	Sire	Longevity	YOB
2	1	0	0	1942	273
3	2	0	0	2106	273
4	3	0	0	2781	273
5	4	2	1	2051	275
6	5	3	1	2593	275
7	6	0	0	2399	273
8	7	2	1	4717	276
9	8	2	1	757	276
10	9	3	1	919	276
11	10	4	1	2655	277
12	11	4	1	2	277
13	12	5	1	518	277
14	13	5	1	422	277

Table 4.1: In the example above, a pedigree file is shown using a text editor, with no blanks, and comma (,) separated values (left); It is also shown how it can be visualized using a spreadsheet program such as LibreOffice Calc or Excel (right).

4.2. Setfiles

A setfile contains a list of names of pedigree files to be analyzed. They are used to run automatically several pedigree files under the same running conditions (see examples below).

In this setfiles, a header in the first row must contain the key word "setfile". The successive rows will contain a list of pedigree file names (without .csv extension).

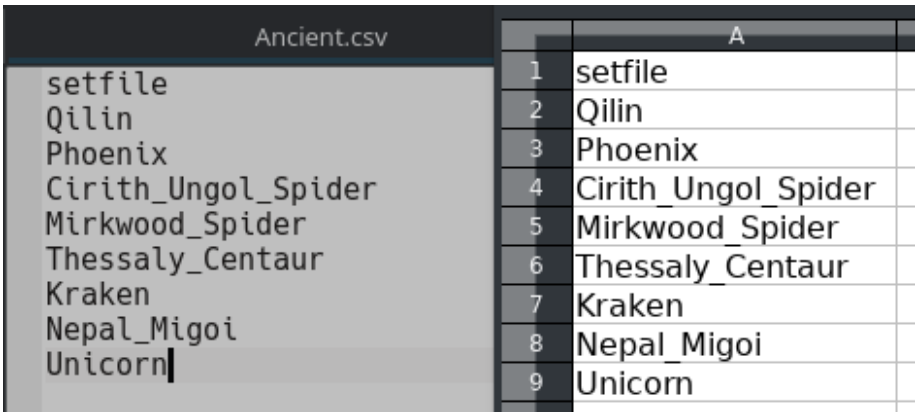


Table 4.2: Setfile (type 1) example: a series of pedigree files Qilin.csv, Phoenix.csv, ... together using the same running parameters, visualized using a text editor (left) or a spreadsheet program (right).

5. START USING PURGd

PURGd is a command terminal program. It can be run in GNU/Linux and Windows command prompts such as bash or cmd.exe, respectively. A guide and examples to run PURGd from Windows is provided in Box 1.

PURGd uses the following syntax:

```
./PURGd  -- method  [ --option ]  datafile
```

where italics indicate the following arguments to be typed by the user:

- **method** refers to the LR (`--lr`) or NNLR (`--nnlr`) methods implemented. An additional method (`--d` *value*) lets to calculate *g* along with other inbreeding coefficients (see section 6) for a given d value.
- **datafile** is the name of the input file or the setfile, preceded by its absolute or relative path.
- **options** need not be specified. Then PURGd assumes IP as the default purging model. Ancestral inbreeding models can be used instead specifying the

corresponding model as an option: Ballou (−−ffa), Boakes & Wang (−−fa) or Ballou and Boakes & Wang (−−faffa) model.

For example,

```
./PURGd −−d 0.27 .. /input/Ancestral.csv
```

runs PURGd assuming *d*=0.27 for the data in the file Ancestral.csv, located in the input folder, in order to calculate *g* for all the individuals.

Alternatively, entering

```
./PURGd −−help
```

will print a short manual to use it in the terminal.

The appearance and procedure, as well as the output files, are the same in both instances. An example of use is shown below:

```
./PURGd −−nnlr −− ffa .. /input/Ancestral.csv
```

In this example, PURGd uses the NNLR method (--*nnlr*) assuming Ballou's model (--*ffa*) on the setfile Ancestral.csv, located in PURGd input folder (note that PURGd is called from the bin folder).

After entering all arguments and options, the software will print a short summary in the terminal as it runs. It also will indicate when the software stops running. Output files and databases will be saved in the output and/or the db folder (section 6).

BOX 1. Using PURGd with Windows

- Download and unfold PURGd in any folder.

