Inferring the age and strength of mutations evolving under positive selection

THÈSE N^{O} 7500 (2017)

PRÉSENTÉE LE 8 JUIN 2017
À LA FACULTÉ DES SCIENCES DE LA VIE
UNITÉ DU PROF. JENSEN
PROGRAMME DOCTORAL EN BIOTECHNOLOGIE ET GÉNIE BIOLOGIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Nicky Louise ORMOND

acceptée sur proposition du jury:

Prof. B. Deplancke, président du jury
Prof. J. D. Jensen, directeur de thèse
Prof. L. Excoffier, rapporteur
Prof. T. Flatt, rapporteur
Dr S. Joost, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2017

Acknowledgements

This work was supported by grants from the Swiss National Science Foundation, and a European Research Council (ERC) Starting Grant, to JDJ.

I would like to thank Jeffrey Jensen for his insightful guidance and counsel throughout my thesis, for his commitment to ensuring that laboratory members continue to learn and develop, and for his amazingly quick responses to requests for feedback and help. It has been an exciting journey and a privilege to work here. I am also grateful to members of the Jensen laboratory, both past and present, and in particular to Matthieu Foll, Daniel Wegmann, Claudia Bank, Stefan Laurent, Susanne Pfeifer, Sebastian Matuszewski, Severine Vuilleumier, Valeria Montano and Kristen Irwin for generously giving their time and help whenever asked. I would also like to thank my family for their unwavering support.

Lausanne, 9th November 2016

Abstract

A fundamental goal of population genetics is to determine and quantify the interplay of mutation, natural selection, genetic drift and migration in shaping allelic frequency changes that underpin evolutionary change. In this dissertation, I present computational, empirical and experimental approaches to address this goal. In the first chapter, I develop an approximate Bayesian computation (ABC) approach that co-estimates the selection strength and age of fixed beneficial mutations in single populations, by integrating a range of existing diversity, site frequency spectrum, haplotype and linkage disequilibrium based summary statistics. This approach is then extended to models of selection on standing variation in order to co-infer the frequency at which positive selection began to act upon the mutation. In the second chapter, I apply this method to an empirical study of convergent adaptation of blanched dorsal phenotypes in two lizard species to the newly formed White Sands system of New Mexico. Estimates of the age of the beneficial mutations underpinning the evolution of cryptic coloration are younger than the related geological shift and support a model of adaptation from *de novo* mutation. In the third chapter, I analyze the experimental evolution of H1N1 influenza virus populations under a combined protocol of two drugs with different modes of action: oseltamivir and favipiravir. Results indicate a complex interplay of mutation, selection and genetic drift, where selective sweeps around oseltamivir resistance mutations hitchhike deleterious mutations owing to the mutagenic effect of favipiravir to fixation. This effect reduces viral fitness and accelerates extinction via Muller's ratchet, but at the risk of spreading both established and newly emerging oseltamivir resistance mutations.

Keywords

Positive selection, allele age, Approximate Bayesian Computation, adaptation, convergent adaptation, experimental evolution, influenza H1N1, drug resistance, genetic hitchhiking, Muller's ratchet, Hill Robertson interference

Résumé

Un des objectifs principaux de la génétique des populations est d'identifier et de quantifier les interactions entre mutation, sélection naturelle, dérive génétique et migration, c'est à dire les quatre forces évolutives déterminant les changements de fréquences alléliques à la base du processus évolutif. Dans cette thèse de doctorat, je présente de nouvelles approches statistiques empiriques et experimentales qui permettent de répondre a cet objectif. Dans le premier chapitre, je développe une méthode ABC (approximate Bayesian computation) permettant l'estimation conjointe du coefficient de sélection et de l'âge des mutations avantageuses dans des populations seules. Cette approche se base sur une simplification des données brutes à l'aide de statistiques décrivants le spectre de fréquences alléliques, la diversité haplotypique et le déséquilibre de liaison. Cette méthode ABC est ensuite adaptée pour l'analyse d'un modèle de sélection dit « SGV » (standing genetic variation) dans lequel l'apparition de l'allèle avantageux précède l'apparition de la contrainte sélective, dans le but d'estimer la fréquence initiale de l'allèle avantageux. Dans le deuxième chapitre, j'applique cette méthode à une étude empirique sur l'adaptation convergente de phénotypes de camouflage à un nouvel environnement dans deux espèces de lézards du Nouveau Mexique. Les estimations obtenues avec cette nouvelle méthode suggèrent que les allèles avantageux responsables de l'évolution du camouflage sont survenus après les modifications environnementales correspondantes et suggèrent donc une histoire adaptative caractérisée par l'apparition de nouvelles mutations bénéfiques. Dans le troisième chapitre, j'analyse l'évolution expérimentale de populations du virus de la grippe (influenza) H1N1 soumis à deux traitements différent: l'oseltamivir et le favipiravir. Les résultats indiquent une interaction complexe entre mutation, sélection et dérive génétique, et mettent en évidence un balayage sélectif, du aux mutations de résistance a l'oseltamivir, qui porte et fixe les mutations délétères causées par l'action mutagénique du favipiravir. Cet effet nuit a la capacité réplicative du virus et accélère le processus d'extinction via un processus connu sous le nom de « Cliquet de Muller » (Muller's ratchet), mais avec un risque accru de transmission de mutations de résistance a l'oseltamivir repandues et nouvellement emergentes.

Mots-clés

Sélection avantageuse, age de l'allèle, ABC, adaptation, adaptation convergente, évolution expérimentale, influenza H1N1, balayage sélectif, cliquet de Muller, interférence Hill Robertson

Contents

Chapter 1 Introduction

Evolution represents the "cumulative change over time in the characteristics of a population of living organisms" (Charlesworth & Charlesworth 2010). Some of these changes may be very rapid, such as the evolution of drug resistance in a virus or bacterial population. Other evolutionary changes, such as the modification in physical traits of a species, may extend over much longer time scales. Evolutionary change pre-supposes the existence of variability within populations generated by mutation, which, along with genetic drift, natural selection, and migration, all serve to shape the frequency of such genetic variation. Population genetics is the study of this allele frequency change, with the goal of quantifying the relative contribution and interplay of these forces. In this dissertation, I present computational, empirical, and experimental approaches for addressing this fundamental question.

Computational: One approach to understand the frequency at which positive selection acts, is to estimate the age and strength of beneficial alleles, with these ages providing a proxy for rate. In the first chapter, I develop an Approximate Bayesian Computation (ABC) approach to estimate these parameters in single populations. While many tests have been designed to identify beneficial mutations (see reviews of Thornton *et al.* (2007), Bank *et al.* (2014)) , comparatively few approaches exist to infer the age of these variants. Despite the rapid sojourn time to fixation of beneficial mutations, existing methods primarily estimate age and selection parameters using haplotype-based approaches tailored to the case of mutations segregating at low or intermediate frequency (Slatkin 2008; Peter *et al.* 2012; Chen & Slatkin 2013; Chen *et al.* 2015). In addition, many of the approaches developed rely on time-sampled data. However, apart from experimental population or ancient genomic data, most data is a collected at a single time point. Thus, there is a gap for methods that co-estimate age and selection parameters for fixed beneficial mutations using single time point, single population data.

Our method addresses that gap, integrating a range of existing diversity, site frequency spectrum, haplotype and linkage disequilibrium based summary statistics. It builds on the ABC method previously developed by Przeworski (2003) which is used as a benchmark to assess performance. We show that for strong selective sweeps on *de novo* mutations the method can estimate allele age and selection strength even in non-equilibrium demographic scenarios. We extend our approach to models of selection on standing variation, and co-infer the frequency at which positive selection began to act upon the mutation. This work was published in *Molecular Ecology* in 2016. Code for implementing the method is available through http://jensenlab.epfl.ch/.

Empirical: In the second chapter, I utilize this method in an application to ecology: namely, to estimate the age of beneficial alleles underpinning cryptic coloration in a newly evolved habitat. The study of populations that have recently colonized new environments is of interest because environmental changes can generate strong signatures of selection that are within the time limits of detection from polymorphism data. The recently formed White Sands system of southern New Mexico offers an outstanding example of rapid and convergent adaptation, with a variety of species having evolved blanched forms on the dunes

that contrast with their close relatives in the surrounding dark soil habitat. In this study, we focus on two of the White Sands lizard species, *Sceloporus cowlesi* and *Aspidoscelis inornata*, for which previous research has linked different mutations in the melanocortin-1 receptor gene (*Mc1r*) to the parallel evolution of blanched coloration. In such studies, establishing whether a mutation pre-dates or post-dates a known environmental change provides an indication of whether positive selection operated on new (*de novo*) or pre-existing (standing) variation, a fundamental question underlying the nature of adaptation which can be addressed using our method. Importantly, the colonization of new environments frequently entails demographic changes that may confound signatures of selection. Here, the advantage of our ABC approach is that it can be applied to the *Mc1r* mutations which are segregating at a high frequency in the light phenotype populations, and that both the different colonization histories and dominance effects can be explicitly modelled when co-estimating allele age and selection strength.

For both species, we find that the estimates of allele age are consistent with the respective colonization histories of the two species. In both cases these indicate an age younger than the White Sands (2,000-7000 years), supporting a model of converging adaption from *de novo* mutation. This work was also published in *Molecular Ecology* in 2016.

Experimental: Apart from ecological applications, the study of population dynamics has important implications in medicine. For example, the evolution of drug resistance is a critical public health problem. Influenza in particular inflicts a heavy death toll annually, with approximately 36,000 deaths in the United States alone. Widespread resistance to the most frequently administered drug, oseltamivir, has focused interest in developing therapeutics with novel mechanisms of action that are less susceptible to resistance. Favipiravir is a new compound that is currently in phase three clinical trials. It acts by increasing the genome wide mutation rate of influenza virus (IAV). Combining oseltamivir and favipiravir is a clinically proposed strategy with demonstrated synergies in mouse models of influenza, but the effects of combining the two drugs on virus population evolution are unknown. In chapter 4, I analyse experimental populations of IAV subject to a combined protocol of oseltamivir and favipiravir. To the best of our knowledge, this is the first study attempting to elucidate the underlying evolutionary processes behind the synergistic effects of oseltamivir and favipiravir in H1N1 virus populations.

We describe an interplay between mutation, selection, and genetic drift, which ultimately leads to mutational meltdown. In particular, we find that selective sweeps around oseltamivir resistance mutations reduce genome-wide variation but also hitchhike deleterious mutations to fixation owing to the increased mutational load generated by favipiravir. This latter effect reduces viral fitness and accelerates extinction via Muller's ratchet, but at the risk of spreading both established and newly emerging oseltamivir resistance mutations, if transmission occurs before the virus populations are eradicated. The analysis portion of this work is currently in preparation for submission to *PLoS Genetics*, and the consideration of the underlying models of mutational meltdown, lethal mutagenesis, and error catastrophe resulted in an opinion piece that has been accepted for publication at *Virus Evolution*.

Thus, the work presented in this thesis illustrates how the combination of approaches and questions from population genetics, statistical inference, ecology, and medicine may all serve to illuminate fundamental principles of evolutionary biology.

References

Bank C, Ewing GB, Ferrer-Admettla A, Foll M, Jensen JD (2014) Thinking too positive? Revisiting current methods of population genetic selection inference. *Trends Genet* **30**, 540-546.

Charlesworth B, Charlesworth D (2010) *Elements of Evolutionary Genetics* Roberts and Company Publishers.

Chen H, Hey J, Slatkin M (2015) A hidden Markov model for investigating recent positive selection through haplotype structure. *Theor Popul Biol* **99**, 18-30.

Chen H, Slatkin M (2013) Inferring selection intensity and allele age from multilocus haplotype structure. *G3 (Bethesda)* **3**, 1429-1442.

Peter BM, Huerta-Sanchez E, Nielsen R (2012) Distinguishing between selective sweeps from standing variation and from a de novo mutation. *PLoS Genet* **8**, e1003011.

Przeworski M (2003) Estimating the time since the fixation of a beneficial allele. *Genetics* **164**, 1667-1676.

Slatkin M (2008) A Bayesian method for jointly estimating allele age and selection intensity. *Genet Res (Camb)* **90**, 129-137.

Thornton KR, Jensen JD, Becquet C, Andolfatto P (2007) Progress and prospects in mapping recent selection in the genome. *Heredity (Edinb)* **98**, 340-348.

Chapter 2 Inferring the Age of a Fixed Beneficial Allele

Ormond L, Foll M, Ewing GB, Pfeifer SP, Jensen JD. Mol Ecol. 2016 Jan;25(1):157-69.
doi:10.1111/mec.13478

2.1 Introduction

Selective sweeps are believed to have played a role in shaping genomic patterns of variation across a wide range of species. Estimating the parameters underlying this process, including the beneficial allele age and associated selection strength, can provide deeper insights into the mode and tempo of adaptation. With regards to allele age in particular, one question that has remained of particular focus is whether specifically identified beneficial mutations correspond with the timing of an environmental change experienced by the population in question – be it the colonization of a novel habitat or a sudden geological event. This question is often posed in the context of whether adaptive events more commonly draw on new or standing genetic variation – and indeed significant debate remains around this topic (Jensen 2014). Adaptation from new mutations may be said to be "mutation limited", in that the appropriate mutation would need to occur after the shift in selective pressure. Thus, the ability to accurately infer the age and the starting frequency at the onset of selection of identified beneficial mutations relative to known environmental shifts will be key for advancing this debate.

Many tests have been designed to identify the action of selection in the genome from patterns of polymorphism (see review of Thornton *et al.* (2007) and Bank *et al.* (2014)). These rely on frequency changes in linked neutral sites induced by a selective sweep, a process known as "genetic hitchhiking" (Kaplan *et al.* 1989). Polymorphism based signals are relatively fleeting and are typically visible only on a time scale of 0.1 N_e generations or less, for an effective population size N_e (Przeworski 2003). Yet the majority of approaches are intended to only identify beneficial fixations, and comparatively few approaches exist for inferring the age of these variants. Over the last few years, method development has largely focused on time-sampled datasets, and much progress has been made in this area (e.g. McVean 2002; Malaspinas *et al.* 2012; Mathieson & McVean 2013; Foll *et al.* 2014; Steinrücken *et al.* 2014). However, apart from experimentally evolved or clinical populations, or the handful of ancient genomes, the great majority of available data is collected at a single time point (i.e., present), and there is thus a compelling incentive to improve single time point methods.

Despite the fast transit time characterizing beneficial fixations, the majority of single time point methods to date have aimed to estimate these parameters for segregating, rather than fixed, beneficial mutations using haplotype structure (e.g. Slatkin 2008; Peter *et al.* 2012; Chen & Slatkin 2013; Chen *et al.* 2015) . Most recently, Chen *et al.* (2015) used a Hidden Markov Model to explore haplotype structure, and developed a likelihood estimation approach assuming strong selection (and thus a deterministic allele tra-

jectory) for currently segregating beneficial mutations. For fixed mutations, the state-of-the-art approach was proposed by Przeworski (2003) to estimate the age of a fixed beneficial mutation in an approximate Bayesian (ABC) framework based on a combination of diversity, site frequency spectrum (SFS) and haplotype statistics. We continue this focus in order to develop an improved estimator for the age of fixed beneficial mutations using the past decade of statistical method development, and utilize the Przeworski (2003) estimator as a performance benchmark.

Most notably, the characteristic pattern of linkage disequilibrium (LD) generated by a complete selective sweep suggests the opportunity to utilize this in an ABC framework. Simulation and theoretical studies (e.g. Stephan *et al.* 2006; Jensen *et al.* 2007; McVean 2007; Pavlidis *et al.* 2010) have described strong LD at linked sites on either side of the beneficial fixation, but not spanning the selected site. In addition, there is a reduction in LD across the target of selection. Kim & Nielsen (2004) designed a statistic ω_{max} that captures this complex pattern, with Jensen *et al.* (2007) subsequently demonstrating that ω_{max} exhibits different density distributions under selective sweep models in both equilibrium and non-equilibrium populations.

Here, we explore the combination of frequency spectrum- and linkage disequilibrium-based expectations as an approach to improve our ability to estimate the age of a fixed beneficial mutation based on observed patterns of polymorphism. We develop an ABC-based method that is demonstrated to outperform existing approaches. This approach is not intended to identify loci under selection from genome-wide scans: rather, it is applicable to previously identified loci. We extend this approach to co-estimate allele age and selection strength assuming that selection acts on a *de novo* mutation. Next, we relax the assumption of selection on a *de novo* mutation in order to co-estimate the starting frequency of the segregating allele with allele age and selection strength. Finally, we apply these developed methodologies to explore the selective history of cryptic coloration in a wild deer mouse population, and compare our newly developed estimates with previous published inference.

2.2 Methods

We present three sets of methods. Firstly, we infer allele age T (the time since the allele fixed) alone assuming that the selection coefficient s is known and that a model of selection on *de novo* mutations applies. Secondly, s and T are co-estimated while continuing to assume a model of selection from *de novo* mutation. Thirdly, we co-infer the starting frequency f at which the previously neutral allele was segregating in the population at the onset of selection s and the age at which selection begins T_s. Although the underlying assumption is that a test of selection has been applied using other tools, we demonstrate that this approach has power to correctly infer neutrality as well.

Approximate Bayesian Computation (ABC)

A standard ABC approach was applied following Tavaré *et al.* (1997) and Beaumont *et al.* (2002). We used the R package *abc* (Csillery *et al.* 2012) and implemented the method in the following series of steps:

1. Simulations

For each scenario considered, 5 x 10^5 simulations were generated using the program *msms* (Ewing & Hermisson 2010). Briefly, neutral genealogies are traced backwards in time for a random sample of alleles using standard coalescent theory, incorporating recombination and demographic changes where applicable. Selection is modeled at a single pre-determined locus by applying forward simulations. In this study, selection is assumed to be additive such that genotypes that are homozygous and heterozygous for the selected derived allele have fitness 1+*s* and 1+*s*/2 respectively, whereas genotypes that are homozygous for the ancestral allele have fitness 1. For a sample of *n* chromosomes of length *L*, and assuming an effective diploid population size *N* = 10,000, a coalescent history was constructed assuming a population scaled mutation rate $\theta = 4N_eL\mu$ with mutation rate $\mu = 10^{-7}$ per base pair per generation, and a population scaled recombination rate $\rho = 4N_eLr$ with recombination rate $r = 10^{-7}$ per base pair per generation. Unless a model of selection from standing variation is stipulated, simulations are designed to model selective sweeps from *de novo* mutations arising on a single chromosome in the population, which have ultimately fixed. The strength of selective sweeps is determined using the population scaled parameter $\alpha = 2N_es$. Unless specified otherwise, simulations were run using *L* = 20kb for the inference of *T* alone and using *L* = 10kb for the co-estimation of *s* and *T*, and the selected mutation is positioned in the center of the region. These lengths were chosen to capture the full signature of the selective sweep for the parameters used, based on theoretical results demonstrating an effect over *L* = 0.01 x *s/r* = 10kb for a selective sweep of coefficient *s* = 0.1 and recombination rate $r = 10^{-7}$ crossovers per base pair per generation (Kaplan *et al.* 1989).

Equilibrium populations are modeled as panmictic diploid populations of constant size N_e = 10,000. Allele age *T* is taken to be the time since the allele fixed, using the –SF option in *msms* simulations. For equilibrium demographic scenarios, the prior distributions for *s* and *T* were $log_{10}(s) \sim U(-4, -0.5)$ and $log_{10}(T) \sim U(-4, -0.5)$ where U is a uniform distribution. *T* is reported in units of $4N_e$ generations in keeping with standard coalescent theory. These distributions were chosen in order to span different orders of magnitude from neutrality (where $N_es \leq 1$) to strong selection ($s = 0.3$), and from very recent to distant ages of the selected allele. Przeworski (2002) have shown that $T = 0.1 \times 4N_e$ is approximately the upper limit for detecting selective sweeps, after which the signature in polymorphism data becomes rapidly obscured by subsequent mutation, recombination, and genetic drift.

2. Choice of summary statistics

We used the program *msstats* (Thornton 2003) to calculate a panel of 21 frequently used summary statistics (see Supp. Table 1 for details of statistics) from the standard *msms* Single Nucleotide Polymorphism (SNP) output simulated in step 1 above. Supp. Fig. 1 shows the correlation of a range of informative diversity, SFS and LD based statistics with *s* for recent sweeps ($T = 0.01 \times 4N_e$ generations). For older sweeps ($T = 0.1 \times 4N_e$ generations), we find that the signature of selection becomes rapidly obscured (data not shown). Following Wegmann *et al.* (2009), we employ a partial least squares method (PLS) to incorporate the most informative statistics into our method. PLS is similar to Principal Component Analysis, but determines orthogonal components from a high dimensional set of statistics by maximizing the covariance between the statistics and the variables. Applying PLS has been shown to improve the performance of ABC methods, partly by reducing the dimensionality of the set of summary statistics and partly by removing noise from uninformative statistics (Joyce & Marjoram 2008). Wegmann *et al.* (2009) have shown that incorporating a large number of non-informative summary statistics may bias the resulting posteriors. The *pls* package in R (Bjorn-Helge & Wehrens 2007) was used to calculate PLS components based on a subset of size 10^4 out of the total 5×10^5 simulations. Prior to implementing PLS we apply a Box-Cox transformation (Box & Cox 1964) to normalize the statistics. We adapted a script available through ABC Toolbox for this purpose

(Wegmann *et al.* 2010). Incorporating PLS into our ABC method was shown to reduce relative bias and root mean square error (RMSE), and was therefore used in all ABC calculations.

3. ABC inference of *s* and *T*

In order to evaluate the performance of our ABC method, we selected values of *T* only, or of *s* and *T* over different orders of magnitude, and ran 100 simulations for each selected pair of values that we considered as pseudo-observables. Summary statistics were calculated from the SNP output using *msstats* and transformed into PLS components using the same loadings as in step 2 above.

Posterior distributions for the parameters were generated using an ABC rejection algorithm and a tolerance level of 0.005, which was found to be optimal. 2,500 simulations were therefore retained out of the total number of simulations of 5 x 10^5. Using local linear or ridge regression ABC methods did not significantly improve results (data not shown).

Point estimates for *s* and *T* were calculated from the mode of the joint density posterior distribution using the two-dimensional kernel density function in the MASS package in R (Venables & Ripley 2002). For estimating allele age alone, the mode of the posterior distribution for *T* was calculated to give a point estimate. Relative bias and RMSEs were calculated between these predicted values and the true pseudo-observable values for *s* and *T* (Supp. Tables 2 and 3). Relative bias is defined as the mean difference between the predicted value *y* and the true pseudo-observable y_t divided by the value of the true pseudo-observable y_t. RMSE is defined as the square root of the squared difference between the predicted value *y* and the true pseudo-observable y_t divided by the number of observations *n*:

$$RMSE = \sqrt{\frac{\sum_1^n (y - y_t)^2}{n}}.$$

Non-equilibrium demographic scenarios

For non-equilibrium populations, allele age is taken to be the time since the onset of selection (T_s) by applying the –SI option in *msms*. It is not possible in the current version of *msms* (or in other simulation programs) to model the time since the allele fixed *T* under changing demographic parameters, but only to model the time T_s since the onset of selection. To ensure that the selected allele fixes in simulations, the –SFC option is used to prevent loss owing to genetic drift, and the –oTrace switch is applied to track the frequency of the selected allele in the population through time using a python script. Only simulations where the frequency of the selected mutation is above 0.99 at the time of sampling are retained.

The demographic models are assumed to have been inferred using other methods (e.g., δaδi (Gutenkunst *et al.* 2009), fastsimcoal (Excoffier *et al.* 2013)), and are incorporated in the simulations in step 1 above to run 5 x 10^5 simulations. The selection coefficient *s* is drawn from a log uniform prior as for equilibrium scenarios: $log_{10}(s) \sim U(-4, -0.5)$, and the prior for allele age T_s is adjusted to account for the allele's sojourn time and to ensure that the selected mutation has sufficient time to fix. Based on the analytical derivation of the sojourn time T_{soj} provided by (Stephan *et al.* 1992) of

$$T_{soj} = \frac{2\ln(2N_e)}{s},$$

we adjust the prior for T_s to $log_{10}(T) \sim U(\log10(T_{soj}), \log10(0.3+T_{soj}))$. T_{soj} as calculated here represents the expected sojourn time under equilibrium demography, and is therefore an approximation of the sojourn time under non-equilibrium demography.

Two scenarios were chosen to model size-change events. In both cases, bottlenecks are assumed to occur relatively recently at 0.01 x $4N_e$ in the past. Firstly, we model a shallow and long bottleneck of length 0.02 x $4N_e$ with a 95% reduction in population size, and secondly we model a narrow and severe bottleneck of length 0.002 x $4N_e$ with a 99.8% reduction in population size. These parameters were chosen to be consistent with other studies (Pavlidis *et al.* 2010). In addition, a growth scenario was modeled assuming exponential growth following a bottleneck at 0.01 x $4N_e$ pastward which reduced the population size to 1% of its current size, with a calculated $\alpha = 460.5$. This last scenario was chosen for its similarity to the demographic parameters inferred for *P. maniculatus* deer mice in the Nebraska Sand Hills for the data application presented here (Linnen *et al.* 2013).

Co-estimating allele starting frequency f

The previous sections assume a model of selection acting on a *de novo* mutation. In the third part of our method, we relax this assumption and extend our approach to co-infer the allele frequency *f* when selection begins, along with T_s and *s.* The same steps as for the joint inference of s and T_s in non-equilibrium scenarios described above were applied, but with the additional specification of *f*. The software *msms* allows for *f* to be input using the –SI switch. In simulated samples, *s* is drawn from a log uniform prior $log_{10}(s) \sim U(-4, -0.5)$, and the prior for T_s is adjusted to take account of sojourn time, to $log_{10}(T_s) \sim U(\log_{10}(T_{soj}), \log_{10}(0.3+T_{soj}))$, to give the selected allele sufficient time to fix in the population, as before. The starting frequency *f* is drawn from a log uniform prior $log_{10}(f) \sim U(-4, -0.5)$ spanning the case of selection on a *de novo* mutation (with $N_e=10^4$) to selection on a previously neutral segregating mutation with a frequency of 30%. Point estimates for *s,* T_s and *f* were calculated using the three-dimensional kernel density estimate of the joint posterior mode in the misc3d package (Feng & Tierney 2015).

ω_{max}-ABC methodology

In addition to the *msstats*-based ABC methodology described above, we also derived a methodology to incorporate the statistic ω_{max}. The same steps as for msstats-ABC were implemented with the adjustments detailed in this section.

As described in the introduction, ω_{max} was designed by Kim & Nielsen (2004) to capture the specific LD pattern associated with selective sweeps, and in particular the reduction in LD that occurs across the selected site after a sweep. The statistic ω is defined as

$$\omega = \frac{\left(\binom{l}{2} + \binom{S-l}{2}\right)^{-1} \left(\sum_{i,j \epsilon L} r_{ij}^2 + \sum_{i,j \epsilon R} r_{ij}^2\right)}{(1/l(S-l) \sum_{i \epsilon L, j \epsilon R} r_{ij}^2}$$

At each site l of S polymorphic sites, the statistic splits sites into two groups, from the first to the l^{th} polymorphic site to the left, and from $(l+1)^{th}$ to S polymorphic sites to the right. Within each group, singletons are excluded and the correlation coefficient r^2_{ij} is calculated between the i^{th} and j^{th} sites. The value of l that maximizes ω (ω_{max}) can also be obtained.

Under equilibrium demography, and assuming $T = 0.01 \times 4N_e$, simulations for different selection coefficients generate limited differences in distributions of ω_{max} overall, but do produce a skewed distribution for the top 5% values in selection scenarios compared to neutral simulations (Supp. Fig. 2). This observation holds over different sequence lengths ($L = 10^4$, 5×10^4 and 10^5 bps). This result is consistent with the findings of Jensen *et al.* (2007), who demonstrated via simulation that for large sample sizes ($n = 50$, as in our simulation study) in equilibrium populations, ω_{max} distributions are characterized by a tail of large values in selection scenarios, which increases with the size of selection coefficients.

To incorporate ω_{max} into an ABC framework, 100 simulations were generated in *msms* for each pair of values of *s* and *T* drawn from the priors, but only the top 5% by value of ω_{max} were retained; these were combined with the *msstats* statistics calculated for those simulations. Taking the top 5% of simulations by value of ω_{max} for both the prior and for pseudo-observables replicates the ascertainment process (i.e., significant p-values). Not correcting for such ascertainment in multi-locus genome scans has been shown to generate a high rate of false positives (Thornton & Jensen 2007). The approach of retaining the top 5% simulations by value of ω_{max} is consistent with the idea of an outlier approach where 100 loci are scanned, as done here, and only extreme values in the tails of distributions are retained as possible candidates for sites under selection.

Values of ω and ω_{max} for each simulation were calculated using OmegaPlus (Pavlidis *et al.* 2010).

Application to data on cryptic color adaptation in deer mice

Data on 91 *Peromyscus maniculatus* deer mice was obtained from a previous study by Linnen *et al.* (2013). Briefly, mutations associated with traits underpinning cryptic color adaptation to a light phenotype have been identified in mice living in the Nebraska Sand Hills. A serine deletion at position 128150 on exon 2 has been shown to be associated with several potentially adaptive traits, with a previously estimated selection coefficient of 0.126. Enrichment, sequencing and genotyping are described in Linnen *et al.* (2013). The sequence data was partitioned according to phenotype, and alleles with the serine deletion were extracted from the data set. The data was adjusted to cover a region of 20kb on either side of the deletion. Of the 100 alleles with the serine deletion, 36 were discarded based on a threshold of more than 15% unknown sites. Of the remaining 64 alleles, all cases where the site was unknown for at least one individual were removed. The filtered data contained 418 segregating sites in the 40kb region surrounding the deletion. We explored the impact of changing the filtering to 25% or 10% of individuals with more than 25% unknown sites, but this did not markedly change the results. We applied the ABC method described above for estimating *s* and *T* conditioning on the number of segregating sites *S* as well as θ and ρ. PLS was used to generate components to drive the inference procedure from the msstats statistics after excluding *S* and any invariant statistics. We verified the accuracy of our ABC estimator using simulations with the mouse parameters. Point estimates for *T* and *s* were calculated from the mode of the joint density posterior distribution as before. A point estimate for *T* alone (assuming the previously published estimate of $s = 0.126$) was also derived for comparison purposes. We then explicitly incorporate into our simulations the previously inferred demographic scenario - a bottleneck 2,900 years ago that reduced the population to 0.004 of its original size, followed by an exponential recovery to 65% of its original size (Linnen *et al.* 2013) and estimate *s* and the time of the onset of selection T_s. Finally we co-estimate the starting frequency *f* of the serine deletion with *s* and T_s. We analyzed the data over one additional length, 80kb, with the selected mutation positioned centrally. We obtained a slightly lower sample sizes after applying the filtering process

described above of 48 alleles for the 80kb region. Simulations for the ABC calculation were run assuming N_e=53,080, a mutation rate μ = 3.62 x 10^{-8} and a recombination rate r = 0.62 x 10^{-8} per base pair per generation (all assumptions are from Linnen *et al.* (2013).

Code for implementing the method is available through http://jensenlab.epfl.ch/.

2.3 Results

Inference of allele age (*T*) alone

Initially, we fixed the selection coefficient *s* to be 0.1 (strong selection), 0.01 (moderately strong selection) or 0.001 (weak selection) and we replicated previous results for the inference of allele age only using three statistics (the number of segregating sites *S*, Tajima's *D* and the number of haplotypes), following Przeworski (2003). We sought to improve on these by using the statistics available through *msstats (msstats*-ABC) and by incorporating the ω_{max} statistic (ω_{max} -ABC). All ages *T* are in units of 4*Ne* generations. The choice of whether ω_{max}–ABC or *msstats*-ABC is used will depend on how the location of the mutation has been established, and therefore whether a method that corrects for ascertainment (ω_{max}–ABC) or one that does not correct for ascertainment (*msstats*-ABC) is appropriate (see Discussion). Figure 1A shows the results of inferring allele age *T* for 6 cases (*T* = 0.001, 0.01, 0.05, 0.1, 0.2 and 0.3) assuming strong selection (*s* = 0.1) and a sequence of 20kb. Boxplots represent the distribution of point estimates, which are the modes of posterior distributions. Both *msstats*-ABC and ω_{max}-ABC differentiate age well for 3 orders of magnitude, for *T* = 0.001, *T* = 0.01 and *T* = 0.1, and outperform the previously implemented summary statistics in (Przeworski 2003). Above *T* = 0.05, the age of sweeps is inferred with high accuracy. Relative bias and RMSE estimates support this conclusion (Supp. Table 2). The age of very young sweeps (*T* = 0.001) is underestimated, presumably because the signature of the selective sweeps is not yet apparent in all of the statistics utilized.

For moderate selection (*s* = 0.01), both *msstats*-ABC and ω_{max}-ABC differentiate *T* well between two orders of magnitude rather than three, effectively separating old sweeps ($T \geq 0.1$) from young sweeps ($T < 0.01$) (Fig. 1B). In this case, the estimators significantly improve performance over the statistics employed by Przeworski (2003). In contrast, for weak selection (*s*=0.001), we find that the estimators perform poorly and identify all sweeps as very young (Supp. Fig. 3). Thus, the estimator for *T* alone works well only for strong and moderately strong selection.

We additionally explored the impact of choosing different window sizes surrounding a selected mutation (*L* = 20, 40 and 80 kb) (Supp. Fig. 4). We find that window sizes of 10kb or 20kb provide the best estimates for the parameter ranges investigated here, and that larger window sizes slightly underestimate allele age, due to a dilution of the statistics. This result is in line with theoretical results estimating the size of a swept region subject to a reduction of diversity as $L = 0.01 \times s/r$ (Kaplan *et al.* 1989).

Joint inference of *s* and *T* under equilibrium demography

Here we extended our approach to jointly infer *s* and *T* for fixed mutations using a simple and computationally efficient approach. Simulations demonstrate that for young sweeps (*T* = 0.01) and old sweeps (*T* = 0.1), neutral scenarios (where $N_e s \leq 1$, i.e. *s* = 0.0001 and *s* = 0) can be readily differentiated from selec-

tion scenarios, for both young and old sweeps, using either *msstats*-ABC (Fig. 2C and 2D, Supp. Fig. 5) or ω_{max}–ABC (Supp. Fig. 6 and 7). Additionally, we can infer strong and moderately strong selection (s = 0.1 and s = 0.01) well (Figures 2A and 2B, Supp. Fig. 5-9) using either methodology. One of the weaknesses of both methods is that weak selection (s = 0.001), whether for old or young sweeps, can be misinferred as stronger, older selection (s = 0.01 or s = 0.1) (Fig. 2C and Supp. Fig. 5C). This limitation owes to the fact that the patterns of polymorphism for weak sweeps resemble that of older, stronger sweeps. We find that ω_{max}–ABC is a more accurate estimator of weak selection than *msstats*-ABC (Supp. Table 3 and Supp. Fig. 9).

With regards to allele age, both methods are able to differentiate old sweeps (T = 0.1) from young sweeps (T = 0.01). We find that the methods do not have the power to accurately infer the age of young sweeps but only to establish whether sweeps are either T = 0.01 or younger. Results of inference for very young sweeps (T=0.001) are similar to the results of inference for moderately young sweeps (T=0.01) (data not shown). In contrast, for older sweeps, the additional time may enable different statistics to be impacted at different rates, and for a subset of these statistics to return towards equilibrium. Simulation studies have shown that statistics reliant on intermediate frequency alleles such as Fay and Wu's H decay rapidly after a selective sweep and retain very little signal at 0.1 x $4N_e$, whereas statistics reliant on singletons such as Tajima's D retain a signal longer (Przeworski 2002). The decay of both of these types of statistics at different rates most likely underpins the accuracy of both estimators to infer T and s for sweeps of T = 0.1.

If a model of *de novo* mutation is assumed ($f=1/2N_e$), a pseudo-observable of selection from a rare mutation ($f \leq 0.01$) will have an inferred allele age as the time at which selection starts to act on that mutation. In contrast, inference from a pseudo-observable of selection from high levels of standing variation (f>0.01) will be erroneous and usually indicate an older age (data not shown).

Co-estimating allele starting frequency *f* under equilibrium demography

In this section, we relax the assumption that selection proceeds from *de novo* mutation, and allow the frequency of the selected allele to be co-inferred along with the time at which selection starts T_s (in contrast to the previous section, in which the time T since fixation is inferred), and the selection coefficient s. Analytical derivations by Stephan *et al.* (1992) predict that if $f < 1/2N_es$ and selection is strong, the reduction in linked neutral diversity associated with selection from rare mutations should resemble that from selection on *de novo* mutations. Subsequent analysis has shown that selection from either a *de novo* mutation or from a rare mutation results in a classical "hard sweep" pattern where a single copy of the mutation is swept to fixation (Orr & Betancourt 2001; Hermisson & Pennings 2005). In contrast, selection from high levels of standing variation ($f >> 1/2N_es$) results in the fixation of multiple haplotypes in a "soft sweep" pattern (the other common definition of a soft sweep, in which haplotype diversity is the result of multiple beneficials, is not considered here). In line with theoretical expectations, simulations by Przeworski *et al.* (2005) showed similar patterns of reduction in diversity for $f=1/2N_e$, f=0.001 and f=0.01 (where $N_e=10^4$ and s=0.05), but almost no reduction in diversity for selection from high levels of standing variation (f=0.05 and f=0.20). Here we show results that are consistent with these previous findings. For strong and moderately strong selection (s=0.1 and s=0.01) in equilibrium populations, we find that the ABC estimator performs well for inferring f, s and T_s as long as the pseudo-observable satisfies the condition $f<1/2N_es$, (i.e., the cases where $f=1/2N_e$, f=0.001, f=0.01) (Fig. 3 and Suppl. Fig. 10). We find marginally better inference for s=0.1 than for s=0.01 right up to f=0.01, which appears to be the cut-off for accurate inference. For weak selection (s=0.001), T_s and s are well inferred but f is not (Supp. Fig. 10D).

In contrast, when the pseudo-observable sweep is characterized by $f>>1/N_e s$, we generally identify that $f>>1/2N_e s$ - but with the drawback that s and T_s are poorly co-estimated. As levels of standing variation increase, rising haplotype diversity means the method infers older, weaker sweeps than are the case for the pseudo-observables.

Robustness to non-equilibrium demography

Non-equilibrium demography can mimic signatures of selection (e.g. Przeworski 2002; Jensen *et al.* 2005) and compromise inference. However, an advantage of the ABC approach is that demographic parameters can be explicitly modeled in simulations, and therefore demography can be taken into account in the inference method. We explored the robustness of the method for inferring firstly, s and T_s, and secondly s, T_s and f under three non-equilibrium scenarios where demographic parameters are explicitly known and modeled: 1) a shallow and long bottleneck of length 0.02 x $4N_e$ with a 95% reduction in population size, 2) a narrow and severe bottleneck of length 0.002 with a 99.8% reduction in population size, and 3) an exponential growth scenario following a sharp 99% reduction in population size with α = 460.5. The bottlenecks are modeled to occur at T = 0.01 x$4N_e$. Our results are described for *msstats*-ABC but similar results were obtained for ω_{max}–ABC.

In the case of inferring s and T_s, Fig. 4 and Supp. Fig. 13 show the results for the third scenario of a bottleneck followed by exponential growth. We find that for old (Fig. 4A and Supp. Fig. 11A), young (Fig. 4B and Supp. Fig. 11B) and very young (Fig. 4C and Supp. Fig. 11C) sweeps, both T_s and s are well inferred. We note that it is difficult to distinguish weak sweeps, where s=0.001 (Supp. Fig 11D and E), from neutral scenarios (Fig. 4D and Supp. Fig. 11F). Similar results were obtained for the two bottlenecks scenarios (data not shown). One of the reasons for the estimator's strong results in non-equilibrium populations – and indeed its limitation - is that T_s rather than T is inferred. In our methodology the prior for T_s is set as a function of s by using an estimate of sojourn time T_{soj} in equilibrium populations. This analytical derivation most likely overestimates T_{soj} for alleles fixing in bottlenecked populations, and therefore, the full potential parameter space for T_s is not covered by our adjusted prior. This shortcoming could be corrected by using another set of simulations, rather than the analytical derivation of Stephan *et al.* 1992, to estimate the minimum T_{soj} under a specific demographic scenario for a mutation of strength s. This would give a broader and more accurate prior from which to draw T_s.

When co-inferring f with T_s and s, we find the performance of the estimator deteriorates under non-equilibrium demography (Suppl. Fig. 12). For both young and old sweeps, inference is only reliable if the method correctly identifies a *de novo* or very rare mutation ($f \leq 0.001$). It is difficult to correctly infer f and therefore to establish whether co-estimates of T_s and s are robust. As for equilibrium populations, a high level of standing variation is inferred as an older, weaker sweep than is the case for the pseudo-observable, due to high levels of diversity.

Data Application: mouse coat color evolution

We applied our methods to a previously published data set for 91 *P. maniculatus* deer mice living in the recently formed Nebraska Sand Hills (estimated age 8,000 years) (Linnen et al. 2013). The data set was adjusted to cover SNPs over 20kb on either side of a serine deletion on exon 2 which has been implicated in

several traits associated with cryptic color adaptation to a light phenotype for predator avoidance. After filtering for the serine deletion and genotyping quality, we retain SNP data from 64 alleles for analysis, and remove any further unknown sites. Firstly, our aim was to co-estimate *s* and *T* assuming an equilibrium population of 53,080. Secondly, we estimate allele age *T* alone assuming a previously published estimate of *s*=0.126 (Linnen *et al.* 2013). Thirdly, we explicitly model the demographic scenario that had been previously inferred in our simulations (of a bottleneck 2,900 years ago which reduced the population to 0.04% followed by an exponential recovery to 0.65% of the original population size). Lastly, we co-estimate *f* with s and T_s. We used msstats-ABC as this is consistent with the initial identification of the selected site described in Linnen *et al.* (2013). We simulated pseudo-observables with the specific mouse parameters to establish how well our methods work before applying these to the data set.

Assuming an equilibrium population, the joint inference of *s* and *T* showed a young, moderately strong selective sweep, with an inferred *s* of 8.7×10^{-3} ($1.1 \times 10^{-4} – 3.3 \times 10^{-2}$) and $T \leq 0.01$ (Fig. 5A), using a window size of 80kb to ensure that diversity patterns are fully captured. Applying a window of 40kb reduced the signal of the sweep (Suppl. Fig. 13A). If *s* is assumed to be 0.126, as estimated in Linnen *et al.* (2013), the inference of allele age alone gives the same result of a young or very young sweep with $T \leq 0.01$ (Supp. Fig. 14).

If the demography inferred in Linnen *et al.* (2013) is explicitly included in the simulations for the ABC calculation, simulations using pseudo-observables show that signals from selective sweeps with T_s coincident with or older than the bottleneck are usually quenched, leading to the inference of neutral scenarios (data not shown). In contrast, *s* and T_s for strong sweeps that are younger than the bottleneck (of the order of T_s=0.005) are accurately inferred. This result illustrates the importance of using simulations to establish the limits of inference for specific scenarios. In applying our method to the mouse data, we infer a strong, recent sweep, with $s = 1.7 \times 10^{-1}$ (1.5×10^{-4} - 3.0×10^{-1}) and $T_s = 1.1 \times 10^{-3}$ ($5.8 \times 10^{-4} – 8.8 \times 10^{-1}$) (Fig. 5B). These results are consistent with those obtained under equilibrium demography but with a stronger estimate of *s.* Using a length of 40kb, we find the qualitatively similar result of a strong recent sweep (Supp. Fig. 13B). We also find that simulated pseudo-observable sweeps that are either coincident or older than the mouse bottleneck are sometimes correctly inferred over this length, which is an improvement over the 80kb length (data not shown), but are mostly inferred as neutral.

Our method is subject to the limitation that an estimate of sojourn time under equilibrium demography is used to set the prior for T_s, under the assumption that the mutation fixes. In our simulations, very recent values of T_s are therefore only associated with strong *s.* Here, we have checked with simulations that sojourn time is longer than under the equilibrium scenario, and therefore that the prior for T_s is broader than required, to reduce this source of error.

Co-inferring *f, s* and T_s jointly supports a recent, strong sweep acting on a *de novo* or rare mutation (s=$8.6\text{x}10^{-2}$($1.5\text{x}10^{-4}$-$2.9\text{x}10^{-1}$), T_s=$2.1\text{x}10^{-3}$($7.5\text{x}10^{-4}$-$9.0\text{x}10^{-1}$) and f=$2.6\text{x}10^{-5}$($1.1\text{x}10^{-5}$-$1.9\text{x}10^{-1}$)) (Fig. 5C). Simulations underpinning this estimate incorporate the demographic scenario from Linnen *et al.* (2013). In comparison with the age of the Sand Hills (i.e., 0.075 in units of 4*Ne* generations, assuming one generation every 6 months), these results support previous claims of selection acting on a young *de novo* mutation subsequent to the environmental change.

2.4 Discussion

We present ABC methods that estimate allele age, selection strength and starting frequency for fixed mutations using single population, single time-point datasets. We demonstrate that it is possible to distinguish between different orders of magnitude of the selection coefficient *s*, between old and young sweeps, and between *de novo/*rare and common starting frequencies. There are significant differences between the ABC method that integrates ω_{max} and the *msstats*-ABC, which undermine a direct comparison between the two methods. Namely, one takes account of ascertainment bias while the other does not. The ω_{max} approach was designed to be consistent with an approach for identifying sites under selection using the top ω_{max} values. Our simulations show that ω_{max} marginally outperforms a simple msstats-ABC approach, particularly in estimating parameters for weak sweeps, as it is able to leverage a statistic that captures the specific LD pattern existing immediately after a selective fixation, but this is conditional on it being the appropriate method for the data analyzed.

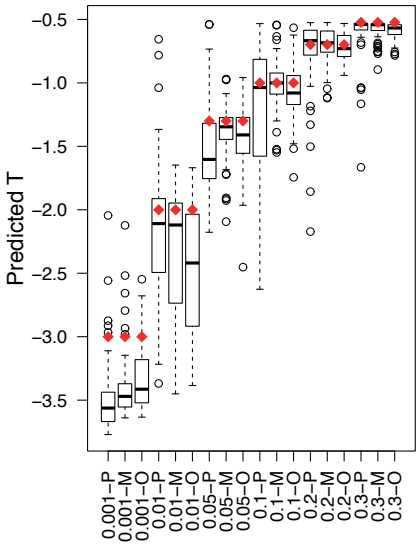
One of the major advantages of an ABC approach is that demography can be explicitly accounted for in simulations, which removes a source of error in estimating the strength of selective sweeps. Here we illustrate this by explicitly including the previously estimated demographic model for our ABC estimation of *s* and *T* in deer mice. We find results that are consistent with those obtained under the assumption of an equilibrium population, but with slightly stronger estimates of selection. We also find that we can distinguish cases of selection on *de novo* and rare mutations from selection on common standing variation resulting in soft sweeps, and in the first case we are able to co-infer T_s and *s* to within an order of magnitude, assuming equilibrium demography. Here, we find the most likely model to be one of selection on *de novo* or rare mutation. This is consistent with our estimates of allele age and provides support for the previously published notion of mutation-limited adaptation underpinning cryptic coloration in deer mice (Linnen *et al.* 2009; Linnen *et al.* 2013; Poh *et al.* 2014).

Many haplotype methods such as iHS rely on a comparison between haplotype lengths for ancestral and derived alleles, and therefore have power to detect selected mutations at low or intermediate frequencies (Voight *et al.* 2006). Beyond this frequency level, power declines because these methods depend on a comparison with alternative allele haplotype structure. For example, the method published by Chen *et al.* (2015) applies to alleles under strong selection that are not yet fixed. Peter *et al.* (2012) use a range of haplotype and SFS based statistics including EHH and iHS to estimate allele age and selection coefficients for segregating mutations in models of *de novo* mutation and standing variation. The importance sampling method developed by Slatkin (2008) is specifically designed to identify *s* and *T* for low frequency alleles such as the A-allele of G6PD in Africa. In contrast to these methods predicated mainly on haplotype structure, our methods use SFS based statistics that are sensitive to different parts of the SFS, as well as LD- and haplotype-based statistics that recover to equilibrium at different rates. Our methods thus fit an important niche, and may be utilized to infer the relative age, strength and frequency of fixed beneficial mutations relative to the timing of environmental shifts – in order to quantify, for example, the age of variants conferring cryptic coloration following the last ice age, as seen here in the mouse example and in the Laurent, Pfeifer *et al.* example in lizard populations also appearing in this issue.

2.5 Figures

Fig. 1 Inference of allele age *T* alone. Boxplots compare results from msstats-ABC (marked M) and ω_{max} – ABC (marked O) with the Przeworski 2003 ABC method (marked P). Boxplots represent the modes of posterior distributions for inferring *T* alone for 100 pseudo-observables. The value of *s* is assumed to be known: A. *s*=0.1 (α = $2N_e s$ = 2000) and B. *s*=0.01 (α = $2N_e s$ = 200). *T* is drawn from a log uniform prior: $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods, with *L* = 20kb. Red diamonds indicate the true values for each case (*T*=0.001,0.05,0.01,0.1,0.2,0.3).

A.



B.

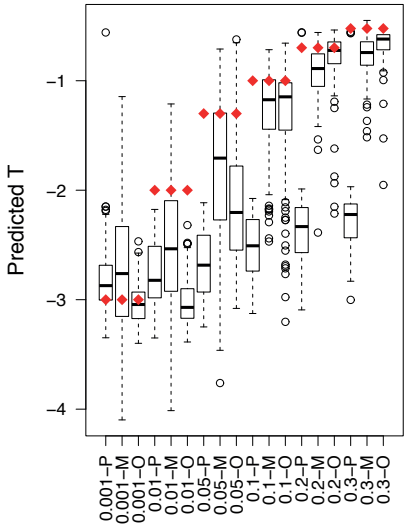
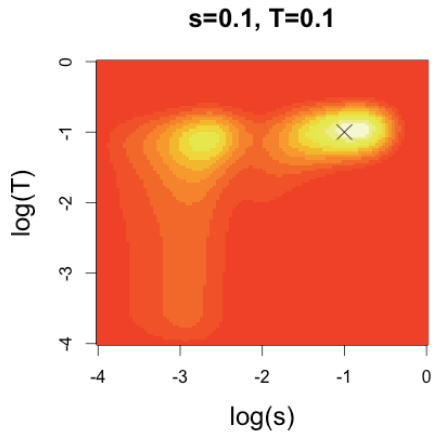
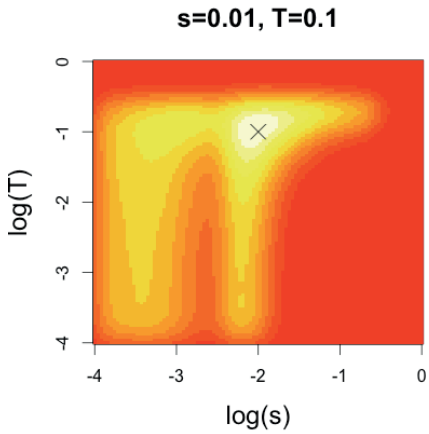


Fig. 2 Joint inference of *s* and *T* in equilibrium populations for old sweeps (*T*=0.1) (msstats-ABC). Figures show the cumulative joint posterior density plots for 100 pseudo-observable simulations over different orders of magnitude of the selection coefficient *s*, for old sweeps (*T*=0.1) and A. *s* = 0.1 ($\alpha=N_es=10^3$); B. *s*=0.01($\alpha=N_es=10^2$); C. *s*= 0.001($\alpha=N_es=10$). The bottom two panels represent neutral scenarios with D. *s*=0.0001 ($\alpha=N_es=1$); and E. *s*=0. The white, yellow and red colors mark areas of high, moderate and low joint density respectively. Black crosses indicate the true values of pseudo-observables. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods.

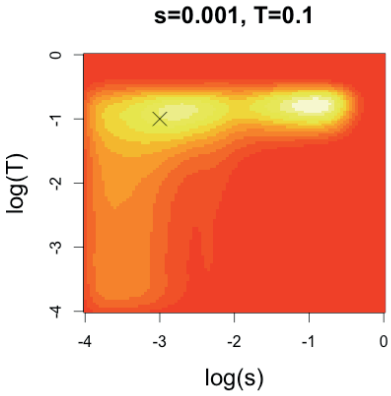
A.



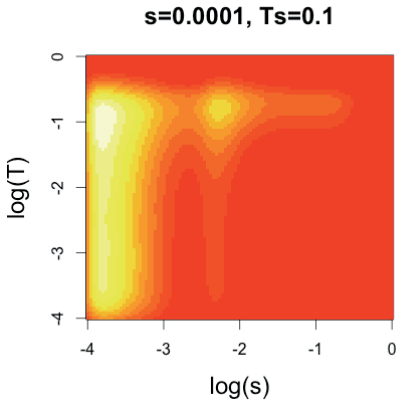
B.



C.



D.



E.

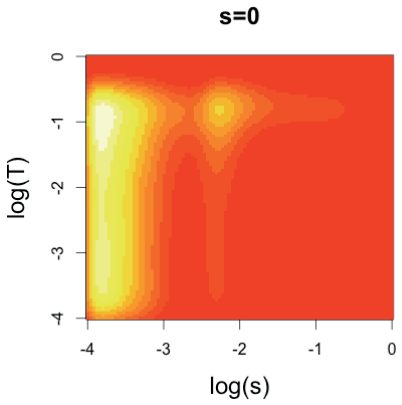


Fig. 3 Joint inference of *s, T_s* and *f* in equilibrium populations. Figures show the predicted values for 100 pseudo-observables for the example of *s*=0.01, T_s=0.060 and *f*=0.0001, 0.001, 0.01, 0.05, 0.1. Estimates of *s,* T_s and *f* were obtained from the mode of the joint posterior density. Red lines indicate the known values of the pseudo-observables. T_s represents the time since selection began acting on the allele (calculated using T_s= time to fixation (*T*=0.01) + sojourn time T_{soj})

F

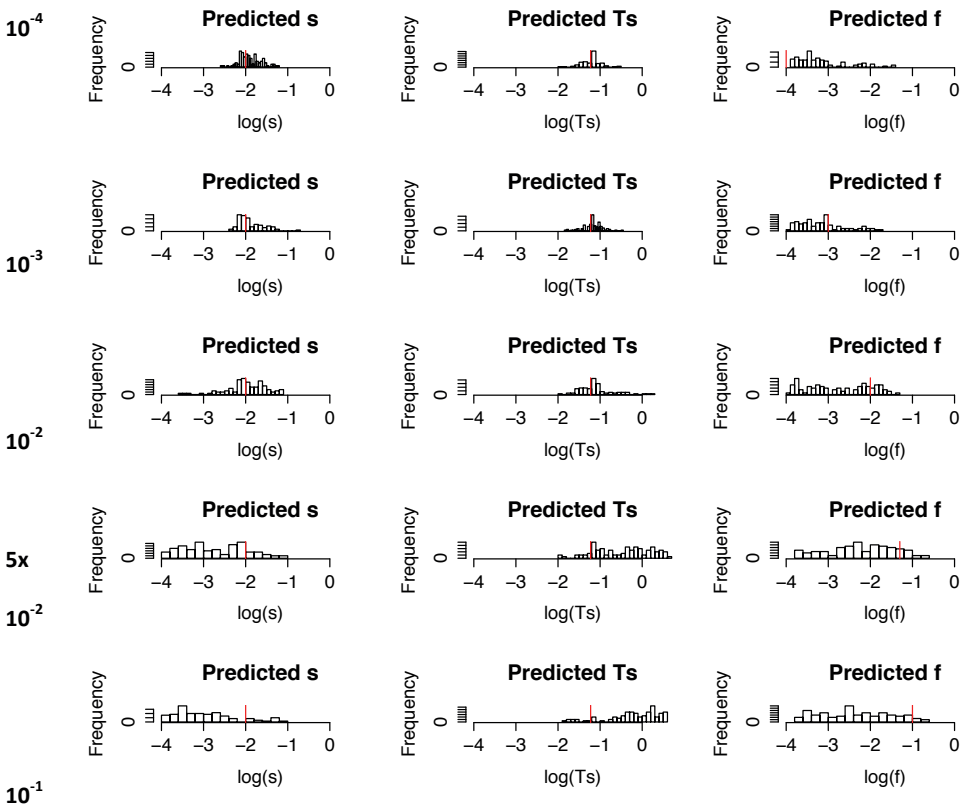


Fig. 4 Joint inference of *s* and T_s in demographic model for a strong bottleneck followed by exponential growth (demographic model 3) using *msstats* -ABC. Figures show the cumulative joint posterior density plots for 100 pseudo-observable simulations. *s* is drawn from a log uniform prior: $\log_{10}(s) \sim U(-4, -0.5)$ and T_s from an adjusted log uniform prior: $log_{10}(T_s) \sim U(\log_{10}(T_{soj}), \log_{10}(0.3+T_{soj}))$. For the pseudo-observables, T_s is calculated from $T_s=T+T_{soj}$ where T is the time since fixation and the sojourn time $T_{soj}=(2\ln(2N_e)/s)/4N_e$. The white, yellow and red colors mark areas of high, moderate and low joint density respectively. Black crosses indicate the true values of pseudo-observables.