- Save your data file, with the correct format, into the input folder. Let´s assume the name of your data file is groundhog.csv

- Click on the Start Menu, then click on the Run option, type **cmd.exe and click OK. That will open a** console **window where you will type lines of commands.**

-To run PURGd you need to move to the bin folder that contains the program using the cd (change directory) command. To do this, open the Windows File Explorer**,** go to the folder where you extracted PURGd. Click on the PURGd icon in the bin folder with the right button and then click in "properties" and copy the location full path. Then go to the console, type cd and paste the path. For example

cd C:\Users\Mary\PURGd_1.1.0\bin

and press intro. **Now the commands you write in the console work in the bin directory**

- Then you can run PURGd from the console. For example, to analyse the data in your groundhog.csv file using the IP model and the NNLR method, you would type

./PURGd --nnlr ../input/ groundhog.csv

where " ../input/" calls the data from the input directory.

If for example you want to use Ballou´s model instead of the IP one you will need to set that option, typing

./PURGd --nnlr--ffa ../input/ groundhog.csv

These analyses use the program settings in the setting.txt file, as explained below.

5.1. Program settings

The bin folder contains a settings.txt text file where the program setting parameters are saved. The program needs it to be read in each run, so be sure it is not cleared or modified in a wrong format.

This file can be modified with a text editor, and allows to define some options that may make the program slower or change its behaviour, but that can perform tasks of interest:

```
# ##################################################################
# Write setting parameters down here, after the equal sign.
# ##################################################################

W0=w

BG=n

MATERNAL=0

USE_ADDITIONAL_FACTORS=0

ADDITIONAL_FACTORS=

NAME_OF_ADDITIONAL_FACTORS=

SCALE=1

STAT=A

ACCURACY=0.01

MAX_BG=10

RANGE_FACTORS=0 10

SAVE_DATABASES=1

RUN_STEPWISE=0
```

W0= : Allows to a introduce a numeric value for the expected fitness for non-inbred individuals, or to calculate it as the average fitness of non-inbred individuals with non-

inbred ancestors (w) (when using the LR method, you should enter a value for the average of ln(w0) or the program wilj compute this average using ln(w) for all non-inbred individuals with non-onbred ancestors).

BG= : Allows to introduce a numeric value for the regression coefficient on *g* (or on F in ancestral inbreeding models), or to estimate it as a parameter (n).

MATERNAL= : Introduce a factor for maternal effects (1) or not (0) by adding the value of *g* (or F in ancestral inbreeding models) in the mother of each individual as a new regressor variable.

USE_ADDITIONAL_FACTORS= : Introduce other additional factors (1) or not (0).

ADDITIONAL_FACTORS= : If used, enter here the column number of additional factors to be used, separated by space.

NAME_OF_ADDITIONAL_FACTORS= : If used, enter here the name of the additional factors, separated by space.

> * Note: Results obtained by including maternal effects or additional factors associated with *g* (as year of birth) should be interpreted with caution, as these factors tend to produce a slight overfitting of the model and can give estimates affected by confounded effects.

SCALE= : Indicates if fitness data are untransformed (1) or logarithmic (2).

STAT= : Indicates the statistic parameter to be used for the Golden Section Search algorithm implemented in the LR method: the adjusted coefficient of determination aR2 ([6]), (R), or the corrected Akaike Information Criterion (A).(AICc) [3].

ACCURACY= : It defines an accuracy value for the search of LS estimate of the purging coefficient and other factors. By default, 0.01 is settled.

MAX_BG= : Defines a maximum value for the search for the BG term in the NNLR method (the minimum value is always assumed to be zero).

RANGE_FACTORS= : Defines the minimum and maximum values to search the effects of each factor in the NNLR method, separated by spaces.

SAVE_DATABASES= : Save databases (1), as described in the output section, or not.

> * Note: only data on analyzed individuals are saved, excluding those with unknown fitness or, in the LR method, those with invalid fitness in logarithmic scale (e.g. fitness 0).

RUN_STEPWISE= : With additional factors, it uses a backward stepwise method to look for the best model in the LR method (1), or not (0).

A description of these setting parameters also appear in the header of the settings.txt file.