A. Inference for a moderately strong, old sweep with pseudo-observable values s=0.01 and T_s=0.15 (calculated from $T_s=T+T_{soj}$ where T=0.1). B. Inference of a strong, very recent sweep with pseudo-observable values s=0.01, T_s=0.006 (calculated from $T_s=T+T_{soj}$ where T=0.001) C. Inference of a strong, very recent sweep with pseudo-observable values s=0.01, T_s=0.006 (calculated from $T_s=T+T_{soj}$ where T=0.001). D. Results of inference where no selected mutation was included in simulations.

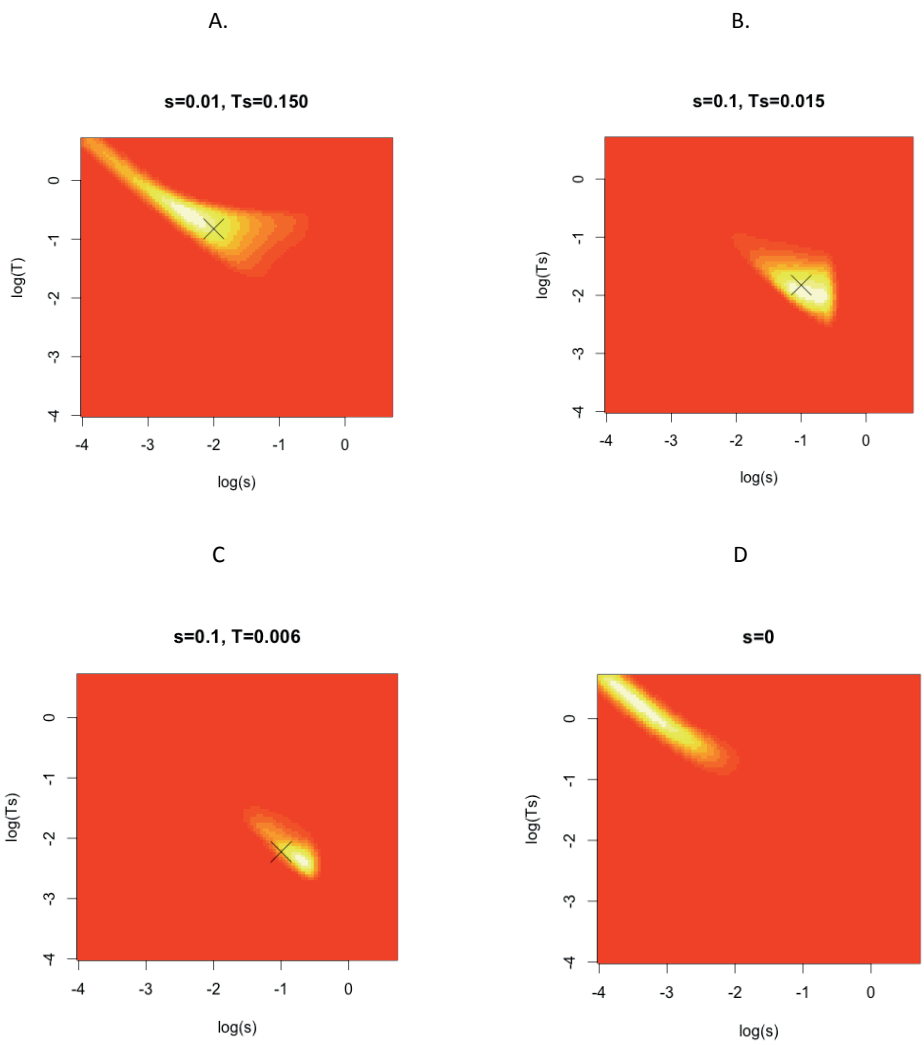
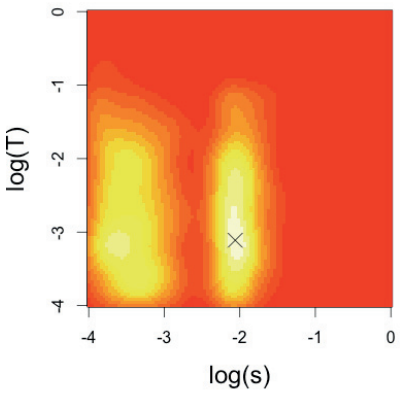


Fig. 5 Joint inference of allele age, selection coefficient and starting frequency for *P. maniculatus*

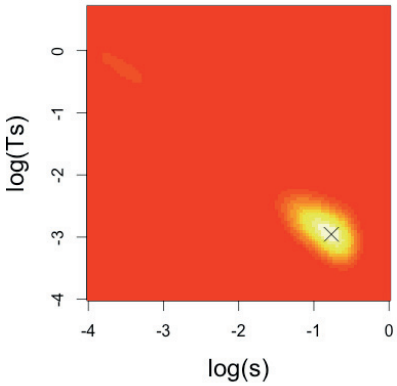
The joint density plots in A and B represent the results of the joint inference for the serine deletion at position 128,150 on exon 2, in A) for *s* and *T* assuming an equilibrium population with N_e= 53,080 and in B) for *s* and T_s with the demographic scenario inferred in (Linnen *et al.* 2013) explicitly included in simulations. In C) *f* is co-inferred with T_s and *s* also assuming the demographic scenario inferred in (Linnen *et al.* 2013); histograms represent the posterior distributions from the ABC inference, with the red lines indicating the mode of the joint posterior density for the three parameters. The density plots are shown for *L*=80kb, with the mutation positioned centrally (x=0.5). Other parameters for deer mice simulations are as described in methods.

For A, the mode of the joint density occurs at s=8.7x10^{-3}(1.1x10^{-4}-3.4x10^{-2}) and T=7.7x10^{-4} (1.2x10^{-4}-1.0x10^{-1})
For B, the mode occurs at s=1.7x10^{-1}(1.5x10^{-4}-3.0x10^{-1}) and T_s=1.1x10^{-3}(5.8x10^{-4}-8.8x10^{-1})
For C, the mode occurs at s=8.6x10^{-2}(1.5x10^{-4}-2.9x10^{-1}), T_s=2.1x10^{-3}(7.5x10^{-4}-9.0x10^{-1}) and f=2.6x10^{-5}(1.1x10^{-5}-1.9x10^{-1})

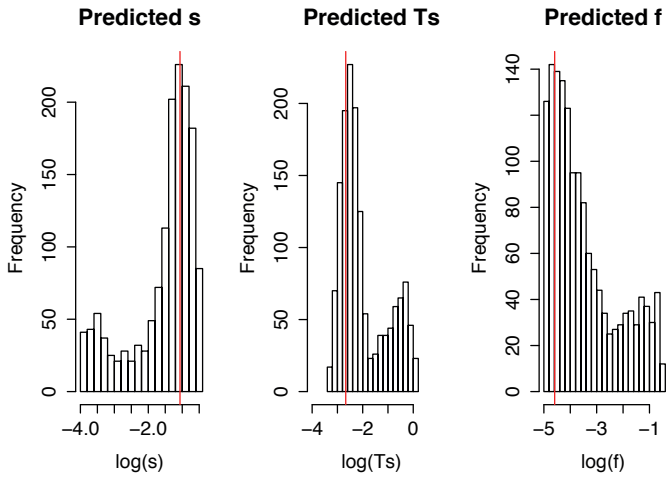
A.



B.



C.



References

Bank C, Ewing GB, Ferrer-Admettla A, Foll M, Jensen JD (2014) Thinking too positive? Revisiting current methods of population genetic selection inference. *Trends Genet* **30**, 540-546.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025-2035.

Box G, Cox D (1964) An analysis of transformations, pp. 211-243. JR Stat Soc.

Chen H, Hey J, Slatkin M (2015) A hidden Markov model for investigating recent positive selection through haplotype structure. *Theor Popul Biol* **99**, 18-30.

Chen H, Slatkin M (2013) Inferring selection intensity and allele age from multilocus haplotype structure. *G3 (Bethesda)* **3**, 1429-1442.

Csillery K, Francois O, Blum M (2012) abc: an R package for approximate Bayesian computation (ABC), pp. 475-479. Methods in Ecology and Evolution.

Ewing G, Hermisson J (2010) MSMS: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics* **26**, 2064-2065.

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLoS Genet* **9**, e1003905.

Feng D, Tierney L (2015) Package "misc3d", CRAN repository.

Foll M, Poh YP, Renzette N*, et al.* (2014) Influenza virus drug resistance: a time-sampled population genetics perspective. *PLoS Genet* **10**, e1004185.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet* **5**, e1000695.

Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335-2352.

Jensen JD (2014) On the unfounded enthusiasm for soft selective sweeps. *Nat Commun* **5**, 5281.

Jensen JD, Kim Y, DuMont VB, Aquadro CF, Bustamante CD (2005) Distinguishing between selective sweeps and demography using DNA polymorphism data. *Genetics* **170**, 1401-1410.

Jensen JD, Thornton KR, Bustamante CD, Aquadro CF (2007) On the utility of linkage disequilibrium as a statistic for identifying targets of positive selection in nonequilibrium populations. *Genetics* **176**, 2371-2379.

Joyce P, Marjoram P (2008) Approximately sufficient statistics and bayesian computation. *Stat Appl Genet Mol Biol* **7**, Article26.

Kaplan NL, Hudson RR, Langley CH (1989) The "hitchhiking effect" revisited. *Genetics* **123**, 887-899.

Kim Y, Nielsen R (2004) Linkage disequilibrium as a signature of selective sweeps. *Genetics* **167**, 1513-1524.

Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE (2009) On the origin and spread of an adaptive allele in deer mice. *Science* **325**, 1095-1098.

Linnen CR, Poh YP, Peterson BK*, et al.* (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312-1316.

Malaspinas AS, Malaspinas O, Evans SN, Slatkin M (2012) Estimating allele age and selection coefficient from time-serial data. *Genetics* **192**, 599-607.

Mathieson I, McVean G (2013) Estimating selection coefficients in spatially structured populations from time series data of allele frequencies. *Genetics* **193**, 973-984.

McVean G (2007) The structure of linkage disequilibrium around a selective sweep. *Genetics* **175**, 1395-1406.

McVean GA (2002) A genealogical interpretation of linkage disequilibrium. *Genetics* **162**, 987-991.

Mevik B-H, Wehrens R (2007) The pls package: principal component and partial least squares regression in R. Journal of Statistical Software.

Orr HA, Betancourt AJ (2001) Haldane's sieve and adaptation from the standing genetic variation. *Genetics* **157**, 875-884.

Pavlidis P, Jensen JD, Stephan W (2010) Searching for footprints of positive selection in whole-genome SNP data from nonequilibrium populations. *Genetics* **185**, 907-922.

Peter BM, Huerta-Sanchez E, Nielsen R (2012) Distinguishing between selective sweeps from standing variation and from a de novo mutation. *PLoS Genet* **8**, e1003011.
Poh YP, Domingues VS, Hoekstra HE, Jensen JD (2014) On the prospect of identifying adaptive loci in recently bottlenecked populations. *PLoS One* **9**, e110579.
Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics* **160**, 1179-1189.
Przeworski M (2003) Estimating the time since the fixation of a beneficial allele. *Genetics* **164**, 1667-1676.
Przeworski M, Coop G, Wall JD (2005) The signature of positive selection on standing genetic variation. *Evolution* **59**, 2312-2323.
Slatkin M (2008) A Bayesian method for jointly estimating allele age and selection intensity. *Genet Res (Camb)* **90**, 129-137.
Steinrücken M, Bhaskar A, Song YS (2014) A NOVEL SPECTRAL METHOD FOR INFERRING GENERAL DIPLOID SELECTION FROM TIME SERIES GENETIC DATA. *Ann Appl Stat* **8**, 2203-2222.
Stephan W, Song YS, Langley CH (2006) The hitchhiking effect on linkage disequilibrium between linked neutral loci. *Genetics* **172**, 2647-2663.
Stephan W, Wiehe THE, Lenz MW (1992) The effects of strongly selected substitutions on neutral polymorphisms: analytical results based on diffusion theory, pp. 237-254, Theoretical Population Biology.
Tavaré S, Balding DJ, Griffiths RC, Donnelly P (1997) Inferring coalescence times from DNA sequence data. *Genetics* **145**, 505-518.
Thornton K (2003) Libsequence: a C++ class library for evolutionary genetic analysis. *Bioinformatics* **19**, 2325-2327.
Thornton KR, Jensen JD (2007) Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* **175**, 737-750.
Thornton KR, Jensen JD, Becquet C, Andolfatto P (2007) Progress and prospects in mapping recent selection in the genome. *Heredity (Edinb)* **98**, 340-348.
Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*, Fourth edn. Springer.
Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. *PLoS Biol* **4**, e72.
Wegmann D, Leuenberger C, Excoffier L (2009) Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics* **182**, 1207-1218.
Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* **11**, 116.

Chapter 3 The Population Genomics of Rapid Adaptation: Disentangling Signatures of Selection and Demography in White Sands Lizards

Laurent S, Pfeifer SP, Settles ML, Hunter SS, Hardwick KM, **Ormond L**, Sousa VC, Jensen JD, Rosenblum EB. Mol Ecol. 2016 Jan;25(1):306-23. doi: 10.1111/mec.13385

3.1 Introduction

The study of populations that have recently colonized novel environments has remained an area of particular interest in evolutionary biology - both because newly encountered environmental pressures can generate strong natural selection, and because adaptation over short time scales can be detectable on the genomic level. However, recent and rapid colonization events are also frequently characterized by severe demographic perturbations (e.g., in population size and migration rates), which may obscure genomic patterns of selection (e.g., Przeworski 2002; Jensen *et al.* 2005; Thornton & Jensen 2007). Although various test statistics have been developed to circumvent this challenge (e.g., Nielsen *et al.* 2005; Jensen *et al.* 2007; Pavlidis *et al.* 2012), the performance of these approaches is often poor under the demographic scenarios underlying colonization, particularly for cases of severe population bottlenecks and ongoing gene flow (Crisci *et al.* 2013; Poh *et al.* 2014). As a result, the underlying demographic history of a population should be explicitly modeled when searching for targets of natural selection.

Although disentangling selection and demography remains difficult, the use of population-level, genome-scale data from recently diverged natural populations can help to discern the relative impact of these factors. The advent of next-generation sequencing technologies together with new computational and statistical techniques to model demographic histories (Thornton & Andolfatto 2006; Gutenkunst *et al*. 2009; Naduvilezhath *et al.* 2011; Excoffier & Foll 2011; Mathew & Jensen 2015) have enabled more accurate inference of demographic history, which can then be used as a null model when scanning for genomic targets of selection. As a result, substantial advances have been made in detecting genes contributing to phenotypic changes in populations that have undergone recent adaptation in the wild such as *Drosophila*, stickleback, mice, and humans (e.g., Reusch *et al.* 2001; Ihle *et al.* 2006; Domingues *et al.* 2012; Linnen *et al.* 2013) and during domestication (e.g., Doebley 2004; Pollinger *et al.* 2005; reviewed in Stinchcombe & Hoekstra 2008; Ellegren & Sheldon 2008; Mackay *et al.* 2009; Stapley *et al.* 2010). However, even in these cases, the underlying genetic variants responsible for the observed changes often remain difficult to identify (Chan et al. 2010).

Melanin-based pigmentation has long been studied as a model for understanding adaptive evolution (Cott 1940; Norris & Lowe 1964), and pigmentation phenotypes are some of the best examples of adaptation where the underlying genetic variants are well characterized. Coloration is involved in a range of

different biological processes from crypsis to mimicry to thermoregulation to sexual signaling, and melanin-based coloration is conserved across many taxa (Thayer 1909; Cott 1940; Norris 1967; Kettlewell 1973; Majerus 1998; Bittner *et al.* 2002; Caro 2005). Although there are a number of genes that affect melanin-based phenotypes, functional changes in the 1kb coding region of the melanocortin-1 receptor gene (*Mc1r*) are responsible for color variation in many species. The melanocortin-1 receptor is an important component of the melanin-synthesis signal transduction pathway in vertebrates (Barsh 1996) and mutations in *Mc1r,* which are responsible for color variation, have been studied extensively in domesticated animals (e.g., dogs (Newton *et al.* 2000), pigs (Kijas *et al.* 1998), horses (Marklund *et al.* 1996), chickens (Takeuchi *et al.* 1996)) and in wild populations (e.g., mice (Nachman *et al.* 2003; Hoekstra et al. 2006), birds (Theron *et al.* 2001; Mundy *et al.* 2004), felines (Eizirik *et al.* 2003), reptiles (Rosenblum *et al.* 2004)). However, in many of the studied species, observed color variation does not correspond to specific environmental pressures, making the connection of *Mc1r* variants to fitness unclear.

The White Sands system in southern New Mexico provides an opportunity to link color variation in wild populations with adaptation to their habitats, and thus to understand the interplay between natural selection and population demography. White Sands is a distinctive landscape of stark white gypsum dunes (~275 square miles), which contrast dramatically with the dark substrate of the surrounding Chihuahuan Desert. There has been dramatic convergence in dorsal color morphology by the White Sands fauna. All of the lizard species that inhabit White Sands and a subset of arthropods and mammals exhibit blanched forms on the gypsum dunes that contrast with dark forms in the rest of their ranges (Smith 1943; Lowe & Norris 1956; Rosenblum 2006). The light coloration of White Sands animals is likely an adaptation for crypsis to avoid detection by visually hunting avian predators, which preferentially predate on poorly background-matched prey (e.g., Dice 1947; Kaufman 1974; Luke 1989).

The demographic context in which natural selection has operated at White Sands has also been dynamic. The white habitat represents a geologically recent change in selective environment, with the bulk of the gypsum deposition having occurred within the last 2,000-7,000 years (Kocurek *et al.* 2007; S. Fryberger unpublished manuscript). Therefore, White Sands populations are expected to result from relatively recent colonizations. Moreover, there are no physical barriers separating the white sands from the surrounding dark desert soils, and the transition between white sand and dark soil habitats occurs abruptly. Thus, adaptation appears to have occurred in many species despite ongoing gene flow (Rosenblum & Harmon 2011).

Here we use population-level genomic data to understand the demographic history and dynamics of natural selection at the genome level for dark and light populations of two White Sands lizards, the Southwestern Fence Lizard (*Sceloporus cowlesi*) and the Little Striped Whiptail (*Aspidoscelis inornata*). In each species, a single *Mc1r* amino acid substitution associated with blanched coloration has been identified through candidate gene studies and functional assays (Rosenblum *et al*. 2004; Rosenblum *et al*. 2010). Population studies and functional assays have also demonstrated that the mutations have different dominance effects in the two species: the blanched allele appears to be dominant in *S. cowlesi* but recessive in *A. inornata* (Rosenblum *et al.* 2010). As a result, this system represents one of the first examples from a natural population where differing predictions regarding hitchhiking effects of recessive and dominant mutations may be directly studied by comparing molecular signatures of natural selection for alleles of the same gene with different dominance effects in two species inhabiting the same novel environment.

In this study, we constructed fosmid libraries and developed a sequence capture approach to obtain >50kb of sequence around *Mc1r* and hundreds of other random genomic locations in each species. We then use model-based statistical inference methods to infer the demographic history of the two popula-

tions using randomly selected genomic regions, investigate the evidence of selection around the candidate sites in *Mc1r*, estimate the age of these mutations relative to the geological age of White Sands, and finally discuss these results in the light of existing population genetic theory. We find that white and dark soil populations show only weak background genomic differentiation but display striking genetic differences in the *Mc1r* gene region. The patterns of variation that we observe at and around the *Mc1r* gene are consistent with strong selective sweeps caused by the non-synonymous mutations previously associated with the blanched phenotype. Furthermore, the signatures of selection at *Mc1r* are consistent with the inferred dominance of the two beneficial mutations.

3.2 Methods

Population Sampling

For both target species, we sampled populations from the two contrasting habitats (white sands and dark soils) at the same localities for both *A. inornata* and *S. cowlesi* to enable demographic inference across the same spatial scale (Figure 1). Populations are not polymorphic for color, therefore all individuals from the White Sands sites exhibited the blanched phenotype, and all individuals from the dark soil sites exhibited the ancestral dark phenotype. We sampled White Sands individuals (referred to as "WS" throughout) from White Sands National Monument (WSNM, Otero County, New Mexico), within an approximately 2km radius along two neighboring trails (Alkali Flat Trail and Backcountry Trail). We sampled dark soil individuals from two localities: 1) a Bureau of Land Management site northeast of white sands (BLM, Otero County, New Mexico; referred to as "DS1" throughout), located ~85 km from WSNM with no separating geographic barrier in between, and 2) from Jornada Long-term Ecological Research Station southwest of White Sands (JLTER, Dona Ana County, New Mexico; referred to as "DS2" throughout), located ~50km from White Sands but separated from WSNM by the San Andres Mountains. The two focal species are patchily distributed in the Tularosa Basin, and the two dark soil sites represent the geographically closest localities where dark color morphs of both species could reliably be sampled. We sampled approximately ten individuals per species per population.

Sequence Capture Assay

For each target lizard species, we first generated a fosmid library from high-quality DNA from a single White Sands individual. We used the diTag fosmid vector pFosDTx. The average *E. coli* insert size was approximately 40kb and each fosmid library contained ~5 million individual constructs. The fosmids were then colony-amplified, and we used homologous recombination to screen the fosmid libraries for the *Mc1r* gene. We isolated clones that spanned a region up to 100kb around the *Mc1r* gene and characterized these clones for insert size and target sequence presence. We also chose 96 random clones in each species. Fosmid library preparation and screening was done at the BACPAC Resources Center (Oakland, CA).

We used both Roche 454 FLX+ pyrosequencing and Illumina MiSeq 2x150bp sequencing to sequence the four *Mc1r* clones and the 96 random clones in each species. We processed raw sequences by removing any potential PCR duplicates, non-target species sequence (e.g., from sequence adapters, PhiX spike-in, *E. coli* and fosmid vector), and by trimming low quality ends. We merged Illumina paired-end reads with FLASH (Magoc & Salzberg 2011) and performed a hybrid (454 + Illumina) assembly using Roche gsAssembler 2.6 (Roche 454, 2011). We retained contigs greater than 2kb in length for further analysis resulting in 364 contigs for *A. inornata* (ranging from 2,011bp to 67,990bp in length) and 290 contigs for *S. cowlesi*

(ranging from 2,309bp to 54,334bp in length). We combined the contigs into reference sequences, which we used to design capture probes for a custom Roche NimbleGen Capture Assay. The resulting probes covered more than 96% of our reference sequence contigs.

We extracted DNA from liver and tail samples using a DNeasy Blood and Tissue Kit (Qiagen). We generated Illumina TruSeq barcoded libraries for all individuals from each species, performed capture protocols according to the Roche NimbleGen specifications, and sequenced using the Illumina MiSeq 2x150bp and Illumina HiSeq 2x100bp platforms.

Sequence Alignment

Raw sequence reads (fastq files) were first preprocessed using a custom in-house pipeline including removal of PCR duplicates, contaminants (phiX), and Illumina sequencing adapters, and trimming of low quality bases (seqyclean parameter –qual 24 24) [GRC_Scripts, http://github.com/ibest/GRC_Scripts; Seqyclean, https://bitbucket.org/izhbannikov/seqyclean]. Preprocessed reads were aligned to the reference assembly for each lane separately using Stampy (version 1.0.22) (Lunter & Goodson 2011). For every individual, aligned reads were merged across different lanes, proper pairs were extracted, and duplicate reads were removed using SAMtools (Li *et al.* 2009), retaining only the read pair with the highest mapping quality. After mapping, the mean coverage across individuals was 112X in *A. inornata* (*n* = 32; Table S1a) and 115X in *S. cowlesi* (*n* = 28; Table S1b). The data set was limited to individuals with >20X coverage; this level of coverage has previously been shown to give good resolution for genotyping heterozygous sites within individual samples (The 1000 Genomes Project Consortium 2010). The resulting data set included 24 individuals for each species (i.e. *A. inornata*: 9 WS; 9 DS1; 6 DS2 and *S. cowlesi*: 9 WS; 9 DS1; 6 DS2). Because single nucleotide variants occur much more frequently than indels in the genomes of most species, most alignment algorithms annotate a single nucleotide variant rather than an indel at positions mismatching the reference genome when mapping individual reads. At positions of true indels, local misalignment of reads can produce an excess of false-positive SNP calls. To identify these positions and to improve variant calls, especially in low-complexity regions of the genome, a multiple sequence alignment was performed using the Genome Analysis Toolkit (GATK) IndelRealigner (McKenna *et al.* 2010; DePristo *et al.* 2011; Van der Auwera *et al.* 2013) to locally realign reads such that the number of mismatching bases is minimized across all reads spanning this locus.

Variant Calling and Filtering

Initial variant calls were made using GATK's HaplotypeCaller via local *de novo* assembly of haplotypes in an active region. Samples were genotyped jointly using GATK's GenotypeGVCFs tool. Besides true variation, these initial variant calls contain false positives due to systematic sequencing artifacts, mismapped reads, and misaligned indels. Such false-positive calls often (i) exhibit excessive depth of read coverage, (ii) show an allelic imbalance, (iii) occur preferentially on a single strand, (iv) appear in regions of poor read alignment, and (v) arise in unusual close proximity to multiple other variants. Thus, the majority of such calls can be detected and rejected using filters based on the above observations. Specifically, initial variant calls were filtered post-genotyping using GATK's VariantFiltration. Variants were removed using the following set of criteria (with acronyms as defined by the GATK package): (i) Three or more variants were found within 10bp (clusterWindowSize=10). (ii) The depth of coverage at the given position (summed across individuals) was <500 or >3000 (DP<500; DP>3000). (iii) There was evidence of a strand bias as esti-

mated by Fisher's exact test (FS>60.0) or the Symmetric Odds Ratio test (SOR>1.0). (iv) The read mapping quality was low (MQ<80). (v) At least one of the samples was not called (NCC>0).