6. OUTPUT FILES

A short summary will be displayed on the screen when the program has finished. More complete results will be saved automatically in the output and db folders.

```
================================================================================
The purging coefficient that best fit the data [Cirith_Ungol_Spider] is:
    - According to AICc        : d=0.120312 ( cAIC = 3366.12 )
================================================================================
The purging coefficient that best fit the data [Mirkwood_Spider] is:
    - According to AICc        : d=0 ( cAIC = 2779.06 )
================================================================================
```

Output files have csv format. If opened with a spreadsheet, they can easily be converted for a friendly view (for example, with Excel in Windows select the first column in the file, go to DATA, and choose “text in columns“ - “delimited“ - “comma“). There are three kinds of output files.

* Note: Output files will take their name from input files, and new analysis with the same input files will overwrite previous output files.

6.1. Output files for d and regression coefficients estimates

Analysis performed using the LR or NNLR method will always save a file with the extension _d_coefficients.csv in the output folder. Two output sets are shown in these files, one for the pertinent analysis performed to estimate d, the other one for an analogous analysis conditional to d = 0 and, therefore, assuming no purging. Comparing these two analyses shows how far fitting improves by considering purging.

The output consists of the following columns:

- Pedigree file: indicates the name of a pedigree file.

- d coefficient: the estimated (or assumed) effective purging coefficient.

- p-value (F): the p-value for the Snedecor F test for the significance of the linear regression model being fitted (only for LR).

- aR2: the adjusted coefficient of determination for the linear regression model fitted (only for LR).

- AICc: the corrected Akaike's Information Criterion, assuming normality for residual errors.

- RSS: the residual sum of squares.

- p-value (bootstrap): the p-value for the bootstrap analysis to contrast whether d = 0. Its value only appears in the row for purging analysis (not in the line for analysis assuming d = 0).

- ln(W0) or W0: the initial non-inbred mean for log-fitness or for untransformed fitness.

- b[factor]: the regression coefficient for each factot (*i.e.*, each regressor) included in the analysis (including the purged inbreeding coefficient term g).

- SD[parameter]: standard deviation for parameters estimated (in the numerical case this is only computed for (W 0)

- p-value (t): the p-value obtained from a t test for the significance of each regression coefficient in the linear regression model.

When using setfiles, the output files save the outputs of several pedigree files.

	A	B	C	D	E	F	G	H	I
1	Pedigree file	Analysis	d coefficient	RSS	AICc	p-value (bootstrap)	W0	SD(W0)	b(g)
2	Mirkwood_Spider	Inbreeding-purging model	0.0021	0.580193	-101623	0.5087	0.810258	0.0632063	-0.3526
3		No purging model	0	0.580044	-103857		0.810258	0.0632063	-0.340051
4	Cirith_Ungol_Spider	Inbreeding-purging model	0.11	1.37672	-49.7773	0.5273	0.846213	0.0275694	-0.75
5		No purging model	0	1.37663	-51999		0.846213	0.0275694	-0.74

Table 6.1.1:Output file for d estimation under the NNLR method using a setfile.

For single pedigree files, if stepwise analysis is activated, results for every possible combination of factors will be saved in the output file.

6.2. Databases

If specified in the program settings (section 5.1), databases are saved in the db folder with _data.csv extension, for both LR and NNLR estimation methods.

A separate database is saved for each pedigree file analyzed. It includes several columns (individuals excluded from analysis won't appear in this output file):

- Identity: The identity of the individual.

- Fitness: As it is used by the analysis, so scale may be changed.

- F: Standard inbreeding coefficient

- g(d): Purged inbreeding coefficient computed using the estimate obtained for d.

- Fa: Ancestral inbreeding coefficient.

Additionally, if maternal and other additional factors are included in the model, successive columns will contain their coefficients (only for factors in the best model when stepwise regression is activated):

- gdam(d): Maternal purged inbreeding coefficient computed using the estimate obtained for *d*.

- Effects of additional factors in the input.

6.3. Output files with inbreeding coefficients computed using d values specified by the user

This output file is generated when calculating g for specified values of d (-d argument in PURGd). An output file with _g(d).csv extension is saved in the db folder.

This output file contains inbreeding coefficient values for every individual in the pedigree, and a fixed number of columns:

- ID: The identity of the individual.
- F: Inbreeding coefficient.
- g(d): Purged inbreeding coefficient, for the d value specified by the user.
- Fa: Ancestral inbreeding coefficient.

7. PERFORMANCE

PURGd is an efficient software that runs quickly, and requires low RAM memory usage, so it can be run in a desktop computer or a laptop. Details on the test on performance are shown below.

We measured the actual CPU execution time as the sum of the user and system time, which means that we do not show results for execution time in real (clock) time, as it can be affected by other processes, including input / output (like entering options in the keyboard). CPU time could be defined as the time used within the process. We also measure memory usage through the maximum resident set size (RSS) memory consumed, which is the portion of main memory (RAM) occupied by the process. These values were calculated for simulated pedigrees with different number of individuals per generation (N) and different depth (t, in generations) using the linear regression and the numerical non-linear regression methods available in PURGd.

CPU time increases linearly with the product $N \cdot t$, that is, with the total number of individuals in the input pedigree file (Figure 7) at a rate that is much higher for the numerical method (~ 0,0117 seconds / individual) than for the regression method (~0,0006 seconds /individual), though both methods are in practice very fast: a big pedigree file of about 5000 individuals can be analyzed in a few seconds using the linear regression method, and in a few minutes using the numerical non-linear regression one.

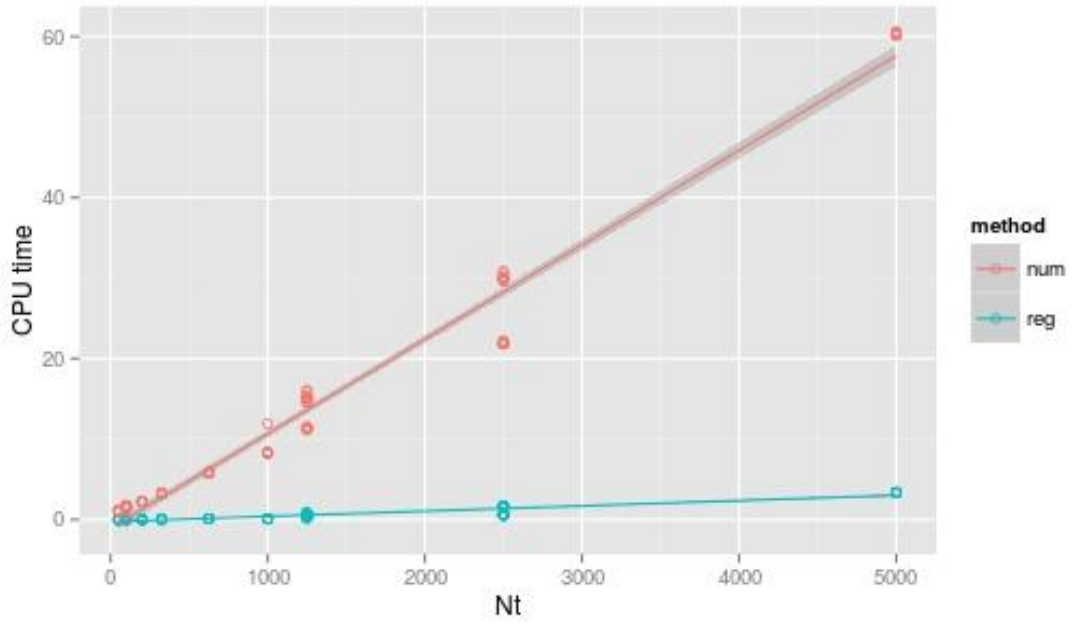


Figure 7: CPU time (in seconds) increase with the number of individuals in the pedigree (N t). The CPU time of the numerical non-linear regression method (num, in red) increases with N t more steeply than that of the linear regression method (reg, in blue).

The maximum RSS used by the bigger pedigree file (N t = 5000) were 18.38MB and 36.01MB for the linear regression and numerical non-linear regression methods, respectively. So it can be concluded that no RAM problems are expected for real pedigree files. Additionally, no memory leak has been detected for this software, so running continuously pedigree files using setfiles will not require any additional memory.

8. ABOUT

Current version of PURGd is 1.1.0 (15/03/2017), compiled with GNU g++ 6.3.1. This Software was developed by:

- Eugenio López-Cortegano
- Jinliang Wang
- Aurora García-Dorado

PURGd is a free software oriented to research, non-commercial use, and it is distributed under the terms described in the PURGd License.txt file. If you use PURGd in your research, cite:

- García-Dorado, A., Wang, J. and López-Cortegano, E. (2016) *Predictive model and software for inbreeding-purging analysis of pedigreed populations*. G3: Genes, Genomes, Genetics 6 (11): 3593-3601.

Users are encouraged to request additional features on the software and to report bugs. In that case, please contact Eugenio López-Cortegano (e.lopez@ucm.es).