After applying the initial filter criteria, the variant data set was limited to biallelic sites using VCFtools (Danecek *et al.* 2011). Genomic positions that fell within repeat regions of the reference assembly were excluded because erroneous alignment of reads to these regions often leads to an increased frequency of heterozygous genotype calls. In particular, five different classes of repeats (i.e., LINE, LTR, DNA, simple repeats, and low complexity regions) were annotated using RepeatMasker (Smit *et al.* 2013-2015) and variants within these regions were excluded from further analyses. To minimize genotyping errors, all variants with either missing data for any individual or genotype quality of less than 20 (corresponding to P[error] = 0.01) for any individual were excluded using VCFtools. Although hard genotype quality thresholds might cause an undercalling of heterozygotes in samples with low or moderate coverage, they have previously been shown to perform well in samples with >20X coverage (Nielsen *et al.* 2011). Variants were also filtered on the basis of Hardy Weinberg Equilibrium (HWE). A p-value for HWE was calculated for each variant using VCFtools, and variants with $p < 0.01$ were removed. Sites for which all individuals were fixed for the non-reference allele were excluded using VCFtools. The resulting call sets contained 13,960 variants (407 SNPs within *Mc1r* and 13,553 within the random contigs) for *A. inornata* and 20,782 variants (691 SNPs within *Mc1r* and 20,091 within the random contigs) for *S. cowlesi*. Genotypes were phased using BEAGLE (version 4; Browning & Browning 2007).

Candidate variants were subject to several filter criteria in order to avoid false positives. As the applied filter metrics can lead to the exclusion of a substantial fraction of sites in the genome, mask files defining which nucleotides were accessible to variant discovery were generated in order to obtain the exact number of monomorphic sites in the reference assembly of each species (used in the demographic estimation) and to avoid biases when calculating summary statistics (e.g., π, Tajima's D, and weighted F_{st}). Mask files were created using the GATK pipeline described in the section 'Variant Calling and Filtering' with the exception that the '-allSites' flag was switched on when running GATK's GenotypeGVCFs tool to include all non-variant loci for which there was data available. The same filter criteria were used with the exception of the variant cluster filter criteria, as this metric cannot be applied in cases of calls at every site in the genome. The filtering resulted in 58% and 45% of the reference assembly being accessible for *A. inornata* and *S. cowlesi*, respectively.

PCA, Population Structure, and Heterozygosity

For both species, variant data sets for the three populations were pruned for linkage, removing SNPs within a 50 SNP window that had $r^2 > 0.2$ (using '-indep 50 5 0.2' in plink) because both PCA (Zheng *et al.* 2012) and the inference of population structure require a set of independent SNPs. Population structure was studied using *structure* (Pritchard *et al.* 2000; Falush et al. 2003; Evanno *et al.* 2005; Falush *et al.* 2007; Hubisz *et al.* 2009), a software that identifies clusters of related individuals from multi-locus genotyping data. *structure* analysis was performed with K= 1-5 (the number of clusters), using an admixture model with correlated allele frequencies. For each K, *structure* was run ten times for 10,000 steps after a burn-in period of 10,000 steps. The best K was chosen such that it maximized the marginal likelihood of the data. Previous work has shown that *structure* is able to identify isolated and relatively homogeneous groups even if they exhibit short divergence times or exchanges with other groups; this is because small, isolated populations often exhibit distinctive allele frequencies due to the fact that genetic drift occurs rapidly (Rosenberg *et al.* 2001). As a result, identified clusters often correspond well to geographically distinct population groups. Finally, heterozygosity was estimated based on the number of heterozygous SNPs per individual.

Demographic Analyses

Population modeling was done using two likelihood methods for comparison: 1) an approach that infers demographic parameters from the joint site frequency spectrum (SFS) using coalescent simulation (fastsimcoal2; Excoffier & Foll 2011; Excoffier *et al.* 2013), and 2) a method in which the likelihood is calculated using a diffusion approximation (δaδi; Gutenkunst *et al.* 2009). SFS for all populations were directly generated from the final variant calls (vcf files) using an in-house script. Since outgroup sequences were unavailable for *A. inornata* and *S. cowlesi,* we used the distribution of minor allele frequencies (i.e., the folded SFS) where the minor allele was considered to be the allele with the lowest frequency across all three populations. Monomorphic sites (defined as the set of all sites that passed all filtering criteria but for which no variants were called using GATK) were included in the analyses. For both species, we maximized the likelihood of the observed SFS under six complex demographic scenarios (parameters shown in Table 2) and identified the best fitting demographic model on the basis of their Akaike's information criterion (AIC) score (Akaike 1974).

Given the genetic differentiation between populations identified by the PCA and *structure* analyses, for each species we tested the three possible tree topologies for the population set including WS, DS1, and DS2. Topologies were tested both without migration and with asymmetrical migration between all population pairs. In the migration models, gene flow was only considered between *T*1 (the time of the most recent population split) and present. Additionally, admixture models were evaluated but in no case better fit the data. Effective population sizes were directly estimated by fixing the mutation rate to 1.5×10^{-9} (Olave *et al.* 2014).

Fastsimcoal2 and δaδi were used on the folded SFS based on 13,553 SNPs and 20,091 SNPs outside *Mc1r* in *A. inornata* and *S. cowlesi*, respectively. For fastsimcoal2 the following options were used: -N 100000 (max. number of simulations) -L 40 (max. number of EM cycles) -M 0.001 (min. relative difference in parameter values for the stopping criterion). For every demographic model, 20 independent estimations with different initial parameter values were run and results for the estimation with the highest likelihood are reported. For δaδi, 50 independent runs at different starting points were executed for each model. Generations were converted to years assuming a generation time of 1.5 years, a reasonable average estimate for this taxonomic group (e.g., Crenshaw 1955).

Selective Sweep Mapping in the *Mc1r* Region

To test for a selective sweep around the *Mc1r* gene, we used the modification of the Kim & Stephan (2002) composite likelihood ratio (CLR) test proposed by Nielsen *et al.* (2005) as implemented in the software *SweeD* (Pavlidis *et al.* 2013). For each species, the folded SFS for all contigs except *Mc1r* was used as a neutral background reference and the CLR statistic was calculated at 10,000 grid points across the contig containing *Mc1r*. The CLR test assumes that the data was sampled at the end of the selective sweep, thus, following Meiklejohn *et al.* (2004), we excluded all WS individuals that were homozygous or heterozygous for the dark allele (as identified by Rosenblum *et al.* (2010)) from the analysis. The statistical thresholds for the test were defined as recommended by Nielsen *et al.* (2005). For each species, we simulated polymorphism data under our best demographic models (as estimated by *fastsimcoal2*) for the WS population and defined the threshold as the 95th percentile of the distribution of highest simulated CLR values.

Additionally, classical summaries of genetic diversity were calculated using a sliding window approach along the *Mc1r* region. We calculated π (nucleotide diversity), Tajima's *D* (Tajima 1989), and Weir and Cockerham's F_{st} (1984) using VCFtools (Danecek *et al.* 2011) taking into account the information about monomorphic sites (using the '--mask' option).

ABC Estimation of Selection Coefficients and Allele Ages

To estimate selection coefficients and allele ages for the putatively selected mutations in *A. inornata* and *S. cowlesi*, we use the approximate Bayesian (ABC) approach of Ormond *et al.* (appearing in this issue). In summary, an approximate Bayesian computation (ABC) approach was applied using a standard rejection algorithm and tolerance threshold of 0.01, implemented in the R program *abc* (Csillery *et al.* 2012), to estimate posterior distributions for selection strength *s* and allele age *T* in both WS populations. 100,000 neutral and non-neutral simulations of the genealogies for White Sands populations in both species were generated using *msms* (Ewing & Hermisson 2010). Simulations used the demographically inferred N_e values, as well as the ρ and μ values used in demographic inference. The lengths *L* of the simulated sequences were taken to match the SNP data available around the region of the putatively selected mutation (a 65kb region for *A. inornata* and a 54kb region for *S. cowlesi*). The positions of the simulated selected mutations in both populations were chosen to match the position of the putative selection targets in the data sets. The prior distributions for the selection coefficient *s* and allele age *T* were $log_{10}(s) \sim U(-4, -0.5)$ and $log_{10}(T) \sim U(-4, -0.5)$, where U is a uniform distribution.

Simulated SNP patterns were output to the program *msstats* to calculate a panel of known summary statistics that are commonly used to characterize selective sweeps. Following Wegmann *et al.* (2009), a partial least squares (PLS) approach was applied to incorporate the most informative summary statistics from *msstats* into the ABC calculation and to filter out noise from uninformative statistics. Summary statistics were also calculated from the actual SNP data for White Sands populations in both species and transformed into PLS components using the same loadings as described above. Point estimates for *s* and *T* were calculated from the mode of the joint density posterior distribution using the two-dimensional kernel density function in the MASS package in R (Venables & Ripley 2002). Finally, credibility intervals were calculated using 95% of the marginal posterior distributions for *s* and *T.*

3.3 Results

Sequence Capture

Our sequence capture approach resulted in a high quality population-level data set. We recovered a large contig containing *Mc1r* for both species. The *Mc1r* contig was 54kb long and contained 691 SNPs in *S. cowlesi* and was 68kb long and contained 407 SNPs in *A. inornata*. We also recovered 289 additional genomic contigs with an average length of 11.3kb in *S. cowlesi* and 363 additional genomic contigs with an average length of 8.6kb in *A. inornata*. These genomic contigs contained 20,091 SNPs in *S. cowlesi* and 13,553 SNPs in *A. inornata.*

Demographic Analysis of the *A. inornata* Populations

All three *A. inornata* populations exhibited similar levels of nucleotide diversity (0.15% - 0.18%, Table 1) and Tajima's *D* values (0.11 – 0.80) (calculated based on the genetic variation observed outside *Mc1r*). Weighted F_{st} values ranging from 0.11 (WS-DS2) to 0.15 (DS1-DS2) outside of *Mc1r* indicated genetic structure between populations. Genetic structure was also identified by the *structure* analysis - supporting the existence of three clusters, which exactly corresponded to the three sampled localities (Figure 2, Figure S1a). The PCA analysis also showed clear separation of all three populations (Figure 3). Notably, weighted F_{st} values were greatly elevated within *Mc1r* in comparisons between White Sands and dark soil populations (0.28 (WS-DS1), 0.36 (WS-DS2)), but not between the two dark soil populations (0.12 (DS1-DS2)).

The demographic model inferred as best fitting the data was the same in both fastsimcoal2 and δaδi (Figure 4), with concordant parameter estimates between the two methods (Table 2). This model suggests a young divergence between WS and DS1 and an older split with DS2, as well as an absence of gene flow between populations. Predictive simulations were generated to test whether the demographic parameters inferred with the two methods were able to correctly predict the patterns of genetic variation observed in the *A. inornata* data set. The results of this analysis demonstrated that both models were well calibrated, predicting π, Tajima's *D*, and F_{st} correctly in all three populations (Figure S2) and the SFS of each individual population (Figure S3).

Demographic Analysis of the *S. cowlesi* Populations

All three *S. cowlesi* populations exhibited similar levels of nucleotide diversity (0.25% - 0.27%, Table 1), with roughly 1.5-fold higher genetic variation than observed in *A. inornata*. Tajima's *D* values ranged from -0.17 to 0.25. The range of pairwise weighted F_{st} values was similar to that found in *A. inornata* (from 0.10 in WS-DS2 to 0.15 in WS-DS1). The *structure* results reported two or three clusters depending on which criterion was used to identify the best value of *K* (Figure 2 and Figure S1B, respectively), highlighting the lower level of differentiation between WS and DS2. However, the three-cluster grouping correctly assigned all individuals to their respective sampling localities. The same pattern of differentiation was observed in the PCA analysis (Figure 3) where WS and DS2 could only be differentiated on the second principal component. As in *A. inornata*, weighted F_{st} was notably higher within the *Mc1r* region relative to the genomic background in comparisons between White Sands and dark soil populations (0.31 (WS-DS1), 0.20 (WS-DS2)), but not between the two dark soil populations (0.16 (DS1-DS2)).

Both fastsimcoal2 and δaδi analyses predicted the same general demographic model as best fitting the data (Figure 4), a model with a closer relationship between WS and DS2 and presence of gene flow between the three populations. Although both methods agreed on the tree topology and on the presence of gene flow, there are, unlike in the case of *A. inornata*, notable differences between the parameters estimated by the two methods (Table 2). To test whether these different parameter estimates affected the predictive power of the model, we conducted predictive simulations using both the fastsimcoal2 and *δaδi* estimates and compared their ability to predict the observed data (Figure S4 and S5). The results demonstrated that both methods were indeed equally well-calibrated as they correctly predicted the distribution of π, Tajima's *D*, and F_{st} in all three populations (Figure S4). Additionally, both approaches predicted similar single population SFS that matched the observed data well (Figure S5). It is important to additionally note that when scaling to calendar years the uncertainty in mutation rate, recombination rate, and generation time in these non-model organisms has important implications (see Supp Table 2 for an example, considering two possible mutation rates and three possible generation times, for both species).

Selection Analysis at the *Mc1r* Locus in *A. inornata*

Several aspects of the genetic variation observed at the *Mc1r* locus in *A. inornata* stand in sharp contrast with the diversity observed in the species' genomic background and are consistent with recent and strong positive selection. First, as noted above, the level of genetic differentiation among populations at *Mc1r* was very high compared to the differentiation observed at neutral fragments (Figure 5). Second, the nucleotide diversity at the *Mc1r* locus in the WS population was below the genome-wide average over a region of about 50kb encompassing the *Mc1r* gene (Figure 6). Furthermore, only two out of nine individuals contributed to most of the variation in this region (CP4 and CP34). If these two individuals are discarded, the seven remaining individuals are almost entirely monomorphic over ~35kb region (Figure S6). This reduction of diversity at *Mc1r* is not observed in the dark soil populations (Figure 6) and is consistent with the expected local reduction of neutral diversity predicted by the selective sweep model (Maynard-Smith and Haig 1974). Note that the estimated and observed neutral diversity indices are very similar between all three populations (Table 1), so we do not expect higher levels of genetic drift in WS to be responsible for the lower variation at *Mc1r*. Third, as expected in a region that experienced recent strong positive selection, Tajima's *D* values in the WS population were negative and below the genomic background average. This was not observed in the dark soil populations where *D* profiles were above average (DS1) or in-line (DS2) with their respective background distributions (Figure 6). To test whether patterns of variation were consistent with a recent event of strong positive selection, we applied the CLR test (Kim & Stephan 2002; Nielsen *et al.* 2005) (Figure 7). The test was only significant for the WS population and yielded an estimated selection coefficient for the advantageous allele of 9×10^{-4}. This estimation was obtained assuming a recombination rate of 1.5×10^{-9} and an effective population size for WS of 23,392 individuals (Table 2). This newly identified selective sweep co-localized with a non-synonymous mutation responsible for a polarity-changing replacement in *Mc1r* ($THR^{170}ILE$; *Mc1r* amino acid residue 170 and contig position 27,132) that has been previously associated with blanched coloration in *A. inornata* (Rosenblum *et al.* 2010). Consistent with the hypothesis that this mutation plays a major role in the local adaptation of the WS population, we found the frequency of the white allele (T) to be 0.83 (i.e., 15 out of 18 gene copies) in WS and absent in DS1 and DS2. This represented the largest difference in allele frequencies between populations in the *Mc1r* region for the WS-DS1 comparison (F_{st} = 0.82) and the second largest for WS-DS2 (F_{st} = 0.78) (Figure S7). The three copies of the dark allele (C) were found in CP34 (homozygote) and CP4 (heterozygote), the two individuals contributing to most of the residual variation in the sweep region. Interestingly, these two individuals appear to share the same blanched phenotype as the other samples from the WS population, suggesting the existence of additional loci contributing to the genetic basis of this adaptive trait.

Selection Analysis at the *Mc1r* Locus in *S. cowlesi*

In *S. cowlesi,* the *Mc1r* region does not contain a signal for selection as strong as that in *A. inornata,* but several aspects of diversity still suggest a strong selective sweep of the putatively beneficial mutation identified by Rosenblum *et al.* (2010). Inspection of polymorphism patterns revealed the presence of strong haplotypic structure in the *Mc1r* alignment (Figure S7). The four WS individuals that were homozygous for the white allele exhibited dramatically reduced genetic variation across the entire contig (55kb) compared to individuals homozygous for the dark allele. Heterozygotes were found in WS only, which stands in agreement with the expected dominance of the white allele. F_{st} values only marginally reflected this strong haplotypic differentiation as the white haplotype was not fixed in WS (Figure S7). The results of the CLR test on this white haplotype subset of WS yielded significant values over the whole sequenced region (Figure 7).

The selection coefficient estimated by the CLR method was 1.1 x 10^{-3}, assuming a recombination rate of 1.5 x 10^{-9} and an effective population size of 20,237 (Table 2).

The putatively beneficial mutation leads to an amino acid replacement ($HIS^{208}TYR$) that induces the blanched phenotype in individuals that are homozygotes or heterozygotes for the white allele (Rosenblum *et al*. 2010). This dominant allele has been shown to have a frequency of 0.46 in the White Sands population but to be absent in the dark soil populations. In agreement with the study of Rosenblum *et al.* (2010), this mutation occurred with a frequency of 0.61 in our WS sample (four individuals were homozygotes and three heterozygous) but was absent from both DS1 and DS2. Interestingly, two WS individuals who were homozygous for the dark allele had a light phenotype, suggesting other loci besides *Mc1r* contribute to the blanched phenotype in *S. cowlesi* as well.

ABC Estimation of Selection Coefficients and Allele Ages

Using an ABC-based approach, we estimated selection coefficients for the putative target of 0.19 based on the inferred fastsimcoal2 model and 0.15 based on the inferred δaδi model in *A. inornata*. In *S. cowlesi*, we estimated selection coefficients for the putative target of 0.12 based on the fastsimcoal2 model and 0.05 based on the δaδi model (Figure 8). Interestingly, the age of the beneficial mutation was inferred to be young in both species, with estimates ranging from 0.0002 to 0.003 4*N* generations in *A. inornata* (mean estimate of 900 years), and from 0.01 to 0.001 4*N* generations in *S. cowlesi* (mean estimate of 1200 years). However, the posteriors suggest that it may be difficult to distinguish between the ages of these two relatively young sweeps – rather, it is simply possible to distinguish between young and old. Given the estimates of effective population sizes however, both sweep patterns are inferred to be significantly younger than the geological age of the White Sands.

3.4 Discussion

The White Sands lizards of southern New Mexico offer an outstanding system to understand the demographic and adaptive history associated with colonization of young and novel habitats. Given that multiple species have independently and convergently adapted to the light colored sand, and that the genetic basis of adaptation is known in some species, this system offers insight into the topology of the adaptive landscape and the interplay between selection and demography during colonization.

Demography

Our findings using hundreds of anonymous loci to estimate the neutral demographic history of the two species pairs investigated here provide insight into both shared and unique aspects of colonization of a novel habitat. In both *A. inornata* and *S. cowlesi*, all populations sampled were distinguishable from each other, indicating that there is genetic structure over this habitat gradient at a fine spatial scale. In both species, the White Sands populations also showed a comparable amount of genetic diversity as dark soil populations. There is evidence in both species for population size reductions relative to the ancestral population, but this effect was no stronger in White Sands populations than in dark soil populations. Thus, we did not find evidence for dramatic genetic bottlenecks at White Sands, consistent with earlier work in this system that drew from more limited genetic data (Rosenblum et al. 2007).

We inferred recent divergence times between White Sands and dark soil populations in both species, consistent with the known geological history of the formation. However, the inferred split time of light and dark populations of *A. inornata* (in which the most recent split was between WS and DS1 ~4,500 years ago) was considerably younger than in *S. cowlesi* (in which the most recent split was between WS and DS2 ~7,400 years ago). The difference in divergence time estimates between the two species might reflect a more recent colonization of the dunes by *A. inornata* or a limited sampling of dark soil populations for this study. Finally, patterns of gene flow differed between the two focal species. In the species with the more recent split, *A. inornata*, no migration was detected, while in *S. cowlesi* there was strong evidence for ongoing migration.

It is additionally of note that while the two commonly used demographic estimators utilized here agree on important features of tree topology and the presence/absence of migration, there are notable differences particularly in *S. cowlesi* with regards to specific parameter estimates. Perhaps the most important of which is the age of the WS population divergence, with *δaδi* inferring a more recent split (post-dune formation) from DS2, and fastsimcoal2 inferring a more ancient split (prior to dune formation). By conducting simulations of the best estimated models of the two approaches, we have demonstrated that both well-explain the observed data – highlighting the fact that there are areas of the demographic parameter space that are equally able to predict our observations. Future work will focus on a more comprehensive sampling of dark soil populations to better understand the ancestry of White Sands populations and refine divergence time estimates.

Selection

Apart from the inherent interest in characterizing the demographic history of colonization in the White Sands lizards, our demographic estimates also provide an important null model for tests of positive selection. We evaluated patterns of molecular evolution around the *Mc1r* gene, a region hypothesized to be under selection, relative to the genomic background. Based on F_{st} analyses, it was clear that *Mc1r* is much more strongly differentiated between light and dark individuals than was expected based on differentiation across the rest of the genome. In addition to F_{st} analyses, we also took a CLR approach to evaluate the likelihood of selection across the *Mc1r* region. In the two dark soil populations sampled for each species, no significant test value was found. However, in light populations of both *A. inornata* and *S. cowlesi*, strongly significant rejections of neutrality were identified. Interestingly, in both cases, the likelihood surface peaks were centered around two previously described functional variants proposed to play a role in the light phenotype (Rosenblum *et al.* 2010).

Given the differing colonization histories of these species, it was next of interest to infer the age of the putatively beneficial mutations highlighted by the CLR approach. Taking a newly proposed approximate Bayesian methodology (presented in this issue, see Ormond *et al.*), we inferred the strength of selection and the age of the light alleles in both species taking in to account both the inferred effective population size and the dominance of the underlying mutations. The strength of selection was estimated to be roughly equally strong in both species, and the ages of both beneficial mutations were estimated to be considerably younger than the age of the White Sands formation itself. This result is highly consistent with the demographic results in *A. inornata*, in which the WS has a recent population split time, consistent with the beneficial allele age. While there is a discrepancy in estimated population split time in *S. cowlesi* between our two inference methods as discussed above, the estimated age of the beneficial allele is indeed young – suggesting that the adaptive event in both species occurred considerably after the formation of the dunes.

Establishing the frequency with which selection acts on rare vs. common variants has become an important focal point in evolutionary genetics. Our results suggest adaptation on *de novo* or rare mutations in both White Sands species. This result is of particular interest given prior theoretical work suggesting that adaptation from standing genetic variation is more likely at White Sands, particularly for the recessive case (Nuismer et al. 2012). However, our approach, which leveraged hundreds of SNPs over more than 50kb in the *Mc1r* gene region, provides a more robust analysis of strength of selection and allele age than has been previously possible. Our inference of selection on rare mutations is based on: 1) the strong CLR signal in both light populations, a test which only has power to detect hard selective sweeps, 2) the inferred allele ages being considerably younger than the age of the White Sand formation (i.e., the timing of the shift in selection pressure), and 3) the light allele not being observed as segregating in the dark populations. As it is likely that the light phenotype would indeed be deleterious on the dark soil, it is likely segregating at mutation-selection balance in the dark populations. Given the necessary frequency at which selection on a standing variant results in a soft sweep (i.e., a multiple haplotype fixation) versus a hard sweep (i.e., a single haplotype fixation) as described by Orr & Betancourt (2001) and Jensen (2014), a soft sweep model is thus unlikely in these populations. However, we have reason to believe that there are as-of-yet unknown variants contributing to the phenotype; therefore, a model of polygenic adaptation can not be ruled out.

Given the inference of selection around previously identified and functionally validated variants, along with the in-depth demographic estimates, this system provides another unique perspective – namely, the role of dominance. The beneficial light allele appears to be dominant in *S. cowlesi* but recessive in *A. inornata*. Though the mean fixation time of both a recessive and dominant favored allele is similar for large *N* and *Ns* (van Herwaarden & van der Wal 2002), the impact of a selective sweep in both cases is expected to differ. When selection acts on a recessive beneficial mutation, genetic drift dominates the early phase of the allele trajectory, as the mutation remains invisible to selection until it reaches a frequency via genetic drift at which it may appear in a homozygous recessive state. However, once achieving that frequency, a deterministic trajectory may be entered bringing the allele to fixation. Conversely, for selection acting on a dominant beneficial mutation, the variant may be visible to selection immediately and begin sweeping, however genetic drift will dominate the later phases of the trajectory as the wild type recessive allele will be invisible to selection when in the heterozygous state. As described by Teshima & Przeworski (2006), these dynamics result in a stronger reduction in diversity near the selected site for recessive mutations, but a wider reduction for dominant mutations. Figure 9 presents simulation results for these two models, for the species-specific parameters estimated here. Though this analysis only represents a single example, our results are indeed qualitatively consistent with theoretical predictions - particularly when considering the width of the significant likelihood surface.

Conclusion

Our study provides a number of important insights into the process of parallel ecological adaptation in a novel and geologically young environment. While the times of colonization may differ between the focal species, both appear to have colonized the White Sands area after their geological formation, utilizing different genetic mechanisms within *Mc1r*. Importantly, while both functional and population genetic evidence strongly support the role of these two identified variants in shaping the light phenotype, our results also suggest that other as-of-yet unknown mutations likely play a role in the light adaptation. Thus, future research will seek to identify these additional mutational targets, and will expand to other light/dark species pairs across this ecotone in order to more fully characterize the generality of these conclusions.

3.5 Tables

Table 1

		A. inornata	*S. cowlesi*
	Length (bp)	1,776,757	1,429,855
WS	n	9	9
	S	9,983	12,308
	π	0.0018	0.0025
	D	0.31	-0.01
DS1	n	9	9
	S	7,750	11,693
	π	0.0015	0.0025
	D	0.80	0.25
DS2	n	6	6
	S	9,497	11,946
	π	0.0018	0.0027
	D	0.11	-0.17
F_{ST}	WS-DS1	0.13	0.15
	WS-DS2	0.11	0.10
	DS1-DS2	0.15	0.13

Table 1: Summary statistics of the genetic variation observed in the set of genome-wide random contigs. All statistics were calculated directly from the joint site frequency spectra (sites that did not pass quality control were masked out).

Table 2

	A. inornata		*S. cowlesi*	
	fsc2	δaδi	fsc2	δaδi
N_ANC	327,022	324,114	394,383	420,678
N_WS_DS1	46,478	123,264	--	--
N_WS	49,586	23,392	173,803	20,237
N_DS1	16,104	7,443	58,323	15,232
N_DS2	83,463	87,418	256,011	32,068
T1	9,627	4,538	435,660	7,401
T2	15,462	18,168	453,086	254,652
T_SIZE_DS1	--	--	10,656	7,401
NE_BOT_DS1	--	--	308,492	1,076,768
NE_WS_DS2	--	--	66,690	1,515,572
M_WS_DS1	--		4.01E-06	2.95E-05
M_WS_DS2	--	--	2.73E-06	3.02E-06
M_DS1_WS	--	--	9.30E-07	1.75E-05
M_DS1_DS2	--	--	2.59E-06	2.75E-05
M_DS2_WS	--	--	5.74E-06	3.09E-06
M_DS2_DS1	--	--	2.99E-06	2.75E-05

Table 2: Parameter estimates inferred by fastsimcoal2 and δaδi under the best demographic models for *A. inornata* and *S. cowlesi*. Times are given in years and sizes in number of individuals. A graphical representation of these two demographic models (and parameter definitions) can be found in Figure 4.

3.6 Figures

Figure 1: Map of study area

Photographs and sampling localities for *Aspidoscelis inornata* and *Sceloporus cowlesi* from contrasting habitats. Blanched colour morphs are found at White Sands (indicated by the white bar) and dark colour morphs are found in the rest of the species' ranges (indicated by the black bar). Both species were sampled from the same three localities in New Mexico: White Sands National Monument (WS) in Otero County (blue), a dark soil Bureau of Land Management site (DS1) in Lincoln County (red) and a dark soil Jornada Long-term Ecological Research site (DS2) in Doña Ana County (green).

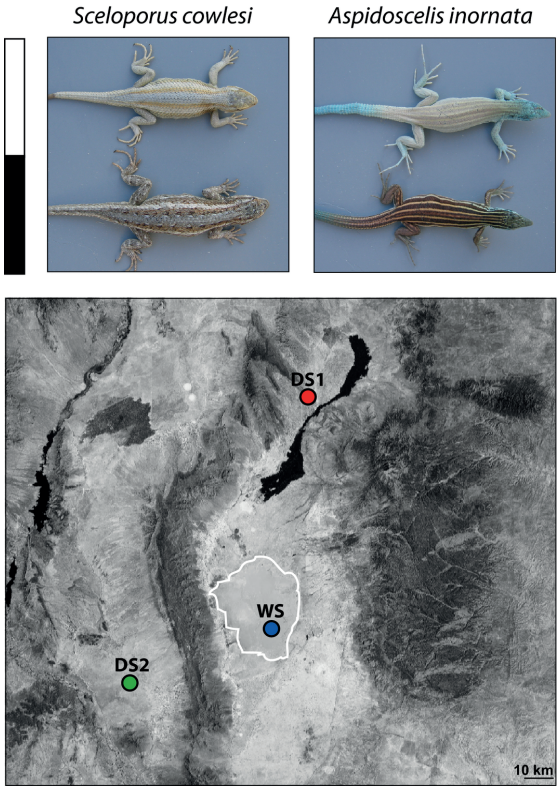


Figure 2: *structure* results for the three lizard populations. Estimated population structure as inferred by *structure* in *A. inornata* (top) and *S. cowlesi* (bottom). Each individual is represented by a bar partitioned into *K* colored segments; the color of each bar's label indicates the source population, blue for WS, red for DS1, and green for DS2. These segments represent the estimated membership fractions of the individual in *K* clusters. At each *K*, ten *structure* runs were performed which generated nearly identical individual membership coefficients. Figures shown for a given *K* are based on the highest probability run.

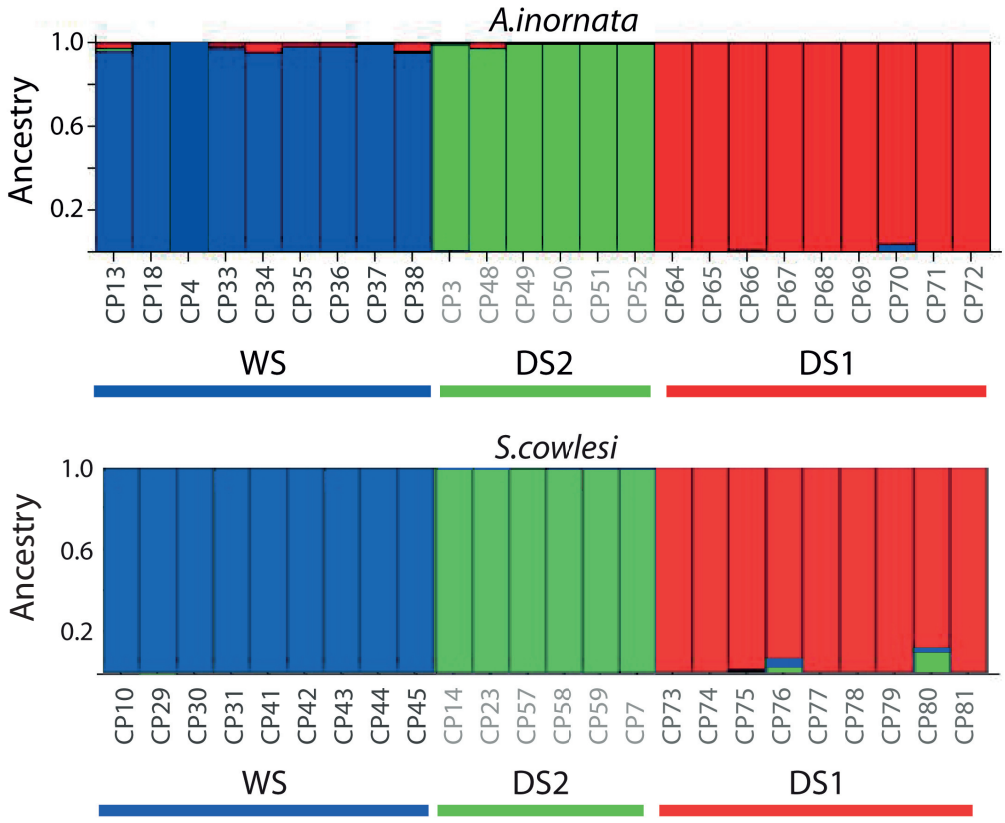


Figure 3: PCA of the three lizard populations (corresponding to the colors in Figure 2; WS (blue), DS1 (red), and DS2 (green)) for *A. inornata* (left) and *S. cowlesi* (right). Data was thinned to exclude SNPs with an r^2 > 0.2 in order to avoid a strong influence of SNP clusters in the PCA. Percentages indicate the percent of variance explain by each principle component.

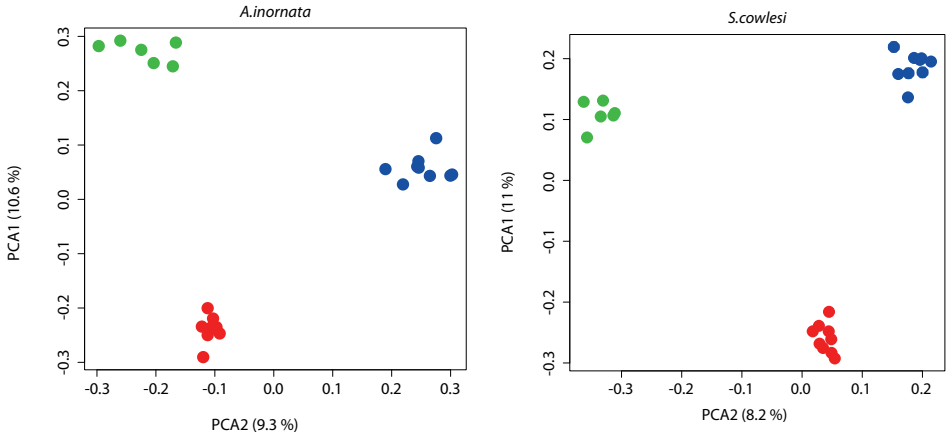


Figure 4: Best demographic model for *A. inornata* and *S. cowlesi* as inferred by both *δaδi* and fastsimcoal2, with parameters defined.

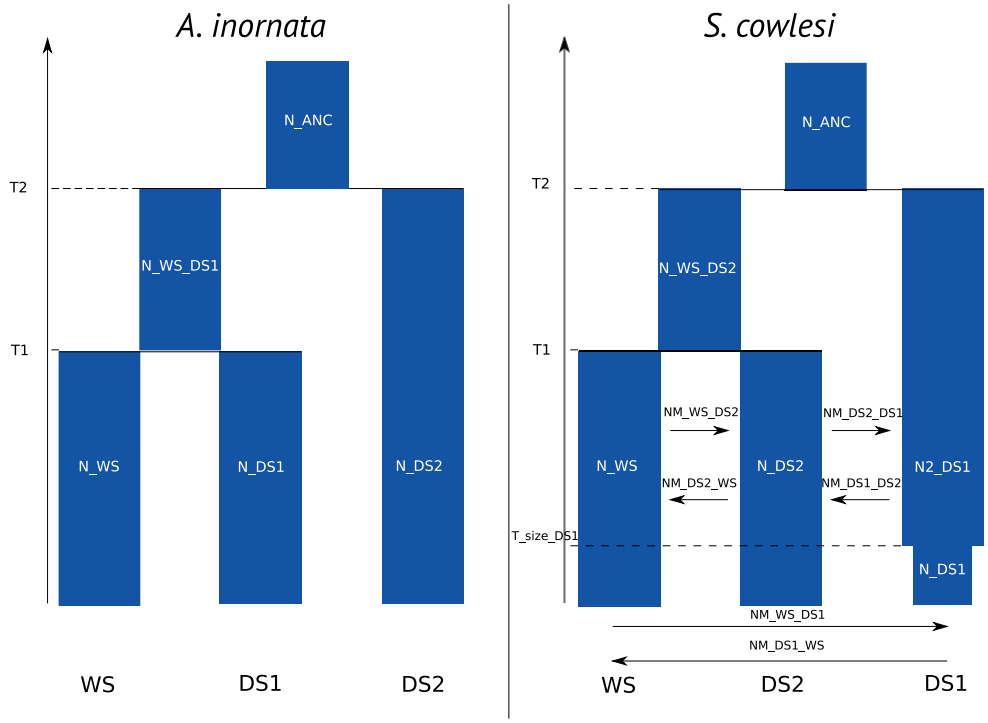


Figure 5: Genetic differentiation at the *Mc1r* locus in *A. inornata*. Sliding window profile of Weir and Cockerham's (1984) estimator of F_{st} for all pairs of populations as calculated by VCFtools. Window size was 1,000bp and step size 250bp. Sites that did not pass quality control were masked (using the '--mask' option). The solid horizontal lines represent the average weighted F_{st} across all windows in the genomic background. The dashed horizontal line represents the .975th quantile of the same distribution. The red vertical solid line indicates the position of the non-synonymous putatively beneficial *Mc1r* mutation reported by Rosenblum *et al.* (2010).

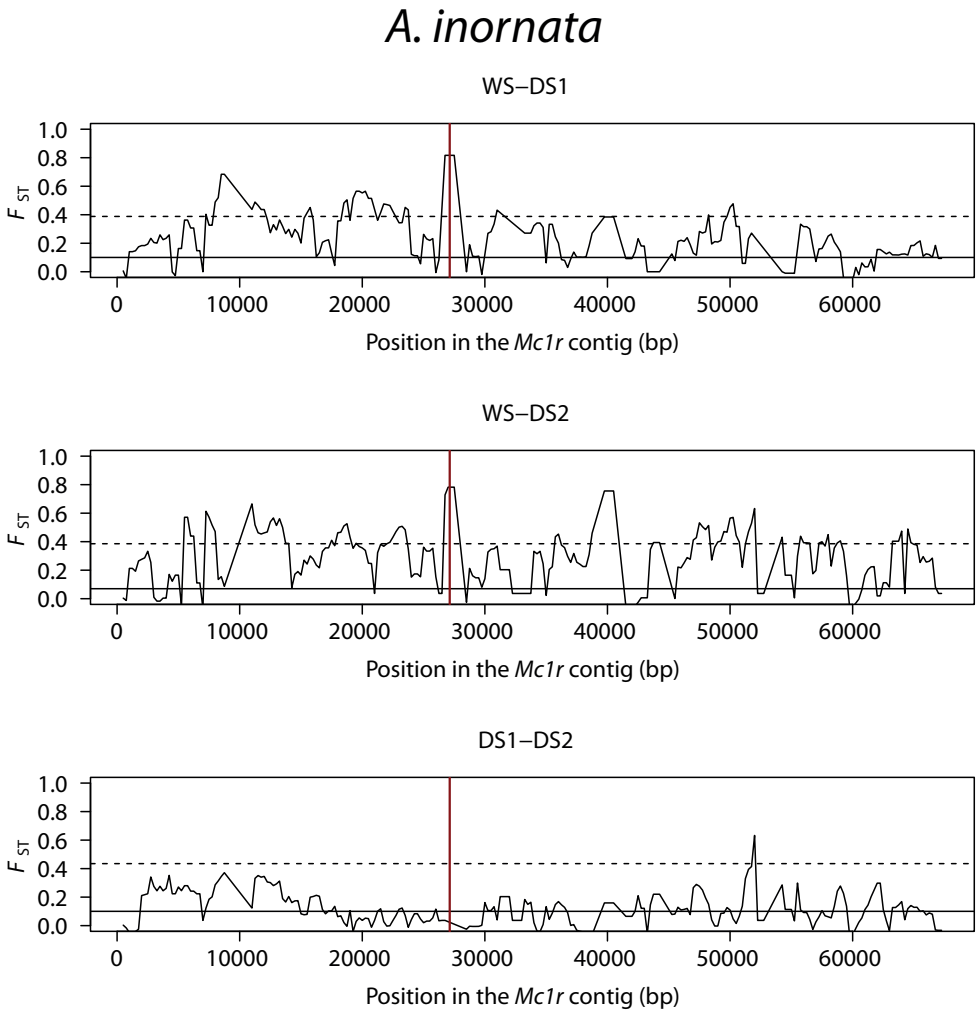


Figure 6: Nucleotide diversity (π) and Tajima's *D* at the *Mc1r* locus in *A. inornata* populations. Sliding window profile of nucleotide diversity (π) and Tajima's *D* in the *Mc1r* region as calculated by VCFtools (window size 1,000bp and step size 250bp). Sites that did not pass quality control were masked (using the '--mask' option). The solid horizontal lines represent the average values of these statistic calculated across all windows in the genomic background. The dashed horizontal lines represent the 0.025th and 0.975th quantile of the same distribution. The red vertical solid line indicates the position of the candidate non-synonymous *Mc1r* mutation reported by Rosenblum *et al.* (2010).

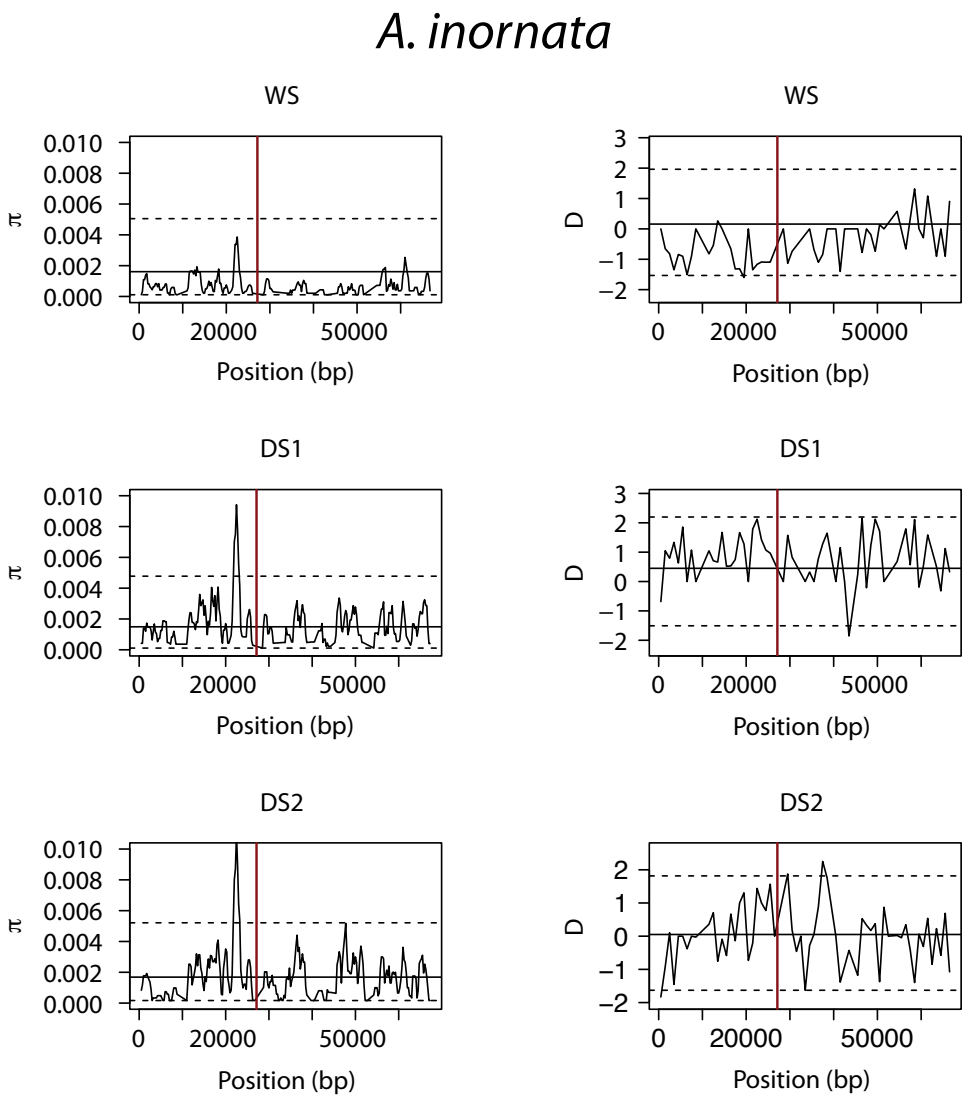


Figure 7: Likelihood surfaces of the CLR test calculated by *SweeD* for *A. inornata* and *S. cowlesi*. The dashed horizontal line is the significance threshold of the test for WS (see materials and methods). The red vertical solid line indicates the position of the non-synonymous *Mc1r* mutations reported by Rosenblum *et al.* (2010). Individuals that were homozygous for the dark allele of this mutation were excluded from the analysis (see main text).

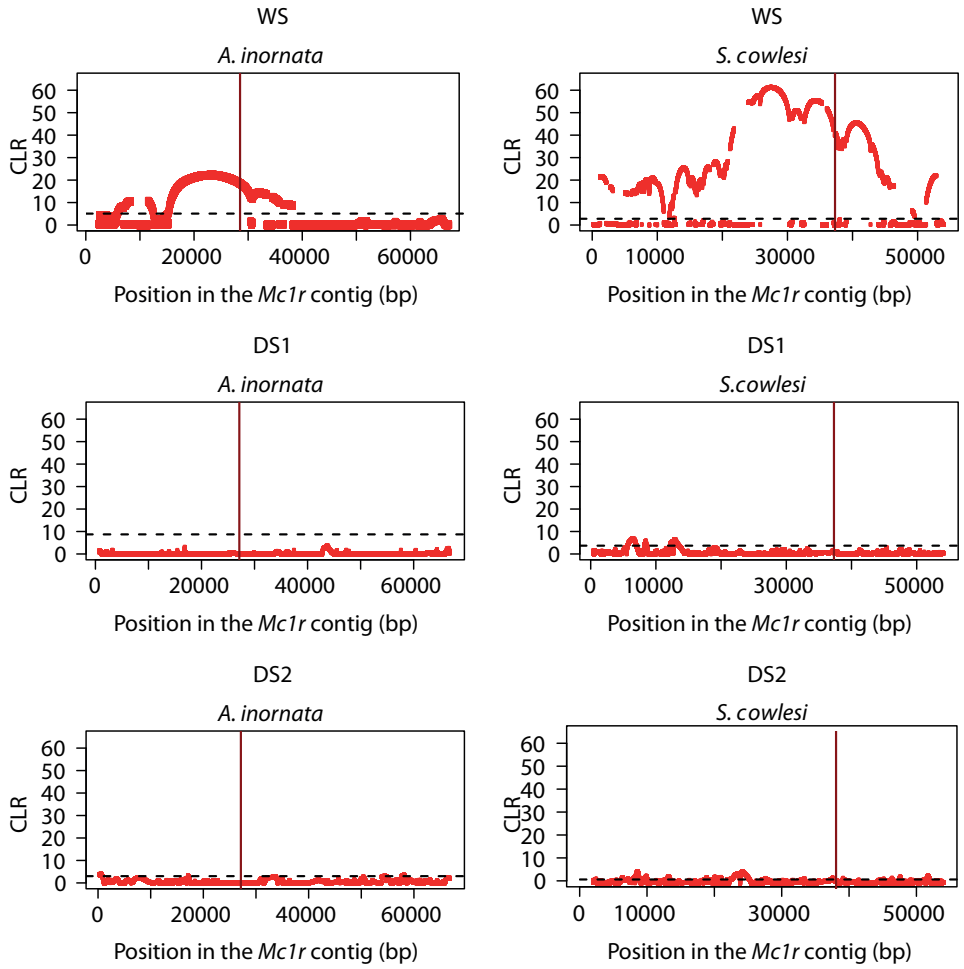
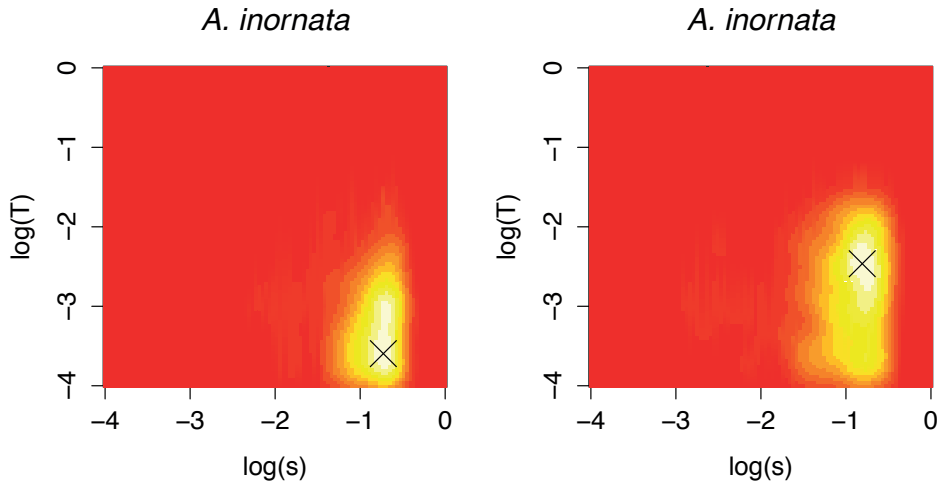


Figure 8: Estimation of selection coefficient *s* and allele ages *T.* Figures show the joint posterior density plots for *s* and *T.* Given somewhat different demographic histories estimated between the fastsimcoal2 (Panel A & C) and *δaδi* (panel B & D) software, both results are shown in each species. The white, yellow and red colors indicate areas of high, moderate, and low joint density, respectively. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Black crosses indicate the modes of the joint posterior distributions, which are at *s*=0.19 (CI: 0.006 – 0.3), *T*=0.00025 x $4N_e$ generations (CI: 0.0001 – 0.002) based on fastsimcoal2 estimates and *s*=0.16 (CI: 0.0009 – 0.3), *T*=0.0034 x $4N_e$ generations (CI: 0.0001 – 0.002) based on *δaδi* estimates for *A. inornata;* and *s*=0.13 (CI: 0.0002 – 0.2), *T*=0.014 x $4N_e$ generations (CI: 0.0001 – 0.003) based on fastsimcoal2 estimates and *s*=0.05 (CI: 0.0003 – 0.2), *T*=0.001 x $4N_e$ generations (CI: 0.0001 – 0.004) based on *δaδi* estimates for *S. cowlesi.*

A) B)



C) D)

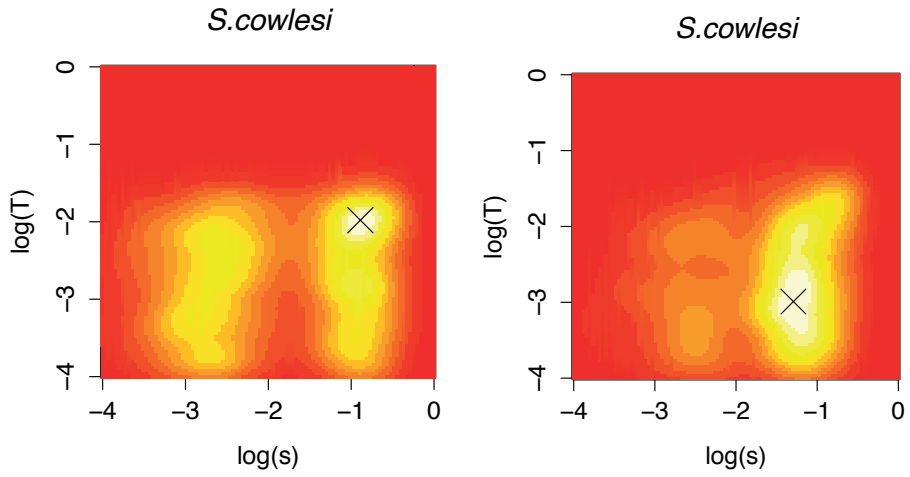
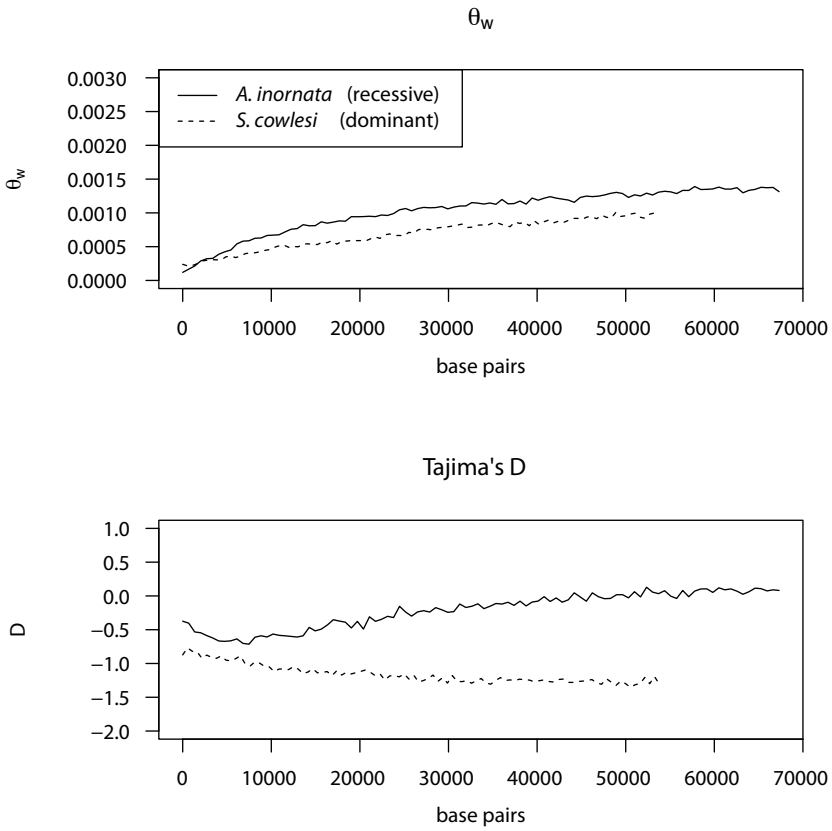


Figure 9: Predictive simulations under the sweep models inferred in this study by the ABC method of Ormond *et al.* (this issue). The missing values for *S. cowlesi* in the right part of the graph are due to different contig length for the *Mc1r* region between the two species. Simulations were performed using the msms program (Ewing & Hermisson 2010)



References

Akaike H (1974) A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, **19**, 716-723.

Barsh GS (1996) The genetics of pigmentation: from fancy genes to complex traits. *Trends in Genetics*, **12**, 299-305.

Bittner TD, King RB, Kerfin JM, Gatten RE (2002) Effects of body size and melanism on the thermal biology of garter snakes. *Copeia*, **2002**, 477-482.

Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, **81**, 1084-1097.

Caro TIM (2005) The adaptive significance of coloration in mammals. *BioScience*, **55**, 125-136.

Chan YF *et al*. (2010) Adaptive evolution of pelvic reduction in stincklebacks by recurrent deletion of a *Pitx1* enhancer. *Science*, **327**, 302-305.

Cott, HB (1940) Adaptive coloration in animals. Methuen, London.

Crenshaw JW (1955) The life history of the southern spiny lizard, *Sceloporus undulates undulates* Latreille. *The American Midland Naturalist*, **54**, 257-298.

Crisci JL, Poh Y, Mahajan S, Jensen JD (2013) The impact of equilibrium assumptions on tests of selection. *Frontiers in Genetics*, **4.**

Csillery K, Francois O, Blum MGB (2012) abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology & Evolution*, **3**, 475-479.