This work was funded by grant CGL2015-53274-P and by an FPI research fellowship (BES-2012-055006) from MINECO (Spanish Government).

9. BIBLIOGRAPHIC REFERENCES

[1] BALLOU, J.D. (1997). Ancestral inbreeding only minimally affects inbreeding depression in mammalian populations, Journal of Heredity 88 (3): 169-178.

[2] BOAKES, E. & WANG, J. (2005). A simulation study on detecting purging of inbreeding depression in captive populations, Genet. Research 86: 139-148.

[3] BURNHAM, K.P. & ANDERSON, D.R. (2002). Information and likelihood theory: A basis for model selection and inference in Model selection and multimodel inference, (pp. 49-96). Springer.

[4] GARCÍA-DORADO, A. (2012). Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection, Genetics 190: 1461-1476.

[5] GARCÍA-DORADO A., WANG J., LÓPEZ-CORTEGANO, E (2016). Predictive model and software for inbreeding-purging analysis of pedigreed populations, G3: Genes, Genomes, Genetics 6 (11): 3593-3601.

[6] WEISBERG, S. (2005). Multiple regression in Applied linear, regression (pp.47-68). Wiley Interscience.

APÉNDICE TERCERO

Supplementary Material

ON THE PREDICTIVE VALUE OF THE EFFECTIVE PURGING COEFFICIENT ESTIMATED FROM PEDIGREED POPULATIONS

Eugenio López-Cortegano[1], Diego Bersabé[1], Jinliang Wang[2] and Aurora García-Dorado[1]

[1] Departamento de Genética. Facultad de Biología. Universidad Complutense. Madrid 28040. Spain.

[2] Institute of Zoology. Zoological Society of London. Regent's park, London NW1 4RY. United Kingdom