Danecek P *et al.* (2011) The Variant Call Format and VCFtools. *Bioinformatics* **27**, 15, 2156-2158.

DePristo M *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, **43**, 491-498.

Dice LR (1947) Effectiveness of selection by owls of deer mice (*Peromyscus maniculatus*) which contrast in color with their background. Contributions from the Laboratory of Vertebrate Biology of the University of Michigan 34:1–20.

Doebley J (2004) The genetics of maize evolution. *Annual Reviews of Genetics*, **38**, 37-59.

Domingues VS, Poh Y-P, Peterson BK, Pennings PS, Jensen JD, Hoekstra HE (2012). Evidence of adaptation from ancestral variation in young popultions of beach mice. *Evolution*, **66**, 3209-3223.

Eizirik E *et al.* (2003). Molecular genetics and evolution of melanism in the cat family. *Current Biology*, **13**, 448-453.

Ellegren H, Sheldon, BC (2008). Genetic basis of fitness differences in natural populations. *Nature*, **452**, 7184:169-75

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 261126-20.

Ewing G, Hermisson J (2010) MSMS: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics*, **26**, 2064-2065.

Excoffier L, Foll M (2011) fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, **27**, 1332-1334.

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLoS Genetics*, **9**, e1003905.

Falush D, Stephens M, and Pritchard JK (2003) Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Falush D, Stephens M, and Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular *Ecology Notes*, **7**, 574– 578.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Iinferring the joint demographic history of multiple pouplations from multidimensional SNP frequency data. *PLoS Genetics*, **5**, e1000695.

Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, **169**, 2335-2352.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science*, **313**, 101-104.
Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources,* **9**, 1322-1332.
Ihle S, Ravaoarimanana I, Thomas M, Tautz D (2006) An analysis of signatures of selective sweeps in natural pouplations of the house mouse. *Molecular Biology & Evolution*, **23**, 790-797
Jensen JD (2014). On the unfounded enthusiasm for soft selective sweeps. *Nature Communications*, **5**, 5281.
Jensen JD, Kim Y, DuMont VB, Aquadro CF, Bustamante CD (2005) Distinguishing between selective sweeps and demography using DNA polymorphism data. *Genetics*, **170**, 1401-1410.
Jensen JD, Wong A, Aquadro CF (2007) Approaches for identifying targets of positive selection. *Trends in Genetics*, **23**, 568-577.
Johnson JB, Omland KS (2004) Model selection in ecology and evolution. *Trends in Ecology & Evolution*, **19**, 101-108.
Kaufman DW (1974) Adaptive coloration in *Peromyscus polionotus*: Experimental selection by owls. *Journal of Mammalogy*, **55**, 271-283.
Kettlewell B (1973) The Evolution of melanism: the study of a recurring necessity; with special reference to industrial melanism in the Lepidoptera. Clarendon Press, Oxford.
Kijas JMH *et al.* (1998) Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics,* **150**, 1177-1185.
Kim Y, Stephan W (2002) Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*, **160**, 765-777.
Kocurek G *et al.* (2007) White Sands Dune Field, New Mexico: age, dune dynamics and recent accumulations. *Sedimentary Geology*, **197**, 313-331.
Li H *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* **25**, 2078.
Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE (2009) On the origin and spread of an adaptive allele in deer mice. *Science*, **325**, 1095-1098.
Linnen CR, Poh Y-P, Peterson BK, Barrett RDH, Larson JG, Jensen JD, Hoekstra HE (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, **339**, 6125
Lowe CH, Norris KS (1956) A subspecies of the lizard *Sceloporus undulates* from the White Sands of New Mexico. *Herpetologica*, **12**, 125-127.
Luke CA (1989) Color as a phenotypically plastic character in the side-blotched lizard, *Uta stansburiana* Ph.D. dissertation. University of California, Berkeley.
Lunter G, Goodson M (2011) Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research,* **21**, 936.
Mackay TFC, Stone EA, Aryoles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, **10**, 565-577.
Magoc T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genomic assemblies. *Bioinformatics*, **27**, 2957-2963.
Majerus M (1998) Melanism: evolution in action. Oxford University Press, Oxford.
Marklund L, Moller MJ, Sandberg K, Andersson L (1996) A missense mutation in the gene for melanocyte-stimulating hormone receptor (MCIR) is associated with the chestnut coat color in horses. *Mammalian Genome*, **7**, 895-899.
Mathew LA, Jensen JD (2015) Evaluating the ability of the pairwise joint site frequency spectrum to co-estimate selection and demography. *Frontiers Genetics,* in press.
McKenna A *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297.
Meiklejohn CD, Kim Y, Hartl DL, Parsch J (2004) Identification of a locus under complex positive selection in *Drosophila simulans* by haplotype mapping and composite-likelihood estimation. *Genetics*, **168**, 265-279.
Mundy NI, Badcock NS, Hart T, Scribner K, Janssen K, Nadeau NJ (2004) Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science*, **303**, 1870-1873.

Nachman MW, Hoekstra HE, D'Agostino, SL (2003) The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences*, **100**, 5268-5273.
Naduvilezhath L, Rose LE, Metzler D (2011) Jaatha: a fast composite-likelihood approach to estimate demographic parameters. *Molecular Ecology*, **20**, 2709-2723.
Newton JM *et al.* (2000) Melanocortin 1 receptor variation in the domestic dog. *Mammalian Genome,* **11**, 24-30.
Nielsen R et al. (2005) Genomic scans for selective sweeps using SNP data. *Genome Research*, **15**, 1566-1575.
Nielsen R *et al*. (2011) Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet* **12**, 443-451.
Norris KS (1967) Color adaptation in desert reptiles and its thermal relationships. *Lizard Ecology: A Symposium*, 162-229, University of Missouri Press, Columbia.
Norris KS, Lowe CH. (1964) An analysis of background color-matching in amphibians and reptiles. *Ecology*, **45**, 565-580.
Nuismer SL, MacPherson A, Rosenblum EB. (2012) Crossing the threshold: gene flow, dominance, and the critical level of standing genetic variation required for adaptation to novel environments. *Journal of Evolutionary Biology,* **25**, 2665-2671.
Olave M, Avila LJ, Sites JW, Morando M (2015) Model-based approach to test hard polytomies in the *Eulaemus* clade of the most diverse South American lizard genus *Liolaemus* (Liolaemini, Squamata). *Zoological Journal of the Linnean Society*, **174**, 169-184.
Orr HA, Betancourt AJ (2001) Haldane's sieve and adaptation from the standing genetic variation. *Genetics*, **157**, 875-884.
Pavlidis P, Jensen JD, Stephan W, Stamatakis A (2012) A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. *Molecular Biology & Evolution*, **29**, 3237-3248.
Pavlidis P, Zivkovic D, Stamatakis A, Alachiotis N (2013) SweeD: likelihood-based detection of selective sweeps in thousands of genomes. *Molecular Biology & Evolution*, **30**, 2224-2234.
Poh Y-P, Domingues VS, Hoekstra HE, Jensen JD (2014) On the prospect of identifying adaptive loci in recently bottlenecked populations. *PLoS One*, **9**, e110579
Pollinger JP, Bustamante CD, Fledel-Alon A, Schmutz S, Gray MM, Wayne RK (2005) Selective sweep mapping of genes with large phenotypic effects. *Genome Research*, **15**, 1809-1819.
Pritchard JK, Stephens M, and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155,** 945–959.
Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics*, **160**, 1179-1189.
Reusch TBH, Haeberli MA, Aeschlimann PB, Milinski M (2001) Female sticklebacks count alleles in a strategy of sexual selection explaining MCH polymorphism. *Nature*, **415**, 300-302.
Rosenberg NA *et al*. (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics*, **159**, 699-713
Rosenblum EB (2006) Convergent evolution and divergent selection: lizards at the White Sands ecotone. *The American Naturalist*, **167**, 1-15
Rosenblum EB, Harmon LJ (2011) "Same same but different": replicated ecological speciation at White Sands. *Evolution*, **65**, 946-960.
Rosenblum EB, Hickerson M, Moritz C (2007) A multilocus perspective on colonization accompanied by selection and gene flow. *Evolution,* **61**, 2971-2985.
Rosenblum EB, Hoekstra HE, Nachman MW (2004) Adaptive reptile color variation and the evolution of the *Mc1r* gene. *Evolution*, **58**, 1794-1808.
Rosenblum EB, Rompler H, Schoneberg T, Hoekstra HE (2010) Molecular and functional basis of phenotypic convergence at white lizards at White Sands. *Proceedings of the National Academy of Sciences USA*, **107**, 2113-2117.
Smit AFA *et al.* (2013-2015) RepeatMasker Open-4.0 <http://www.repeatmasker.org>.
Smith HM (1943) The White Sands earless lizard. *Zoological Series of the Field Museum of Natural History*, Chicago.

Stapley J *et al*. (2010) Adaptation genomics: the next generation. *Trends in Ecology and Evolution*, **25**, 705-712.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158-170.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

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

Teshima KM and Przeworski M (2006) Directional positive selection on an allele of arbitrary dominance. *Genetics* 172; 713-8.

Thayer GH (1909) Concealing—Coloration in the Animal Kingdom: An Exposition of the Laws of Disguise through Color and Pattern. Macmillan, New York.

The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061-73.

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

Thronton K, Andolfatto P (2006) Approximate Bayesian inference reveals evidence for a recent,s evere bottleneck in a Netherlands population of *Drosophila Melanogaster*. *Genetics*, **172**, 1607-1619.

Thronton KR, Jensen JD (2007) Controlling the false-positive rate in multilocus genome scans for selection. *Genetics*, **175**, 737-750.

Van der Auwera GA *et al.* (2013) From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols* **43**, 11.10.1-11.10.33.

Van Herwaarden, OA, and van der Wal NJ (2002) Extinction time and age of an allele in a large finite population. Theor Pop Biol 61: 311-318.

Venables WN, Ripley BD (2002) Modern applied statistics with S. Springer, New York.

Wegmann D, Leuenberger C, Excoffier L (2009) Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, **182**, 1207-1218**.**

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Zheng X, *et al*. (2012) A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data, *Bioinformatics*, **28**, 3326-8.

Chapter 4 On the Combined Effect of Oseltamivir and Favipiravir in Treating Influenza Virus

Ormond L, Liu P, Matuszewski S, Renzette N, Bank C, Zeldovich K, Bolon DN, Kowalik TK, Finberg RW, Jensen JD, and Wang JP. 2017. The combined effect of oseltamivir and favipiravir on influenza A virus evolution. Manuscript in preparation.

4.1 Introduction

Influenza A virus (IAV) inflicts a heavy disease burden worldwide, including 36,000 deaths in the United States annually. Developing effective drugs to combat this burden is a public health priority. The most frequently used drug, oseltamivir, was designed as a competitive inhibitor of the viral surface neuraminidase (NA) glycoprotein responsible for binding host cell sialic acid to enable the release of virus progeny (Moscona 2005). Oseltamivir binding requires altering a hydrophobic pocket in the NA region and can be destabilised by a single mutation near the active site, of which several are possible (Varghese *et al.* 1998; Collins *et al.* 2008). Early studies *in vitro* and *in vivo* identified high fitness costs associated with these mutations and lent support to the view that the development of resistance was unlikely in clinical settings (Ives *et al.* 2002). The most common resistance mutation in H1N1 strains, H275Y, was observed infrequently during clinical testing (Gubareva *et al.* 2001) but spread rapidly worldwide during the 2007/2008 influenza season (Moscona 2009) and remains a clinical concern (Ghedin *et al.* 2012; Meijer *et al.* 2014; Takashita *et al.* 2015). H275Y confers resistance to oseltamivir but lowers virus fitness by reducing the amount of neuraminidase that reaches the cell surface (Bloom *et al.* 2010). The higher than expected fitness of mutants carrying H275Y is likely due to the presence of compensatory mutations that increase cell surface expression and enzymatic activity of neuraminidase (Bloom *et al.* 2010; Bouvier *et al.* 2012; Ginting *et al.* 2012; Butler *et al.* 2014)

The rapid spread of oseltamivir resistance has increased interest in developing drugs with an alternative mechanism of action and lower susceptibility to resistance. Favipiravir is a mutagenic drug, which inhibits viral RNA-dependent RNA polymerase (RdRp) and dramatically increases the IAV mutation rate (Fig. 1), potentially pushing the virus towards extinction (Baranovich *et al.* 2013; Furuta *et al.* 2013). Favipiravir is effective against a range of RNA viruses including influenza and is currently in phase 3 clinical trials. It is safe for use in humans as human cells do not have RdRp domains, and it has a distinct mode of action from both oseltamivir and M2 drugs. Most importantly, to date no resistance

mutations have been identified, either because the process may be biologically complex or because viral extinction occurs too rapidly for resistance to evolve.

Previous work in the Jensen laboratory has analysed the evolution of IAV experimental populations treated with oseltamivir only (Foll *et al.* 2014) or with favipiravir only (Bank *et al.* 2016). Foll *et al.* (2014) developed a time-sampled approach (WFABC) to identify targets of selection in oseltamivir treated IAV populations, and to distinguish these from genetic drift (Fig. 2). Results confirmed that resistance to oseltamivir could be readily achieved in a single mutational step and re-identified the NA H275Y mutation. A number of other previously described (*e.g.*, HA D112N mutations) and novel (*e.g.*, MP1 E23Q and A41V) mutations were found to have roles in improving cell infectivity and virion budding, and with potential epistatic interactions with H275Y. E23Q was highlighted as a possible compensatory mutation, off-setting the cost of adaptation of H275Y reflected in the reduced amount of neuroaminidase to reach the cell surface. In addition, Foll *et al.* (2014) analysed the distribution of fitness effects in the absence and presence of oseltamivir. A heavy tail of beneficial mutations in the presence of the drug reflects the presence of resistance and compensatory mutations, underpinning the rapid adaptive potential of IAV under treatment with oseltamivir (Fig. 3).

Bank *et al.* (2016) describe the accumulation of deleterious mutations in experimental IAV populations treated with escalating doses of favipiravir over 15 passages. Beyond a critical threshold, the viral population proceeds into a phase of rapid decline and escalating mutation load that ultimately leads to extinction (Fig. 4). This outcome depends on a high dosage of favipiravir. In two replicates where the drug dosage was held constant or withdrawn after passage 10, a reduction in the negative growth rate (relative to the no drug control) was observed, providing the first evidence of viral adaptation to low concentrations of favipiravir (Fig. 5).

Using drugs in combination is an established clinical strategy aimed at preventing or delaying resistance by rapidly depleting pathogen populations before resistance can emerge (*e.g.*, against tuberculosis (Mitchison 2012)). Thus, there is interest in combining oseltamivir and favipiravir in the treatment of influenza. Synergistic benefits have been obtained using favipiravir and oseltamivir in combination against influenza A H1N1pdm in *in vivo* studies in mice, as well as against an H275Y oseltamivir resistant model of infection in mice (Smee *et al.* 2013). The two drugs have very different mechanisms of action and the effects of combining these are unknown.

Over the past few decades, a number of studies have shown the impact of mutagenic drugs on RNA virus extinction (*e.g*. foot-and-mouth disease virus (FMDV) (Sierra *et al.* 2000; Pariente *et al.* 2001) HIV-1 (Loeb *et al.* 1999; Loeb & Mullins 2000) and lymphocytic choriomeningitis (LCMV) (Grande-Pérez *et al.* 2002)). Different hypotheses have been evoked to explain these effects. RNA viruses, including the H1N1 virus studied here, are frequently cited as being subject to an error threshold, as they are characterized by high mutation rates, short replication times and large populations sizes. The error catastrophe model (Eigen 1971, 2002; Holmes 2003) that this set of characteristics invokes, postulating a distinct mutational threshold above which replication is no longer possible and extinction occurs, implies an upper limit on genome size and mutation rate. In keeping with this expectation, most RNA viruses are characterized by very short genomes (3000-30,000bps, Domingo *et al.* 2001), and evidence is indeed accumulating that mutation rates in these viruses may rest very near the error threshold. As a means of explaining the extinction process upon crossing this limit, models of lethal mutagenesis have been proposed (Bull *et al.* 2007; Wylie & Shakhnovich 2012), in which this

extinction is described as a deterministic (and thus population size-independent) process of accumulation of deleterious mutations which erode viral fitness until a point of population collapse is reached. As a separate literature, the population genetics community has developed models of mutational meltdown (Lynch & Gabriel 1990), rather focusing on finite population processes governing extinction, including Muller's ratchet (a stochastic process (Muller 1964; Felsenstein 1974)) and Hill-Robertson interference (a deterministic process, (Hill & Robertson 1966; McVean & Charlesworth 2000)).

Muller's ratchet describes the stepwise loss of the fittest class of individuals (the "least loaded" class) in a non-recombining population due to genetic drift, and thus the reduction in absolute fitness due to the accumulation of deleterious alleles, or drift load (Whitlock & Bourguet 2000). In a finite population, Muller's ratchet eventually results in the extinction of the population, if it is not opposed by compensatory or beneficial mutations. The ratchet operates at a speed that depends on the effective population size, the selection coefficient of deleterious mutations and the size of the least loaded class at mutation-selection equilibrium and accelerates exponentially with an increasing mutation rate (Haigh 1978; Gordo & Charlesworth 2000a, b). Hill-Robertson interference (Hill & Robertson 1966) describes the reduction in the efficacy of selection owing to linkage between selected alleles and their genetic backgrounds. In an effect known as weak-selection Hill-Robertson interference (WSHRI) (McVean & Charlesworth 2000), the spread of weakly beneficial mutations may be slowed by linkage with weakly deleterious mutations. In another form of interference, beneficial mutations arising on different backgrounds will compete for fixation. In small, non-recombining experimental viral populations subject to an increasing mutation rate, both Muller's ratchet and Hill-Robertson interference are likely to be significant, in addition to the deterministic erosion of viral fitness owing to the accumulation of deleterious mutations.

To evaluate these processes and assess the clinical advantage of a combined protocol, we treated IAV populations with escalating doses of oseltamivir and favipiravir ("combined drug" populations) over 10 passages, and compared these with three control-paired replicate IAV populations treated with oseltamivir only, as well as with results from the earlier favipiravir and no drug treated populations. This experimental set-up offers an excellent platform to dissect the complex dynamics contributing to meltdown, and in particular the roles of genetic hitchhiking, Muller's ratchet and Hill-Robertson interference. The oseltamivir-only replicates and favipiravir population provide a way of separating out these components and of identifying the dynamics that are particular to the combination. The availability of multiple replicates affords our study additional power.

Our results demonstrate that the combined drug populations become extinct or nearly extinct in ten passages, but with some important differences in the underlying dynamics compared to the favipiravir-alone population. Viral growth in oseltamivir-treated replicates remained high and stable throughout all passages, due to the emergence of resistance mutations. In the combined drug replicates, we observe an accumulation of deleterious and neutral mutations that lead to viral extinction. Intriguingly, extinction proceeded at a faster rate than for the favipiravir population, despite a lower number of segregating mutations. We find evidence that deleterious mutations hitchhike to fixation with strongly selected oseltamivir resistance mutations, a process that reduces viral fitness and accelerates extinction. The selective sweeps depress genetic variation and explain the lower number of segregating mutations. Additionally, evidence of a small and rapidly declining effective population size support a role for Muller's ratchet. Finally, Hill-Robertson interference patterns specific to the combined drug (clonal competition between beneficial mutations) and favipiravir populations (WSHRI) emerge.

There is no evidence that the oseltamivir resistance mutation NA H275Y arises earlier in combined drug replicates than in oseltamivir replicates. In two out of three combined drug replicates, H275Y fixed during the same passage as in the corresponding oseltamivir control. Interestingly, two other putative NA resistance mutations were identified that fixed rapidly in combined drug replicates 2 and 3. Our results suggest that the mutational effect of favipiravir combined with oseltamivir rapidly explores sequence space, generating other possible NA resistance mutations and clustering beneficial mutations on the same haplotype, although at the cost of a high linked deleterious mutation load that ultimately drives viral extinction. Thus, we find that the combined drug protocol potentially drives an earlier extinction point for viral populations than the favipiravir protocol alone, but at the risk of spreading both established and new oseltamivir resistance mutations.

4.2 Background : Two sides of the same coin : a population genetics perspective on lethal mutagenesis and mutational meltdown

Matuszewski S, **Ormond L**, Bank C, & Jensen JD. *Virus Evolution*, accepted.

Fisher (1930) argued that an intermediate mutation rate is likely optimal for populations to survive, as it ensures a constant input of beneficial mutations providing the 'fuel' for adaptation, while limiting the impact of accumulating deleterious mutations. Subsequently, the effects of high population mutation rates and the related risk of population extinction have been explored heavily in the theory literature, spanning the fields of population genetics as well as virology.

This literature has spawned a number of concepts describing the extinction of populations owing to the excessive fixation of deleterious mutations. From population genetics, the mutational meltdown model (Lynch & Gabriel 1990) has been proposed, invoking previously developed evolutionary processes including Muller's ratchet (Muller 1964; Felsenstein 1974) and Hill-Robertson interference (Hill & Robertson 1966; McVean 2000). Relatedly, from the study of viral evolution, the models of error catastrophe (Eigen 1971; Eigen 2002; Holmes 2003) and lethal mutagenesis (Bull *et al*. 2007; Wylie & Shakhnovich 2012) have emerged.

Despite the different fields in which these ideas were developed, there is a considerable amount of parallelism between these notions. For example, mutational meltdown is generally discussed within the context of 'small' population sizes in which stochastic effects play an important role, whereas lethal mutagenesis is generally discussed within the context of 'large' population sizes driven by deterministic factors, though recent extensions of the latter model have relaxed this assumption to incorporate stochastic effects (Wylie & Shakhnovich 2012). Hence, the extent to which these models are overlapping descriptions of related processes or events, or are even subsets of one another, is unclear. This has resulted in an inherent confusion when invoking these models to describe biological observations.

We here seek firstly to clearly define and review these models in terms of parameter requirements, predicted effects, and biological relevance, and then compare them side-by-side in the light of their similarities and differences. Fundamentally, we propose that the notion of lethal mutagenesis

largely describes the deterministic subset of the dynamics of mutational meltdown, and thus call for a more integrated focus on the underlying processes driving extinction.

Error Catastrophe

The concept of "error catastrophe" was originally developed in the context of a general theory of molecular evolution pertaining to the origin of life (Eigen 1971). This theory sought to answer the question of how primitive organisms can maintain sufficiently accurate genetic information during reproduction. Subsequently, Eigen & Schuster (1979) developed quasispecies theory to depict the dynamics of nucleic acid molecules under selection and mutation. A key feature of quasi-species theory is the idea of an error threshold, which sets a hard limit on the maximum mutation rate and thus ensures sufficient accuracy of replication.

A system of differential equations can be used to model dynamical evolutionary processes, such as the growth of a bacterial population under different "types" of mutation and selection or the replication of RNA viruses (Eigen 1971; Eigen & Schuster 1979). The analysis assumes an asexual, infinite population. Assuming a probability W_{ii} of error free reproduction of type *i*, and probability W_{ik} that type *i* is formed from type *k*, the description of the change in frequency of type *i* with time *t* is given by:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = (W_{ii} - \bar{E}(t))x_i(t) + \sum_{k\neq i} W_{ik}x_k(t) \qquad (1)$$

(Equation 6 of Eigen and Schuster 1979)
where x_i is the frequency of type *i* and $\bar{E}(t)$ represents the mean fitness of the total population (Biebricher & Eigen 2005). Thus, the change in mutant frequency *i* involves both selection (as represented by the first term of this equation) and mutational gain (as represented by the second term, which takes into account offspring of type *i* produced from parental type *k* through errors in the replication process).

Quasispecies theory (Eigen and Schuster 1979) builds on this analysis in the framework of a sequence space (similar to the notion of protein/sequence space developed by Maynard Smith (1970)). In this context, a quasispecies is defined as "a given distribution of macro-molecular species with closely interrelated sequences, dominated by one or several (degenerate) copies" (Eigen & Schuster 1979). If an equilibrium or steady state is reached (implying $\mathrm{d}x_i/\mathrm{d}t = 0$) and ignoring mutational gain terms above ($\sum_{k\neq i} W_{ik}x_k$), the population proportions of the different "types" making up a quasispecies can be derived as

$$\boldsymbol{x_i = \frac{W_{ii} - \bar{E}_{k\neq i}}{E_i - \bar{E}_{k\neq i}}} \qquad (2)$$

(Equation 10 in Biebricher *et al.* 2005).

In addition to assuming an infinite population (though see Ochoa & Harvey 1998), this analysis effectively assumes a fitness landscape with a single peak (though see Bonhoeffer & Stadler 1993). This means that there is a (single) master sequence *m* with a probability W_{mm} of error free reproduction

that is significantly higher than the probability of reproduction of all other types $W_{kk,}$ (which allows the mutational gain term to be ignored).
This derivation implies the existence of an error threshold. If W_{mm} becomes equal or less than $\bar{E}_{k\neq i}$ the master sequence population would collapse. Given the average fitness of the master sequence $\bar{\sigma}_m$, its length v_m and its fidelity or accuracy of reproduction $Q_{mm}=\bar{q}_m^{v_m}$, the population proportion of the master sequence (from the above equation) can be rewritten following Biebricher & Eigen (2005) as

$$\bar{\boldsymbol{x}}_{\boldsymbol{m}} = (\bar{\boldsymbol{\sigma}}_{\boldsymbol{m}}\boldsymbol{Q}_{\boldsymbol{mm}} - \mathbf{1})/(\bar{\boldsymbol{\sigma}}_{\boldsymbol{m}} - \mathbf{1})\,. \qquad (3)$$

The error threshold $(1-\bar{q}_m)$ can then be derived (Eigen 1971; Eigen & Schuster 1979) as:

$$\mathbf{1} - \boldsymbol{q}_{\boldsymbol{m}} \leq \frac{\ln(\bar{\sigma}_m)}{v_m}. \qquad (4)$$

For the master sequence to be maintained in the population the product of its mean fitness and it accuracy of reproduction needs to exceed unity ($\bar{\sigma}_m Q_{mm}$> 1; allowing the error threshold to be recovered; Eigen 2002). Thus, effectively, any loss in accurate reproduction of the master sequence must be compensated by the relative fitness $\bar{\sigma}_m$ to avoid a collapse. The error threshold depends both on the fitness of the master sequence and on the sequence length. Because the average accuracy of reproduction per locus $\bar{q}_m$ is less than one, any increase in length v_m reduces the fidelity of the master sequence $Q_{mm}=\bar{q}_m^{v_m}$ and reduces the error threshold in the above equation.

Simulations by Swetina & Schuster (1982) and Tarazona (1992) of the stationary state of a population of binary sequences have been used to explore changes in the mutant spectrum as the population mutation rate is increased over the error threshold. Consistent with theory, the spectrum becomes markedly more diverse but remains centered around the master sequence until a complete collapse at the error threshold. This has been described as a sharp "all or none" change with the characteristics of a "first order phase transition" (Eigen 2002).

Lethal mutagenesis

Originally introduced and coined to describe a therapeutic strategy for curing vaccine-lacking viral diseases, the concept of lethal mutagenesis refers to the drug-induced increase in viral mutation rate (achieved through incorporating non-complementary nucleotides or nucleotide analogs during the DNA/RNA replication process), which reduces population mean fitness (and increases mutation load) and in turn leads to a decline in viral population size and eventual extinction (Loeb *et al*. 1999; Anderson *et al*. 2004; Bull *et al.* 2007). Although the concept itself has been applied to both RNA and DNA viruses - including human immunodeficiency virus (HIV), influenza A virus (IAV) and hepatitis B virus (HBV) (see Anderson *et al*. 2004) - results were initially discussed in light of Eigen's error catastrophe model (Eigen 1971) owing to the lack of a separate, formal theoretical framework of lethal mutagenesis.