THE RELIABILITY OF IP PREDICTIONS OBTAINED USING LR ESTIMATES

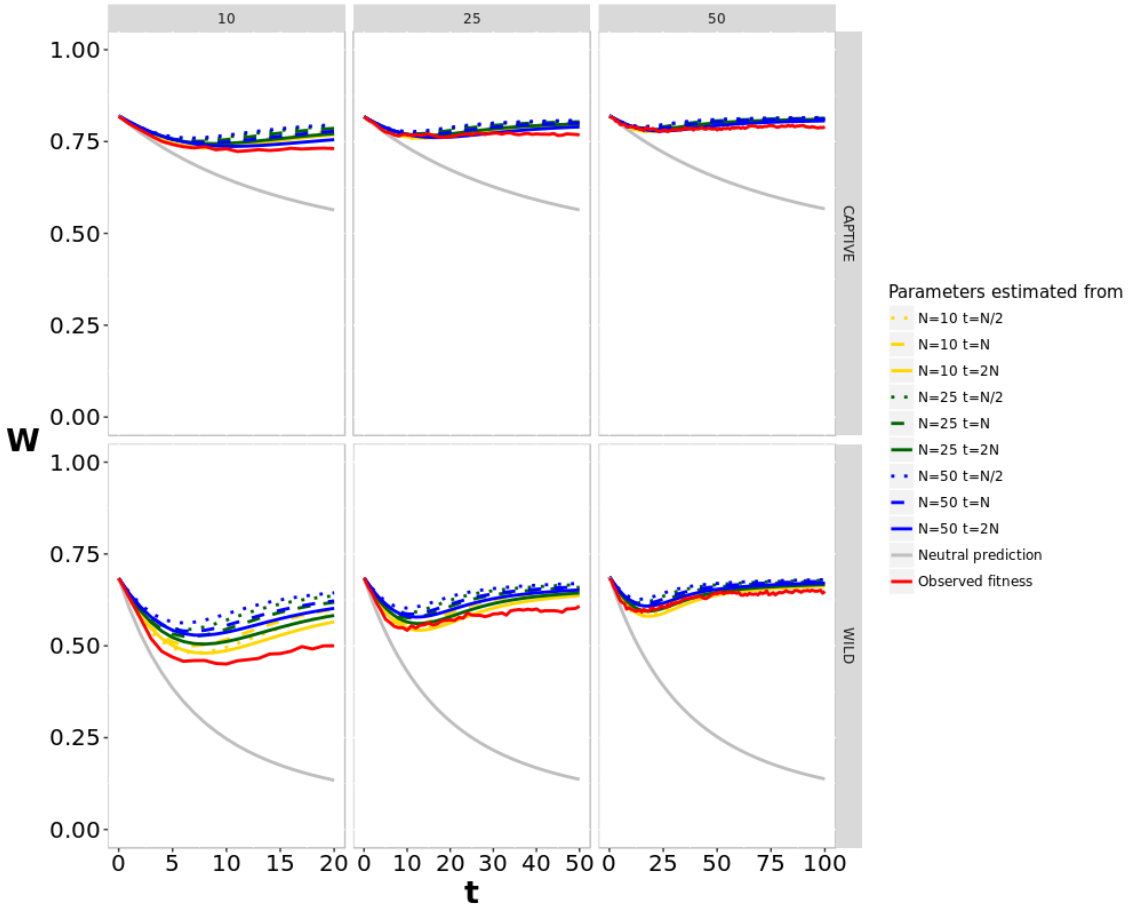


Figure S1: Observed fitness for the CAPTIVE (up) and WILD (down) cases, and the corresponding prediction obtained using LR estimates in the IP model (δ obtained using Equation 2). In each panel, observed and predicted values over t=2N generations correspond to the population size indicated in the column (N=10, N=25 and N=50), and different predictions are plotted using estimates obtained from different data sets, denoted by different colors and strokes as shown in the lateral panel. Neutral predictions, computed assuming no selection and using the inbreeding load observed in the simulated base population (B_{sim}) are also shown.

THE LOGISTIC MODEL

In previous studies of purging estimation, fitness is often evaluated as a categorical binary trait, so inbreeding depression and genetic purging are analyzed using logistic (alias logit) regression.

In order to compare the performance of the NNLR method and the logit approach to analyze binary data, we run the cases described in the main text transforming fitness into binary values. This is achieved by assigning fitness 1 to each individual *i* with a probability equal to W_i , and fitness 0 otherwise. Logit regression through the origin is performed by using $logit(W_0)$ as known intercept in the model.

When using estimates of logit regression coefficient, fitness is predicted as:

$$W_i = \frac{1}{1+e^{-(b_1 \cdot X_{1,i} + \cdots + b_k \cdot X_{k,i})}},$$

where *b* represents the logit regression coefficients (note that the intercept is excluded, as we force regression through the origin), and *X* stands for each independent variable or factor in the model (*F* and $F \cdot F_a$).

Figure S2 shows predictions using Ballou's model based on original fitness data evaluated as a continuous trait (NNLR method) or as a binary variable using both the NNLR and the logit method. Both exponential predictions based on NNLR estimates (one obtained from binary data and the other from original fitness data) are very similar to each other, but logit predictions fit the data worse. Of course, NNLR coefficients estimated from binary data have much higher standard errors (see Tables S1 and S2), as the transformation of original fitness to binary fitness

introduces important sampling error. Standard errors were also larger for logit than for NNLR estimates.

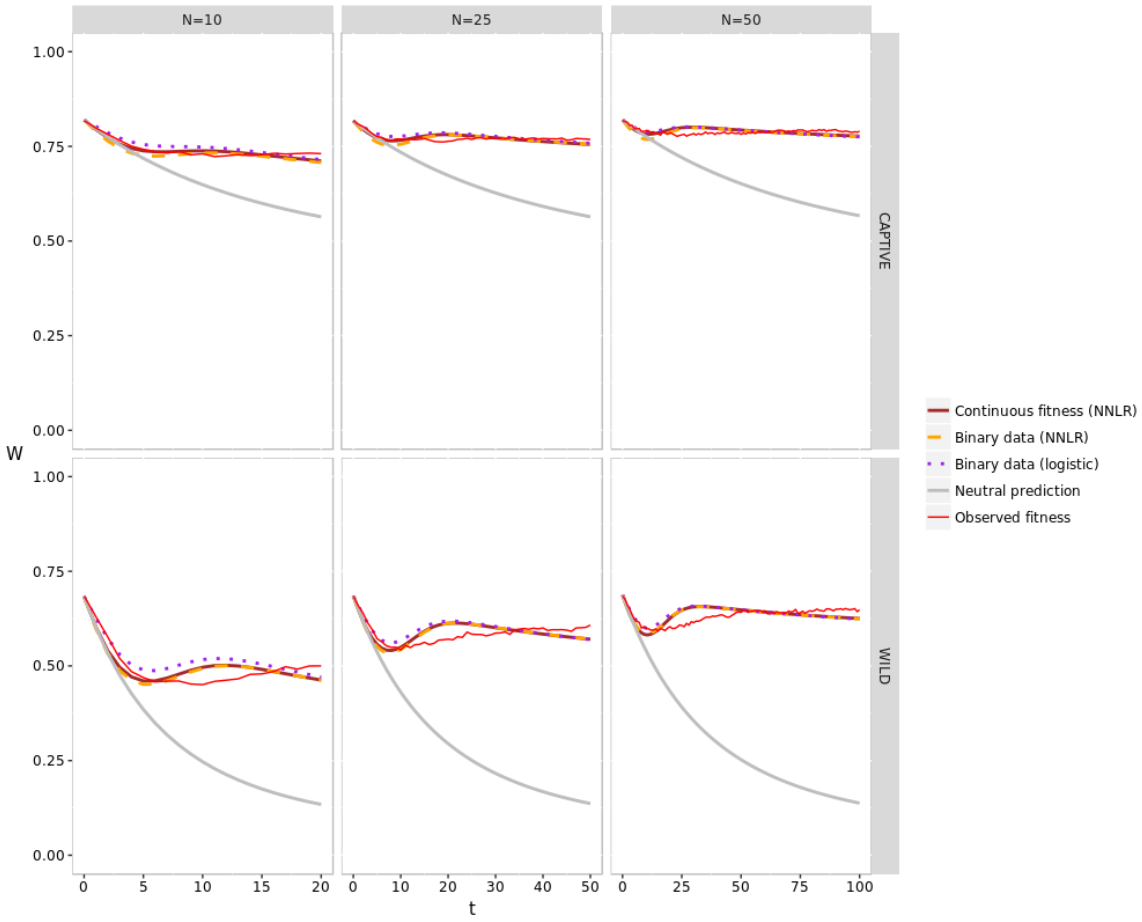


Figure S2: Predictions for fitness in the CAPTIVE and WILD cases computed with Ballou's method using the exponential and the logistic model. Logit Ballou's predictions (dotted lines) are computed using logit estimates computed from binary (0,1) fitness observations (see text for explanation). NNLR predictions are computed using the exponential Ballou's model with NNLR estimates obtained from the same (0,1) fitness data (dashed line). NNLR Ballou's exponential predictions based on NNLR estimated from the original continuous fitness values are also shown for comparison (solid line). Both the estimates of coefficients used to obtain predictions and the observed mean fitness correspond to population sizes (in columns) N=10, N=25 and N=50 during t=2N generations.

	CAPTIVE					
	Continuous fitness (NNLR)		Binary data (NNLR)		Binary data (Logistic)	
	b(F)	*b(F · F_a)*	*b(F)*	*b(F · F_a)*	*b(F)*	*b(F · F_a)*
N=10	-0.6247 ± 0.0214	0.4040 ± 0.0222	-0.7454 ± 0.0357	0.5160 ± 0.0359	-2.2564 ± 0.1166	1.3300 ± 0.1206
N=25	-0.6885 ± 0.0411	0.5644 ± 0.0397	-0.8792 ± 0.0522	0.7566 ± 0.0526	-2.5738 ± 0.1662	2.0162 ± 0.1648
N=50	-0.7228 ± 0.0515	0.6377 ± 0.0523	-1.0169 ± 0.0755	0.9290 ± 0.0744	-2.9368 ± 0.2207	2.4956 ± 0.2240

Table S1: Regression coefficients estimated in the CAPTIVE case for Ballou's model using the NNLR and logistic methods with fitness as a continuous or binary trait in pedigrees from different population sizes (N=10, N=25 and N=50) and *t*=2N generations. Estimates are averaged over replicates, and are given with their empirical standard errors

	WILD					
	Continuous fitness (NNLR)		Binary data (NNLR)		Binary data (Logistic)	
	b(F)	*b(F · F_a)*	*b(F)*	*b(F · F_a)*	*b(F)*	*b(F · F_a)*
N=10	-2.5070 ± 0.0663	1.9002 ± 0.0648	-2.6547 ± 0.1005	2.0480 ± 0.1005	-5.0640 ± 0.1314	3.6847 ± 0.1442
N=25	-2.5705 ± 0.0806	2.2831 ± 0.0799	-2.7120 ± 0.1188	2.4259 ± 0.1099	-5.6883 ± 0.1753	4.9071 ± 0.1748
N=50	-2.6496 ± 0.1065	2.4997 ± 0.1066	-2.5584 ± 0.1558	2.4071 ± 0.1557	-6.0579 ± 0.2762	5.6093 ± 0.2736

Table S2: Regression coefficients estimated in the WILD case for Ballou's model using the NNLR and logistic methods with fitness as a continuous or binary trait in pedigrees from different population sizes (N=10, N=25 and N=50) and *t*=2N generations. Estimates are averaged over replicates, and are given with their empirical standard errors.