In an attempt to synthesize existing empirical and theoretical work, Bull *et al*. (2007) proposed a general theory of lethal mutagenesis, which, as the authors noted, "offered [nothing] … specifically original", but is "rather [an] … application of simple models and the interpretation of those results in the context of empirical methods … that made this [theory] original". This model is composed of three basic features: First, a genotype-to-fitness map $\omega(g)$ characterizes how the number of (deleterious)

mutations affects the genotype's fitness; second, a mutation-rate-to-fitness map addresses how the population mean fitness at mutation-selection balance changes with the rate of mutations U (Haldane 1932; Bürger 1998); finally, a demographic model that links the two above components with an ecological component $R_{max}(e)$ that quantifies the (environmentally-dependent) maximal absolute population growth rate (sometimes also called the maximal fecundity). Therefore, the population *absolute* mean fitness $\overline{W}$ is given by

$$\overline{\boldsymbol{W}} = \boldsymbol{R_{max}}(\boldsymbol{e})\overline{\boldsymbol{w}}(\boldsymbol{\omega}(\boldsymbol{g}), \boldsymbol{U}).$$

Depending on the choice of the underlying demographic model, the number of viral particles in the next (discrete) generation N_{t+1} is then simply a function of the current population size N_t and the absolute mean fitness (*e.g.*, a simple exponential growth model of the form $N_{t+1} = N_t\overline{W}$ or a more complex density-dependent model; see Gabriel & Bürger 1992). However, independent of the choice of demographic model, the population will eventually become extinct if the mean absolute fitness $\overline{W} < 1$, such that the population can no longer maintain itself. This can happen either because the mean *relative* fitness $\overline{w}$ drops below a critical value (*i.e.*, mutation load becomes too high), or because $R_{max}(e)$ drops below unity as a consequence of a change in the environment (*e.g.*, due to the application of a novel drug treatment).

In this general formulation, viral populations are assumed to be large initially such that genetic drift is not affecting mutation-selection balance, recombination is thought to be absent, and beneficial mutations are disregarded (see below for a discussion of recent relaxations of these model assumptions). This simplified model allows the calculation of critical mutation rates U_c, beyond which absolute mean fitness drops below unity and the population will become extinct. In particular, under a multiplicative genotype-to-fitness map $\omega(i) = (1-s)^i$, where the fitness of a genotype is reduced by a constant s per deleterious mutation, the mean relative fitness is $\overline{w} = e^{-U}$ (Kimura & Maruyama 1966). Thus, the population would go extinct if $U > U_c = \mathrm{Log}[R_{\max}(e)]$. Notably, due to environmental dependence of $R_{\max}(e)$, there is no universal critical mutation rate across viral populations – not even for a single species.

As shown by the above back-of-the envelope calculation, lethal mutagenesis is independent of population size – a result made explicit in the original model in noting that it is fundamentally a deterministic process that will operate even in very large populations (Bull *et al.* 2007). Importantly, although the outcome of lethal mutagenesis is deterministic, population dynamics, extinction times, and individual trajectories of mean absolute fitness are not. Thus, demography is nonetheless important here (Nowak & May 2000), where finite population sizes will always induce an additional (drift) load that is not accounted for in these models, but which can have a strong effect on population dynamics owing to the fixation of deleterious mutations further decreasing population mean fitness.

Mutational meltdown

Muller's ratchet (Muller 1964; Felsenstein 1974) describes the stepwise loss of the fittest class of individuals in a population and the associated reduction in absolute fitness due to the accumulation of deleterious alleles, or drift load (Whitlock & Bourguet 2000). In a finite population, Muller's ratchet

eventually results in the extinction of the population if it is not opposed by compensatory or beneficial mutations. Lynch and Gabriel (1990) were the first to combine the study of population dynamics (*i.e*., size changes and absolute growth rates) and population genetics (*i.e.,* allele frequency distributions and relative fitnesses) in order to describe this extinction process, which they termed "mutational meltdown". In a series of papers, the authors proposed and analyzed various related models and discussed the properties and implications of mutational meltdown in both asexual and sexual populations (Lynch & Gabriel 1990; Gabriel *et al*. 1993; Lynch *et al.* 1993; Gabriel & Bürger 1994; Lynch *et al.* 1995a,b).

This process is fundamentally described with respect to the accumulation of mutations over time. The mean number of mutations, $\bar{n}(t)$, depends on the carrying capacity *K*, the absolute growth rate *R*, the deleterious effect of a mutation *s*, and the (deleterious) mutation rate μ, and can be expressed as

$$\bar{\boldsymbol{n}}(\boldsymbol{t}+\mathbf{1}) \cong \bar{\boldsymbol{n}}(\boldsymbol{t}) + \mu - \boldsymbol{s}[\mu + \boldsymbol{\sigma}_{\boldsymbol{n}}^{\mathbf{2}}(\boldsymbol{t})(\mathbf{1} - \frac{\mathbf{1}}{\boldsymbol{K}})]$$

where $\sigma_n^2(t)$ describes the variance in *n* over time, which is generated by mutation and reduced by selection. The dynamics of mutation accumulation, beginning with an isogenic population, can then be split into three phases (see Fig. 1 of Lynch *et al.* 1993): first, mutations are accumulated rapidly, until mutation-selection-balance is reached. This is followed by a constant accumulation of mutations at constant population size (*i.e*., when the population is at its carrying capacity). Once the mean viability drops below 1/*R*, carrying capacity cannot be maintained and population size starts to decline, thus increasing the susceptibility to further accumulate deleterious mutations (which in turn again reduces mean viability), ultimately resulting in rapid population extinction.

One important difference in the dynamics of the meltdown model is the dependence on the order of events in the life cycle (*i.e.,* whether selection acts before or after population size regulation). In the first case, the carrying capacity *K* can be maintained over a long period, resulting in a constant population size and linear accumulation of mutations, followed by a rapid extinction phase (Lynch & Gabriel 1990; Lynch *et al*. 1993). In the second case, each click of the ratchet (*i.e*., when the least-loaded class of individuals is lost) results in fewer offspring; thus the population size declines gradually and the speed of the ratchet increases over time (Gabriel *et al*. 1993). Independent of the type of model, the conclusions from these papers were that extinction times of a few hundreds of generations are expected for small populations, and that the process is slowed by roughly an order of magnitude in sexual populations (Lynch *et al*. 1995b). A simple rule determines the beginning of the final meltdown phase, which was subsequently used in Lande's treatment of the same problem in a quantitative genetics framework (Lande 1994; 1998), and in models of lethal mutagenesis (Bull *et al*. 2007; and see below): the population is doomed to extinction when the mean viability decreases below the reciprocal of the absolute growth rate (*i.e*., the number of offspring an individual can produce; Lynch et al. 1993). An interesting finding is that an intermediate magnitude of the deleterious selection coefficient minimizes the time to extinction through mutational meltdown; this is of particular importance given the recently accumulating empirical evidence for a generally bimodal distribution of fitness effects of new mutations (*e.g.,* Eyre-Walker & Keightley 2007; Hietpas *et al.* 2011, 2013; Bank *et al.* 2014).

Beyond the "extinction threshold", other elements of lethal mutagenesis were indeed first discussed in the seminal papers on mutational meltdown (Lynch & Gabriel 1990; Lynch *et al*., 1993). First-

ly, Lynch *et al.* (1993) describe the conditions under which mutational meltdown is driven by genetic drift *versus* by mutational pressure, and conclude that "when the mutation rate is on the order of 1 per individual per generation […] the [least-loaded] class will be lost due to mutation pressure alone". Secondly, the authors demonstrate that the extinction time is only relatively weakly (logarithmically) dependent on the population size. Finally, Lynch *et al.* (1993) compare the mean fitness reached after the first phase with that of the infinite-population mutation-selection-balance (*i.e.*, the starting point for lethal mutagenesis; Kimura & Maruyama 1966), and observe that it is indeed only slightly larger in the case of a finite population. Thus, though generally associated with small-population size effects, mutational meltdown has been discussed with regards to both high-mutation-rate regimes and large population sizes as well (Lynch & Gabriel 1990).

Comparing and interpreting the models

"Everything should be made as simple as possible, but not simpler."
Albert Einstein

The notion of lethal mutagenesis arose out of the error catastrophe literature to provide a comprehensive framework to describe the deterministic (and thus population size-independent) processes that lead to population extinction via the crossing of a distinct error threshold. In contrast, the notion of mutational meltdown was fundamentally concerned with the stochastic nature of this extinction process, invoking classical population genetic models describing small population size dynamics. However, as discussed, the model of mutational meltdown has also been examined with regards to large population sizes. Further, recent extensions of the model of lethal mutagenesis have begun to consider stochastic effects. For example, Wylie & Shakhnovich (2012) studied the role of population size and mutation rate on extinction times, finding, as expected, that small populations may go extinct very quickly, whereas large populations survive almost indefinitely.

Thus, in some ways the model of lethal mutagenesis has converged with that of mutational meltdown, certainly in terms of appreciating the importance of stochastic processes in driving extinction events. Indeed, the genetic processes underlying population extinction are governed by the *effective* population size, N_e (Wright 1931; Crow 1954; Charlesworth 2009). Estimates of N_e/N taken from over 100 species (excluding viruses), demonstrated that census population size is on average an order of magnitude larger owing to fluctuating population sizes, unequal sex-ratios, and/or variance in reproductive success (Frankham 1995). In particular, the latter has been argued to significantly affect viral populations (Neher & Hallatschek 2013; Irwin *et al.* 2016). The ladder-like genealogy of the influenza A virus hemagglutinin segment, for example, suggests that only a few viruses seed the entire next generation (Grenfell *et al.* 2004). Along the same lines, estimates of effective population size in HIV range from 10^3 to 10^6, but generally show an extraordinarily low N_e/N ratio (Pennings *et al.* 2014).

Another similarity common to both models is the need to incorporate the potential effects of beneficial and compensatory mutations in modifying the rate of fitness decline and time to extinction (see Manrubia *et al*. 2010; Bull *et al.* 2013). Specifically, small increases in mutation rate may improve the ability of populations to respond to novel environmental challenges, and there indeed exist examples of selection for hypermutator strains in bacteria under particular stressors (*e.g.*, Sniegowski *et al.* 1997; Gerrish *et al.* 2013). However, owing to the effects of Hill-Robertson interference, as well as the far greater input of newly arising deleterious relative to beneficial mutations, this concern fundamen-

tally suggests a need to simply quantify the *extent* to which mutation rates must be increased in order to ultimately result in population extinction.

Recent empirical studies attempting to test the genetic models underlying lethal mutagenesis have largely failed to match its (qualitative) predictions (Springmann *et al.* 2010; Bull *et al.* 2013). This is presumably owing to the evolutionary mechanisms neglected by the (original) theory such as adaptive evolution (*i.e*., beneficial/compensatory mutations), interactions between mutations (*i.e.,* epistasis) and non-constant mutational effect sizes (*i.e*., the distribution of fitness effects [DFE]). While Antoneli *et al*. (2013) recently derived a generalization of the lethal mutagenesis extinction criterion that allows for a small fraction of (fixed effect) beneficial mutations, epistasis and the DFE are inherently connected to the genotype-to-fitness and/or the genotype-to-phenotype map. Two alternatives to the frequently used multiplicative fitness model of Kimura & Maruyama (1966) have been proposed: First, in biophysics-based fitness landscape approaches, the DFE is derived from the mutational effects on protein folding and its thermodynamic properties (*e*.g., Chen & Shakhnovich 2009; Stich *et al*. 2010; Wylie & Shakhnovich 2011, 2012). While these approaches may indeed account for an important class of mutational fitness effects in viruses (Wylie & Shakhnovich 2011), they have been criticized for directly equating fitness with protein stability - thus neglecting ecological aspects underlying fitness (Martin & Gandon 2010). As an alternative, phenotypic landscape models naturally accommodate epistasis, variation in mutational effects, and permit compensatory mutations (*e.g.,* Fisher's (1930) Geometric Model; for recent empirical support see Martin & Lenormand 2006a; Cooper *et al.* 2007; Hietpas *et al.* 2013; Achaz *et al.* 2014; Tenaillon 2014). Under these models, mutations, instead of directly affecting fitness, change n (unknown) quantitative traits (*e.g*., cell-to-cell transmission rate or levels of drug tolerance) which are considered to be under (Gaussian) stabilizing selection centered around an optimum. However, despite the conceptual differences between these two approaches, the resulting shapes of the DFEs are surprisingly similar (compare Fig. 3 in Wylie & Shakhnovich 2011 with Fig. 1 in Martin & Lenormand 2006a), perhaps simply emphasizing that variable mutational effects must be accounted for (Bull *et al.* 2013).

The best evidence for the empirical observation of population extinction driven by increased mutation rates comes from the experimental evolution literature in which these stochastic effects are prominent by design – for example, in yeast where population sizes were artificially kept small (Zeyl *et al.* 2001). Investigation in this area is particularly active in the study of RNA viruses, where the impact of mutagenic agents administered either alone or combined with antiviral inhibitors has been widely assessed (*e.g*., Loeb *et al*. 1999; Pariente *et al*. 2001; Crotty *et al*. 2001; Airaksinen *et al*. 2003; Grande-Pérez *et al*. 2005; Bank *et al.* 2016). In other words, this literature has focused on experiments directly modulating either effective population size or mutation rate. It is additionally of note that several empirical papers claiming to study mutational meltdown do not observe extinction of their study population (*e.g*., Rowe & Beebee 2003; Shoubridge & Wai 2008; Allen *et al*. 2009; Willi 2013; Woodruff 2013), which may partly be owing to a confusion of terminology: the *process* of Muller's ratchet and the *event* of mutational meltdown are sometimes used interchangeably.

In order to avoid future confusion, we propose here that focusing on these processes will likely be more informative for quantifying evolutionary dynamics and developing clinically relevant treatment strategies, rather than quibbling about the proper (largely semantic) nomenclature surrounding the extinction event itself. Namely, both models are fundamentally concerned with the notion of mutation-selection-(drift) balance (*e.g.,* Haldane 1937; Kimura & Crow 1964; Bürger 1989) and its induced mutational load (also see the helpful theoretical results of Hermisson *et al.* 2002 that relate this equi-

librium behavior to changing mutation rates). Further, Muller's ratchet and Hill-Robertson interference are fundamentally the processes driving the loss of fitness and eventual extinction. Though these processes have been well characterized in the population genetics literature, further progress must be made to extend these results to account for particular features of virus biology, namely the large variance in reproductive success, the population structure induced by compartmentalization, the interplay of fluctuating population sizes and changing environmental pressures, and the effects of tissue-specific drug permeability.

4.3 Methods

WFABC analysis

In oseltamivir treated replicates, the software developed by Foll *et al.* (2014) was used to estimate global effective population size N_e and the selection coefficients s for derived mutations from time sampled data for the allele frequency trajectories. Because the frequency of the third-most frequent mutation is very low in this dataset, all sites are treated as bi-allelic. Sites with a coverage above 100 were randomly (hypergeometrically) down-sampled to a sample size of 100. Only trajectories with a down-sampled frequency >2.5% were kept for the analysis, to ensure that these were above the estimated sequencing error of 1%. Following Foll *et al.* (2014), mutations with a Bayesian posterior distribution excluding zero of less than 0.5% ($P(s<0|x)<0.5\%$) were deemed to be under positive selection.

Population size estimates

The temporal method of Jorde & Ryman (2007) was used to estimate effective population size between two time points. This method uses *Fs'*, which is a measure of the variance in allele frequencies between two time points adjusted for sampling bias to calculate N_e

$$Fs = \frac{\sum (x-y)^2}{\sum z(1-z)}$$

$$Fs' = \frac{1}{t_{xy}} \frac{Fs\left[1-\frac{1}{2\tilde{n}}\right]-\frac{2}{\tilde{n}}}{\left(1+\frac{Fs}{4}\right)\left[1-\frac{1}{\tilde{n}}\right]}$$

where *Fs* is the estimator for allele frequency variance before adjusting for sampling bias, x and y are the allele frequencies at the two time points, t_{xy} is the number of generations between the two time points, z is the average frequency where $z=(x+y)/2$, and $\tilde{n}$ is the harmonic mean of the sample sizes n_x and n_y at each time point. For each passage, allele counts were hypergeometrically down-sampled to a minimum of 100 or coverage, and sites with counts less than 5 were excluded. Sites were included in the calculation of *Fs'* if one of the observed (randomly downsampled) frequencies was >2.5%. For each pair of passages, N_e was calculated as $1/Fs'$.

Population dynamics

Absolute growth rates for all replicates were obtained from the starting and final population sizes at each passage. The underlying assumption is that each viral plaque is the result of a single infective particle, based on the low multiplicity of infection (MOI) used for each passage. Following Foll *et al.* (2014) and Bank *et al*. (2016), we assume 13 generations of viral populations per passage and calculate the Malthusian growth rate *r* per passage as

$$N(t) = N_1(exp^{rt})$$

where *t* is the number of generations, $N(t)$ is the population size at time *t,* and N_1 is the initial population size at the start of each passage. Relative growth rates are calculated as $r_{combined} - r_{osel}$, which corresponds to an estimate of relative fitness, based on using the oseltamivir replicates as a control (as the oseltamivir replicates exhibit stable growth rates and no drug paired controls were not available). We carried out a linear regression between relative growth rates and time in passage number. An ordinary student's t-test was used establish whether the slope of the line, corresponding to the change in relative growth over time, was significantly different from 0, indicating a significant difference between the treatments. Following Bank *et al.* (2016) the slope of the line was taken to indicate whether there was a continuing decline in relative growth (negative slope) or potential recovery and adaptation to the combined drug protocol (positive slope) over time.

Hierarchical clustering analysis

The hierarchical clustering analysis is based on the squared Euclidian distance between allele frequency trajectories of the candidate mutations using Ward's minimum variance criterion (Ward Jr 1963), starting from the first time point where the frequency was higher than the estimated sequencing error of 1%.

4.4 Results

We analysed the evolution of influenza A/Brisbane/59/2007 (H1N1) in Madin-Darby canine kidney cells under treatment with oseltamivir alone, or under treatment with a combination of oseltamivir and favipiravir, over a total of 10 passages. In the first three passages, IAV was adapted from chicken egg and serially amplified in the MDCK cells with no treatment, as part of an earlier experiment (Foll *et al.* 2014). Stock viral populations from an earlier experiment were used to seed passage 4 to ensure that replicates are identical before administration of the drug treatment. In passages 4 to 8, three replicates of IAV were exposed to increasing doses of a combination of oseltamivir and favipiravir (Fig. 6). The replicates were paired with three populations exposed to increasing doses of oseltamivir only as a control. This ensured that the oseltamivir only replicates are subject to the same experimental conditions as the populations exposed to the combination treatment. A multiplicity (MOI) of 0.01 was used for all replicates, except replicate 1 passage 10 (MOI=0.005). Following Foll *et al.* (2014), 13 viral generations are assumed to occur during each passage. Details of drug treatment, MOI and output PFU for each replicate are shown in Table 2. These results were compared to results obtained from a previous experiment, where two populations of IAV were exposed to favipiravir alone

and to a no drug control over passages 4 to 15 Bank *et al.* (2016). The favipiravir population was treated 2μM of the drug from passage 4, and the dosage was doubled at every passage, which represents a higher dosage than for the combined drug replicates (dosage of 1μM favipiravir in passage 4 and doubling of the dose thereafter). At the end of each passage, samples from each replicate were sequenced using high coverage, whole-genome high-throughput population sequencing.

4.4.1 Genetic diversity and purifying selection

We calculated genetic diversity for the different replicates as the expected heterozygosity (Nei 1973). Genetic diversity was consistently low throughout passages 4 to 10 (Fig. 7). It was slightly lower for the oseltamivir only replicates ($2.0x10^{-3}$ average for replicate 1, $1.8x10^{-3}$ for replicate 2, $1.5x10^{-3}$ for replicate 3) compared to the combination drug replicates ($3.0x10^{-3}$ for replicate 1, $2.6x10^{-3}$ for replicate 2, $2.4x10^{-3}x10^{-3}$). The SFS (shown in Fig. 8 for replicate 1 passages 4 and 8) shows a bias towards low frequency variants, for all replicates and all treatments. Low frequency counts increased by passage 8 because of the mutagenic effect of favipiravir on the number of segregating sites. Only a relatively small number of mutations had a derived allele frequency (DAF) in excess of 40% in all replicates (Table 1), although higher in the combined drug replicates. Together, these points confirm the influence of strong purifying selection in viral populations acting to remove new deleterious mutations, as found by others (*e.g*. Foll *et al.* (2014)), and evidence the action of favipiravir on the number of segregating sites in combined drug replicates.

4.4.2 Evidence for mutation accumulation and Muller's ratchet

All combined drug replicates reached extinction or near extinction by passage 10. The output number of plaque-forming units per ml (PFU/ml) tracks census viral population size at the end of each passage. Output PFU/ml declined to $3.0x10^{2}$ for replicate 1, $8.0\ x10^{3}$ for replicate 4 and $2.4x10^{4}$ for replicate 3 (Table 2). In contrast, output PFU/ml for oseltamivir only replicates remain high and stable throughout all passages ($6.0x10^{5}$ PFU/ml for replicate 1, $5.0x10^{6}$ PFU/ml for replicate 2 and $7.0x10^{6}$ PFU/ml for replicate 3,), indicating that the virus has rapidly evolved resistance to this drug (see section on resistance to oseltamivir for more details).

We calculate relative growth per passage (a measure of the fitness of the viral population) as output/input PFU/ml and show its progression on Fig. 9A. All passages and replicates have the same multiplicity of infection (MOI = 0.01), giving a constant input PFU, except for combined drug 1 passage 10 (with an MOI of 0.005 because of low viral titers). We observe that the relative growth for combined drug replicates declines more rapidly than for the population treated with favipiravir alone, despite a lower dosage of favipiravir. Relative growth for combined replicate 1 has a sharp recovery at passage 7 followed by a rapid decline, while the other two replicates decline steadily. Relative growth for the oseltamivir replicates remains relatively stable, although below the level exhibited by the no drug population. Based on randomly down-sampled frequencies, we calculate the total number of sites segregating above a 1% derived allele frequency (DAF) (the estimated sequencing error) in the

populations at each passage (Fig. 9B) to explore whether a higher segregating mutation load is responsible for the earlier decline. We find that the number of segregating mutations is lower for the combined drug replicates than for the favipiravir replicates. There are peaks in the number of segregating mutations at passage 6 for combined drug 1 and passage 8 for combined drug 2, followed by a rapid reduction. We explain these trends in the next section on genetic hitchhiking. In contrast, oseltamivir replicates exhibit constant low mutation load and stable viral population sizes, confirming that these populations remain in mutation-selection-drift balance.

Combined drug replicate 1 provides the strongest evidence of transition into a phase of rapid population collapse (negative relative growth) and escalating mutation accumulation, similar to the dynamics observed in passage 14 and 15 of the favipiravir population. The total number of segregating mutations increased 8x to 1,800 in passage 10 from 202 in passage 9. For replicates 2 and 3, there is a slowing of relative growth to zero and an increase in the number of segregating sites above the levels exhibited by oseltamivir control replicates, but the scale of the end dynamics are different from replicate 1, and lack the transition into population contraction.

A possible explanation for the more rapid decline of combined drug replicates as opposed to those treated with each drug alone is the accelerated action of Muller's ratchet. Under this theory, the mean number of deleterious mutations per individual accumulates at a constant rate (Muller 1964; Felsenstein 1974), assuming a static mutation rate and constant population size. The rate of this process – the speed of the ratchet - increases exponentially with mutation rate and decreases with population size and with the selection strength of deleterious mutations (Haigh 1978; Gordo & Charlesworth 2000a, b). Background selection accelerates the speed of the ratchet, owing to the presence of strongly deleterious linked mutations, which reduce the effective population size N_e because neutral and weakly selected mutations can only fix if they occur on genetic backgrounds that exclude these strongly deleterious mutations (Gordo & Charlesworth 2001). In combined drug and favipiravir treated populations subject to an influx of mostly deleterious and neutral mutations, we would expect an acceleration in the rate of Muller's ratchet, particularly in the case of declining population size. The evolutionary processes discussed here are determined by effective population size N_e rather than census size (Wright 1931; Charlesworth 2009). Because genetic drift has greater impact in small populations, we explore whether differences in effective population size exist between the favipiravir and combined drug replicates, and use this as a means to indirectly assess the contribution of Muller's ratchet to the overall decline in virus populations, and to quantify the rate of mutation accumulation.

Theoretical work and simulations have shown that the fixation of neutral and weakly deleterious mutations is a robust indicator of the loss of least loaded classes and therefore of the speed of Muller's Ratchet, assuming a haploid asexual population under an influx of deleterious mutations with the same selection coefficient (Charlesworth & Charlesworth 1997; Bergstrom & Pritchard 1998; Gordo & Charlesworth 2001). However, because of strong purifying selection, here we observe a relatively small number of mutations segregating above a DAF of 40% and fixing in the population (Table 1 and Fig. 14), most of which can be attributed to positive selection or to genetic hitchhiking (see next sections). Therefore, the rate of fixation of deleterious mutations estimates the rate of genetic hitchhiking rather than the speed of the ratchet in this case.

The serial passaging of virus populations creates a series of bottlenecks followed by exponential growth (or contraction) over 13 viral generations during each passage, which depress effective population size compared to end of passage census size. The strong purifying selection seen in IAV populations also acts to reduce effective population sizes (Charlesworth *et al.* 1993). We apply two methods to estimate effective population size per passage, and we use these estimates to calculate the rate mutation accumulation. Firstly, we apply the temporal method of Jorde & Ryman (2007), which leverages the variance in allele frequencies between two time points, and adjusts for sampling effects, to generate an unbiased estimator of effective population size, under the assumption of a constant mutation rate (see methods). However, the increasing mutation rate here will generate additional variance in allele frequencies and likely cause effective population sizes to be underestimated. Secondly, we calculate the effective population size at each passage based on the harmonic mean of the population size at each generation (Ewens 1967). However, this method ignores differences in virion budding, which create skewed offspring distributions (Irwin *et al.* 2016) consistent with the fact that only a few virions seed subsequent generations (Grenfell *et al*. 2004). Estimates of effective population size made using this second method will likely be inflated. We therefore calculate the rates of mutation accumulation using these two estimates of effective population size as a lower and an upper limit, respectively.

Because of the variance induced by the increasing mutation rate and hitchhiking patterns (see next section), we find that Jorde & Ryman's (2007) method is not a reliable estimator of effective population size in combined drug replicates, and we use this method to assess average trends only. A rapid drop in effective population size for all three combined drug replicates after passage 6 is observed (Fig. 10), and is probably exacerbated by the hitchhiking effect. Excluding passage 10 for replicate 1, mutation load increased at a rate of 0.49-1.45 segregating mutations per individual per passage (0.49 for replicate 1 passages 4-9, 1.45 for replicate 2 (all passages) and 0.82 for replicate 3 (all passages)). The rate of mutation accumulation is similar in order of magnitude across all three replicates but with stochastic variation (R^2 of 0.172 to 0.767), partly due to the heterogeneity in estimates of N_e. The level of mutation accumulation is significant for replicate 2 only (at the 1% level, p=0.0098, student's t test). At the point of population collapse in passage 10 for replicate 1, mutation load escalates to 77.5 per individual. As expected, the estimates of N_e using the harmonic mean of population per size are several orders of magnitude higher (Fig. 10 D-F) than estimates using Jorde & Ryman's (2007) method, leading to a lower estimated rate of mutation accumulation (0.0027-0.0059 per individual per passage) (significant at the 1% level for replicate 2, p=0.0094 and at the 5% level for replicate 3, p=0.034, student's t test).