IP AND BALLOU ESTIMATES OBTAINED FOR ALTERNATIVE NNLR SEARCH INTERVALS

CAPTIVE		IP		Ballou	
		δ	*d*	b_F	b_{FF_a}
	t=N/2	0.5489	0.1353	-0.6359	0.6139
		±0.0180	± 0.0238	± 0.0210	± 0.0520
N=10	t=N	0.5285	0.2176	-0.6143	0.3649
		±0.0169	± 0.0163	± 0.0185	± 0.0217
	t=2N	0.4868	0.1920	-0.6468	0.4274
		±0.0103	± 0.1624	± 0.0199	± 0.0194
	t=N/2	0.504	0.288	-0.5961	0.4083
		±0.0204	± 0.0230	± 0.0240	± 0.0334
N=25	t=N	0.4783	0.2516	-0.6115	0.4250
		±0.0219	± 0.0179	± 0.0275	± 0.0255
	t=2N	0.4008	0.1874	-0.6813	0.5582
		±0.0193	± 0.0169	± 0.0404	± 0.0391
	t=N/2	0.4893	0.2792	-0.5911	0.3787
		±0.0266	± 0.0263	± 0.0378	± 0.0412
N=50	t=N	0.4284	0.1974	-0.6599	0.5088
		±0.0241	± 0.0221	± 0.0448	± 0.0421
	t=2N	0.3727	0.1474	-0.7322	0.6482
		±0.0203	± 0.0198	± 0.0532	± 0.0537

Table S3: Estimates obtained in the CAPTIVE case. In the IP method NNLR searched for the *d* estimate in the interval $-0.5 \leq d \leq +0.5$. In Ballou's method, NNLR searched for the b_{FFa} estimate in the interval $0 \leq b_{FFa} \leq +10$. Estimates are for different populations sizes (*N=10*, *N=25* and *N=50*) and numbers of generations ($t = N/2$, $t = N$ and $t = 2N$). They are averaged over replicates, and are given with their empirical standard errors.

WILD		IP		Ballou	
		δ	*d*	b_F	b_{FF_a}
	t=N/2	2.2463	0.2340	-2.5570	2.8791
		±0.0537	±0.0222	± 0.0625	± 0.1807
N=10	t=N	2.1076	0.3054	-2.2702	1.4859
		±0.0465	±0.0104	± 0.0539	± 0.0817
	t=2N	1.7903	0.2162	-2.5022	1.9015
		±0.0391	±0.0075	± 0.0653	± 0.0629
	t=N/2	2.0527	0.4076	-2.2320	1.7386
		±0.0566	±0.0112	± 0.0626	± 0.0841
N=25	t=N	1.8372	0.3184	-2.2960	1.8257
		±0.0512	±0.0120	± 0.0708	± 0.0694
	t=2N	1.404	0.2031	-2.5520	2.2616
		±0.0439	±0.0114	± 0.0800	± 0.0793
	t=N/2	1.8741	0.3935	-2.0973	1.6061
		±0.0605	±0.0156	± 0.0741	± 0.0794
N=50	t=N	1.6393	0.3116	-2.3798	2.0946
		±0.0529	±0.0174	± 0.0836	± 0.0818
	t=2N	1.4089	0.2564	-2.6051	2.4546
		±0.0607	±0.0215	± 0.1202	± 0.1219

Table S4: Estimates obtained in the WILD case. In the IP method NNLR searched for the *d* estimate in the interval $-0.5 \le d \le +0.5$. In Ballou's method, NNLR searched for the b_{FFa} estimate in the interval $0 \le b_{FFa} \le +10$. Estimates are for different populations sizes (*N=10*, *N=25* and *N=50*) and numbers of generations ($t = N/2$, $t = N$ and $t = 2N$). They are averaged over replicates, and are given with their empirical standard errors.