Because of the limitations of the above methods, global estimates of effective population size were obtained from WF-ABC (Fig. 11) to enable a comparison between the favipiravir-treated population and the combined drug replicates, with the caveat that the assumptions of WFABC are not fully respected in these populations. We find that estimates of effective population size for combined drug replicates (N_e=239.16 for replicate 1, N_e=216.19 for replicate 2 and N_e=161.64 for replicate 3) are similar to the global estimates for the favipiravir population (N_e=209.15) and cannot account for the differences between the populations. As expected, these global estimates are lower than for oseltamivir only replicates (N_e=519.21 for replicate 1, N_e=269.83 for replicate 2 and N_e=392.16 for replicate 3).

The dynamics of the combined drug replicates appear to fit a classic model of mutational meltdown, with an increase in mutation load until a sharp threshold of rapid population collapse and escalating mutation load is reached. Evidence of small and declining effective population sizes, and of stochastic evolutionary dynamics support a role for Muller's ratchet, in addition to the deterministic process of mutation accumulation owing to the shift in mutation-selection balance induced by the increased mutation rates. However, the rate of viral decline is faster than in favipiravir treated populations and the segregating mutation load is lower. We cannot account for these differences with estimates of effective population size, which govern processes such as Muller's ratchet. Therefore, in the next sections we explore other processes, which may act in conjunction with Muller's ratchet, and in particular the potential effects of selective sweeps in the combined drug populations.

Assessing signs of evolutionary rescue (Alexander *et al.* 2014) is important because of the risk of adaptation of the virus to favipiravir, and because of the potential development of resistance. In a previous study, Bank *et al.* (2016) observed a deceleration in relative decline in two favipiravir populations, where the drug was either held constant or was withdrawn in passages 10-15 (Fig. 5). This was highlighted as the first sign of adaptation of the virus to a low dosage of favipiravir. Here, we estimate relative growth rates for the combined drug replicates, and calculate these as the difference between the absolute growth rates per generation for the combined drug replicates and for the oseltamivir replicates (which are used as control because of stable population sizes). We observe an increasing rate of relative population decline, and find no evidence of adaptation of the virus under a protocol of escalating dosage (Fig. 12). Further studies with two constant doses of favipiravir (alone and in combination with oseltamivir) are underway to establish the level of favipiravir required to ensure viral extinction and prevent adaptation of the virus.

4.4.3 Mutations putatively evolving under positive selection

To identify positively selected mutations, the posterior distributions for estimates of s were utilized (*i.e.* with a posterior density interval for the selection coefficient s excluding zero of less than 0.5%, $P(s < 0|x) < 0.005$)). WFABC differentiates trajectories of mutations under selection from those due to genetic drift under the assumption that all sites are unlinked and independently selected, and that there is a static mutation rate. Because many of the trajectories in the oseltamivir-only, combined drug and favipiravir-only replicates exhibit non-standard trajectories (*i.e.* trajectories that are not typically derived from the diffusion equation), and because the assumptions of a constant mutation rate and of unlinked sites do not always hold, we also track trajectories that exceed a derived allele frequency (DAF) of 40% at any time point, using randomly down-sampled counts to calculate frequencies.

We identified 11 mutations potentially evolving under positive selection in oseltamivir replicates (see Fig. 13 A-C), including the known resistance mutation NA H275Y. Four of these mutations are synonymous and seven are non-synonymous. These seven non-synonymous contending mutations are also the only ones that arise in more than one replicate (outside of the NA region) in either the combined drug replicates or in the favipiravir-only treated population. Neutrality is rejected for all of these mutations in at least one oseltamivir replicate (except for NP D101N, which only exhibits a "standard" trajectory in the favipiravir-only population) and WFABC was used to estimate selection coefficients and Bayesian p-values (Table 1). The varied trajectories and limited clustering of these

mutations in oseltamivir-only replicates supports the assumption that these are unlinked, independently selected sites, although epistatic interactions cannot be excluded.

In contrast, three of the four synonymous mutations to reach a DAF in excess of 40% arise in one oseltamivir replicate only; two cluster with the strongly selected H275Y (HA L73L in oseltamivir 1, and NA P326P in oseltamivir 2) suggesting probable genetic hitchhiking of neutral variants. The third (PA D67D) has a non-standard trajectory and its functional significance is unknown. These synonymous mutations are not discussed further. The fourth synonymous mutation, PA G58G arises in all oseltamivir and combined drug replicates, as well as in the favipiravir population. It is inferred to evolve under positive selection in oseltamivir 1 and fixed in this population (along with MP1 E23Q). It exhibited non-standard trajectories in other oseltamivir and combined drug replicates, as well as in favipiravir and no drug replicates, and clusters with MP1 E23Q, suggesting both a cell adaptation function and a possible epistatic interaction with MP1 E23Q.

Fig. 13 tracks the putative positively selected mutations in combined drug replicates (Fig. 13 D-F) and in favipiravir and no drug populations (Fig. 13 G&H). In addition to H275Y, two other non-synonymous mutations (A454V and E128G) in the NA region that is important for oseltamivir resistance arise in the combined drug replicates only; one of these mutations, NA A454V arises in both combined drug replicates 2 and 3. In addition to the 11 contending positively selected mutations described above, this gives a new total of 13 contending beneficial mutations that are tracked in the combined drug (Fig. 13 D-F) and in the favipiravir treated (Fig. 13G) and no drug (Fig13H) populations (Table 3). While there is no recombination in influenza (but see (Bao *et al.* 2008)), here many of the significant mutations occur in different segments, so that in the absence of recombination, selection of the fittest haplotype could be facilitated by segment reassortment (an established process in H1N1) as well as by mutation.

We explore the possible biological function of these 13 mutations (Table 3) in the light of previous studies. Outside of the NA region, the HA region contains the mutations with the highest selection coefficients: D112N and E78G. This is consistent with studies showing that changes in the HA region counter the deleterious growth effects of H275Y (Bloom *et al.* 2010; Ginting *et al.* 2012). The HA D112N mutation was previously identified by Foll *et al.* (2014) and has been described in other influenza strains and HA serotypes (Daniels *et al.* 1985; Reed *et al.* 2009): it acts by inducing a pH change at the point of endosome and viral fusion, thereby improving IAV infectivity (Thoennes *et al.* 2008). Here it is significant in oseltamivir replicate 2 (s=0.126, Table 3) and in the no drug comparison population, suggesting a role in cell adaption. A newly identified HA mutation, E78G, is present in all oseltamivir and combined drug replicates where H275Y is present (5 out of 6 populations), but not in combined drug replicate 2 where H275Y is absent, indicating a possible epistatic interaction with H275Y, although the trajectories of E78G and H275Y are not always aligned. E78G is significant in oseltamivir replicate 1 and 2 (s=0.114 and s=0.117 respectively, Table 3) and fixes in combined replicate 3, whereas in oseltamivir 2 and combined replicate 1 its trajectory suggests clonal interference (see next section).

Mutations in the M1 region have been suggested to have compensatory benefits upon interacting with H275Y, namely by improving the process of virion budding and helping to overcome the fitness cost of the H275Y mutation, reflected in the lower amount of neuroaminidase to reach the cell surface (Jin *et al.* 1997; Noton *et al.* 2007; Rossman & Lamb 2011). E23Q was identified in both previous sets of experiments (Foll et al. 2014, Bank et al 2016) and is significant in oseltamivir 1 (s=0.057),

favipiravir and no drug populations, with trajectories typical of clonal interference in the other replicates. A37V is very close to the previously identified A41V (Foll *et al*. 2014) and may serve a similar function in improving virion budding; it clusters with D101N in some replicates and E32Q in others and arises in oseltamivir 2 and 3 (*s*=0.038) and combined 1 and 2. NP mutation D101N has been previously screened as resistance mutation to the mutagenic drug ribavirin (Cheung *et al.* 2014) with inconclusive results; here we find that it is present in all populations except the no drug control, including the oseltamivir only populations, and is therefore also likely to have a role in cell adaptation, or possibly in improving the formation of infective virions (Noton *et al.* 2009). Lastly, D125N, a mutation in the NS1 region, is present in oseltamivir replicate 2 and in the no drug control, and is therefore also hypothesised to have a general adaptive function to changing cell conditions.

Although H275Y does not appear more rapidly in combined drug populations than in oseltamivir-only replicates, the two other NA non-synonymous mutations described above (A454V and E128G) only appeared in combined drug replicates (Fig. 13) and exhibit trajectories characteristic of strong selection coefficients (not estimated using WFABC because the assumptions of constant mutation rate and unlinked sites do not hold in those populations). We thus put forward the hypothesis that the mutational input from favipiravir allows the virus to rapidly explore sequence space for alternative oseltamivir resistance solutions within the limits of the time to extinction imposed by the increasing load burden.

4.4.4 Effects of genetic hitchhiking

Under a model of genetic hitchhiking, neutral or weakly selected sites in physical linkage to strongly beneficial mutations will rise in frequency (Maynard Smith & Haigh 1974). Here, we find that the positively selected non-synonymous mutations identified in the oseltamivir replicates show greater evidence of clustering in the combined drug populations (Fig. 13 D-F) than in the oseltamivir-only populations (Fig. 13 A-C) (although there is also evidence of clonal interference of beneficial mutations, see next section). This clustering of beneficial mutations may be seen around the strongly selected NA mutations H275Y, A454V and E128G. Potential beneficial mutations for cell adaptation including D101N and A37V (with trajectories characterized by a low selection coefficient in some replicates) are rapidly driven to fixation by association with H275Y in combined replicate 1 and A454V in combined replicate 2. In contrast, hitchhiking in the oseltamivir replicates occurs in synonymous (assumed neutral) variants (HA L73L with H275 in replicate 1 and NA P326P with H275Y in replicate 2). This suggests that in addition to generating possible alternative oseltamivir resistance mutations, the enhanced mutational input of favipiravir (acting jointly with purifying selection and reassortment) serves to optimise combinations of beneficial mutations on different haplotypes by quickly exploring different combinations in sequence space, and hitchhikes them to fixation with the strongly selected NA mutations.

Tracking the remaining mutations specific to the combined drug replicates and segregating in excess of 40%, revealed that these also cluster with the strongly selected NA mutations and with the ubiquitous HA mutation E78G (H275Y and E78G in replicate 1, A454V in replicate 2, and A454V, E128G, and HA E78G in replicate 3, see Fig. 13 D-F); this was confirmed through hierarchical clustering analysis (Fig. 14). These mutations are unique to each combined drug replicate and are therefore assumed to be largely deleterious or neutral, consistent with studies indicating a bi-modal distribution of fitness effects with only a small tail of beneficial mutations (Eyre-Walker & Keightley 2007; Bank *et al.* 2014).

Analyses suggest that in the combined drug replicates the strongly selected beneficial mutations hitchhike not only other beneficial mutations but also a high mutation load to fixation. The timing of the selective sweeps coincides with a sharp decrease in the total number of segregating mutations after passage 6 in combined replicate 1 and after passage 8 in combined replicate 2 (Fig. 9), suggesting that the sweeps reduce genome-wide variation but at the cost of fixing deleterious mutations that depress viral fitness. A likely explanation is that the rapid trajectory to fixation of the beneficial mutations does not allow purifying selection sufficient time to purge the linked deleterious mutations, particularly as effective population size has diminished and selection is therefore less efficient. In addition, there is a constant input of new deleterious and neutral mutations segregating at a low frequency from the impact of favipiravir. Ultimately this combined high load burden accelerates the decline in viral fitness and precipitates the population towards extinction.

Thus the strength of beneficial mutations (as indicated by the shape of their trajectory and the corresponding WFABC estimates) governs the size of the linked mutation load and potentially accelerates the process of extinction. As shown, the trajectories of mutations in the favipiravir-treated population are more random and diffuse (Fig. 13G), with fewer obvious clusters. The tracked beneficial mutations in the favipiravir population mediate cell adaptation and are less strongly selected than NA mutations in the oseltamivir and combined drug replicates with longer trajectories, giving purifying selection more time to act. The absence of strong selective sweeps reducing genome-wide variation is reflected in the high and escalating number of segregating mutations for favipiravir populations observed in Fig. 9B. Significantly, the favipiravir population reaches extinction by passage 15, compared to passage 10 for combined replicate 1, suggesting a tentative hypothesis that the weaker hitchhiking dynamics are at least partly responsible for the later point of collapse.

4.4.5 Effects of Hill-Robertson interference

There are different types of Hill-Robertson (Hill & Robertson 1966) interference, depending on the relative strength or weakness of the interacting mutations and whether they are deleterious or beneficial. In a non-recombining, asexual population such as H1N1 here, strongly selected beneficial alleles arising on different haplotypes compete for fixation (Fisher 1930, Muller 1932, and see Barton 2010). There is a build-up in negative linkage disequilibrium between these "repulsion haplotypes" (Hill & Robertson 1966; McVean & Charlesworth 2000)(*i.e.* the beneficial mutations and their linked variants are found associated less frequently than by chance), and thus a reduction in the efficacy of selection. If the haplotypes carrying competing beneficial mutations have similar net fitness, they can endure in populations, but if not, the less fit haplotype will become extinct. In contrast, beneficial mutations arising on the same haplotype (because of mutational input or reassortment here) will be in positive linkage disequilibrium and will rapidly fix (Hill & Robertson 1966; McVean & Charlesworth 2000). The sojorn time of a beneficial mutation is $(1/s)\log(4N_e s)$ (Barton 2010) which gives a short window for the association to occur when the selection coefficient s of the most beneficial allele is strong. Assuming multiplicative fitness effects, each beneficial allele must have sufficiently high fitness to overcome the mutation load of linked deleterious mutations (and its fitness will be reduced by a factor of e^{-U}, where U is the combined deleterious mutation strength/rate assuming a multiplicative fitness model).

In the combined drug replicates, we find many examples of non-standard trajectories of beneficial mutations (known to be mutations under weak positive selection in the oseltamivir replicates), which are characterized by a rapid rise and decline. These trajectories suggest patterns of clonal interference between beneficial mutations: haplotypes carrying the strongly selected NA resistance mutations H275Y or A454V outcompete haplotypes carrying weaker mutations mediating cell adaptation, in cases where these weaker mutations do not hitchhike with the resistance mutations. Indeed, in the combined drug replicates, these weak beneficial mutations only fix if they are associated with the resistance mutations. For example, in combined replicate 1, the H275Y haplotype sweeps the associated NP D101N and A37G to fixation; its rise coincides with the decline of haplotypes carrying E23Q, and possibly slows the rise of PA G58G (which clusters with E23Q in many replicates). In combined replicate 2, the rise of the haplotype carrying NA A454V leads to the extinction of MP1 E23Q, MP2 A37G, PA G58G and NP D101N, until NP D101 is reshuffled onto the A454V background and fixed. In combined replicate 3, the rapid spread and fixation of haplotypes carrying NA A454V, NA E128G and HA E78G coincides with the decline of MP1 E23Q (Fig. 13 D-F). There is some evidence of these effects in the oseltamivir replicates but not in the favipiravir or no drug replicates, where the spread of beneficial mutations mediating cell adaptation follow more standard diffusion-based trajectories (except for PA G58G) characteristic of weakly selected mutations (Fig. 13 G&H).

In the favipiravir population, we observe the influence of a different form of interference: that between linked weakly selected beneficial and deleterious mutations (WSHRI) (Hill & Robertson 1966; McVean & Charlesworth 2000). There is limited clustering and significant variance in allele trajectories over the longer lifespan of this population (Fig. 13G). Linkage is likely between weakly selected mutations mediating cell adaptation (and other, non identified weakly beneficial mutations) and the increasing influx of slightly deleterious and neutral mutations. The high variance in allele trajectories is not observed to the same extent in the no drug (Fig. 13H), in the oseltamivir populations (Fig. 13A-C) or in the combined drug populations (Fig.13D-F). Hitchhiking of beneficial, deleterious and neutral variants with the resistance mutations (and with HA E78G) accounts for almost all of the mutation trajectories in the combined drug populations, with the exception of 1) combined replicate 1 passage 10 where genetic drift (Muller's ratchet) leads to the fixation of many neutral and deleterious mutations owing to population collapse in this replicate and 2) combined replicate 2 passage 8 where there is a cluster segregating at less than 40% DAF (not accounted for).

4.5 Discussion

Here we find evidence of mutation accumulation in combined drug populations that reduce viral fitness and ultimately result in the extinction of H1N1 viral populations subject to the combined drug protocol. The results appear to fit a classic mutational meltdown model, characterised by increasing mutation accumulation until a sharp transition point of meltdown is reached, as reported by others on the impact of favipiravir alone on IAV *in vitro* (Baranovich *et al.* 2013; Bank *et al.* 2016). This is particularly evident for replicate 1. Intriguingly, despite a lower mutation load, we observe a more rapid decline in relative growth rate in the combined drug population than in the favipiravir treated population, and try to disentangle the processes behind this.

We find that strongly selected oseltamivir resistance mutations (and one ubiquitous HA mutation, E78G, that has potential epistatic interactions with H275Y) influence the evolutionary dynamics in

combined drug replicates by sweeping deleterious mutations to fixation, along with other weak beneficial mutations (Fig. 13, 14 and summarised in Fig. 15). These selective sweeps cause a reduction in genome-wide variation relative to the favipiravir population, reflected in a lower number of segregating mutations (Fig 9). The timing of these sweeps coincides with sharp drops in the number of segregating mutations after passage 6 in replicate 1 and passage 8 in replicate 2. This striking evolutionary dynamic is not apparent in the favipiravir-only populations, where the identified beneficial mutations mediate cell adaptation and are less strongly selected, giving purifying selection more time to dissociate hitchhiked deleterious mutations. We therefore tentatively hypothesise that the fixation of deleterious mutations hitchhiking with the strongly selected oseltamivir resistance mutations reduces viral fitness and induces an earlier point of extinction (passage 10 for combined drug replicate 1 relative to passage 15 for favipiravir populations). Potentially, we note that the strongly selected mutations required to hitchhike deleterious mutations need not be oseltamivir resistance mutations – but we observe that few mutations except drug resistance mutations would have the necessary fitness benefit.

We identify two novel candidate mutations that only arise in combined drug replicates (NA A454V and NA E128G - although NA E128G clusters with NA A454V in replicate 3 and may also be hitchhiking). We also re-identify several mutations that are known to mediate cell adaptation and to provide compensatory mechanisms to offset the adaptive cost of H275Y mutations, reflected in reduced virion budding.

We attempt to tease out other processes contributing to population collapse in combined drug populations. Global estimates of small effective population size are consistent between the favipiravir and combined drug replicates and support a role for Muller's ratchet. We cannot use the rate of fixation of deleterious mutations to determine the speed of Muller's ratchet because very few mutations fix (Table 2), and because hitchhiking rather than genetic drift appears to drive fixation of those mutations. The oseltamivir resistance mutations (and HA E78G) also have clonal interference effects, preventing weaker beneficial mutations on haplotypes not containing the resistance mutations from spreading in the combined drug populations. In the absence of strongly selected mutations in the favipiravir population, the high variance in allele frequencies suggests a more important role for WSHRI between weak beneficial and deleterious mutations than for clonal interference.

Only mutations that have been identified as under positive selection in the oseltamivir replicates arise in more than one combined drug replicate (except for one NA mutation) – suggesting that these represent the bulk of mutations under positive selection in the combined replicates too, despite the enhanced mutation rate. This indicates that there is a limited repertoire of beneficial mutations that can be accessed by viral populations. This finding is consistent with theoretical work and experimental studies showing that the distribution of fitness effects includes only a small tail of beneficial mutations (Eyre-Walker & Keightley 2007; Bank *et al.* 2014)

To the best of our knowledge, this is the first study attempting to elucidate the mechanisms behind the synergistic effects of oseltamivir and favipiravir in H1N1 virus populations. New four-arm experiments are currently underway that entail escalating doses of oseltamivir combined with fixed low (2μM) and high (4μM) dosages of favipiravir in IAV populations, which will be compared with favipiravir-only and oseltamivir-only treated populations. The results from these experiments will provide welcome evidence to confirm whether the hitchhiking patterns observed here are replicated, and to establish the dosage of favipiravir required to achieve meltdown.

4.6 Tables

Table 1. Summary information on the number of segregating sites

Replicate	Number of sites with DAF> 40%
Comb 1	14
Comb 2	14
Comb 3	24
Osel 1	6
Osel 2	8
Osel 3	5

Table 2 Summary output PFU/ml

Pas-sage	Osel (*μM*)	Favi (*μM*)	MOI	Output PFU/ml					
				Comb 1	**Comb 2**	**Comb 3**	**Osel 1**	**Osel 2**	**Osel 3**
3	0	0	0.01*						
4	0.1	1	0.01	2.0E+06	1.1E+07	1.1E+07	4.0E+06	3.0E+07	3.6E+07
5	0.2	2	0.01	4.0E+04	1.3E+06	7.0E+05	2.6E+06	3.2E+07	3.3E+07
6	0.4	4	0.01	1.0E+04	1.4E+06	1.1E+06	6.0E+06	8.0E+06	2.1E+07
7	0.8	8	0.01	2.8E+06	3.0E+05	1.0E+05	8.6E+06	4.0E+06	9.0E+06
8	1.6	16	0.01	8.0E+04	1.1E+05	2.8E+05	7.0E+06	2.0E+06	6.0E+06
9	3.2	32	0.01	2.0E+04	9.0E+04	3.0E+05	9.0E+05	3.0E+06	3.0E+06
10	6.4	64	0.01#	3.0E+02	8.0E+03	2.4E+04	6.0E+05	5.0E+06	7.0E+06

* MOI of 0.0103 in comb1, comb2, osel1 and osel2

MOI of 0.005 in comb1 and osel1

Table 3 Mutations inferred to be evolving under positive selection

Seg name	Position	Ref base	Mut base	Type S/NS	SNP	WFABC s*	Bayesian P value (p<0)	Previd	Replicates	Functional interpretation
PA	199	G	T	S	G58G	0.026	0.017	Y	All	Cell adaptation
PA	225	C	T	S	D67D			N	osel3	Synonymous
HA	1280	G	A	S	L73L	0.099	0**	N	osel1	Synonymous
HA	1294	A	G	NS	E78G	0.114(1) 0.117(3)	0** 0**	N	osel123 & comb13	Possible epistasis with H275Y
HA	1395	G	A	NS	D112N	0.126	0**	Y	osel2 & nodrug	Cell adaptation
NP	346	G	A	NS	D101N	0.027(1)	0.0349	Y	all except nodrug	Cell adaptation
NA	403	A	G	NS	E128G	§		N	comb3	Resistance mutation
NA	843	C	T	NS	H275Y	0.125(1) 0.209(2) 0.218(3)	0** 0** 0**	Y	osel123 & comb13	Known resistance mutation
NA	998	G	T	S	P326P	0.075	0.002**	N	osel2	Synonymous
NA	1381	C	T	NS	A454V	§		N	comb23	Possible resistance mutation
MP1	92	G	C	NS	E23Q	0.057(1)	0**	Y	All	Compensatory mutation
MP2	848	C	G	NS	A37G	0.038(3)	0.039	N	osel23 & combf12	Compensatory mutation
NS1	399	G	A	NS	A125D	0.036	0.03	N	osel2 & nodrug	Cell adaptation

§ in combined drug replicates only, strength not estimated (the assumptions of a constant mutation rate and of unlinked sites do not hold)

* the numbers in brackets indicate the oseltamivir replicate used for the estimation (where the mutation arises in several replicates)

** significant (p<0.005)

Table 4 Mutations segregating with DAF >40% in the combined drug replicates

excluding the mutations under positive selection listed in Table 3. These mutations are assumed to be neutral or deleterious (see text). Mutations marked with a question mark do not have a known amino acid mutation equivalent using H3 numbering conversion.

Seg	Position	Seg name	Ref base	Mut base	Ref AA	Mut AA	Type	AA number	SNP	Dataset
seg1	46	PB2	C	T	L	L	S	7	L7L	comb3
seg1	212	PB2	G	A	R	K	N	62	R62K	comb3
seg1	780	PB2	G	A	R	R	S	251	R251R	comb3
seg1	867	PB2	A	G	L	L	S	280	L280L	comb2
seg1	1650	PB2	C	T	G	G	S	541	G541G	comb3
seg1	2078	PB2	C	T	A	V	N	684	A684V	comb3
seg1	2132	PB2	G	A	R	K	N	702	R702K	comb3
seg2	730	PB1	C	T	L	L	S	236	L236L	comb3
seg2	1089	PB1	C	T	Y	Y	S	355	Y355Y	comb1
seg2	2058	PB1	C	T	S	S	S	678	S678S	comb3
seg3	323	PA	C	T	A	V	N	100	A100V	comb1
seg3	1431	PA	G	A	L	L	S	469	L469L	comb3
seg3	1854	PA	G	A	E	E	S	610	E610E	comb2
seg3	2175	PA	C	T					?	comb1
seg4	210	HA	C	T	L	L	S	60	L60L	comb1
seg4	767	HA	C	T	Y	Y	S	245	Y245Y	comb3
seg4	1261	HA	G	A	G	D	N	410	G410D	comb2
seg4	1332	HA	A	G	I	V	N	434	I434V	comb2
seg4	1551	HA	G	A	E	K	N	507	E507K	comb2
seg5	210	NP	G	A	R	R	S	55	R55R	comb1
seg5	815	NP	C	T	T	I	N	257	T257I	comb3
seg5	1023	NP	C	T	S	S	S	326	S326S	comb3
seg5	1327	NP	G	A	A	T	N	428	A428T	comb2
seg5	1369	NP	G	A	A	T	N	442	A442T	comb2
seg5	1500	NP	A	G	G	G	S	485	G485G	comb3
seg7	208	MP1	A	G	G	G	S	61	G61G	comb3
seg7	309	MP1	G	A	R	K	N	95	R95K	comb2
seg7	314	MP1	G	A	V	I	N	97	V97I	comb2
seg7	370	MP1	A	G	I	M	N	115	I115M	comb3
seg7	426	MP1	G	A	R	K	N	134	R134K	comb1
seg7	607	MP1	A	G	G	G	S	194	G194G	comb3
seg8	290	NS1	G	A	R	R	S	88	R88R	comb3
seg8	613	NS2	G	A	M	I	N	19	M19I	comb1
seg8	640	NS1	G	A	S	N	N	205	S205N	comb1

4.7 Figures

Fig. 1 Mechanisms of action of NA inhibitors (such as oseltamivir) and of favipiravir

Favipiravir is converted to Favipiravir-RTP by host cell enzymes and selectively inhibits the activity of the influenza viral RNA polymerase. Reproduced from (Furuta *et al.* 2013)

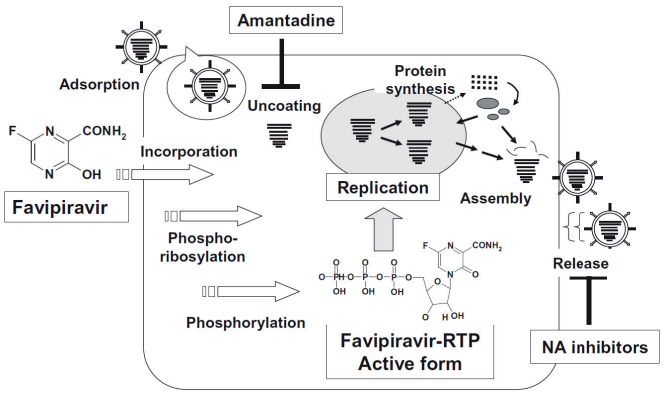


Fig. 2 Evidence of positive selection in the H1N1 genome (reproduced from Foll *et al.* (2014). Plots A and C represent the Bayesian P-values for each SNP in the absence and presence of oseltamivir (the red line is a threshold of P=0.01). The allele trajectories of significant SNPs are plotted in B and D respectively, with estimates of selection coefficients. The eight segments are color coded, and lines are dashed if these represent a second SNP in each segment. Significant SNPs are highlighted on the structure: HA D112N (NS), NA G193G (S), NA H275Y (NS), M1 A41V (NS)

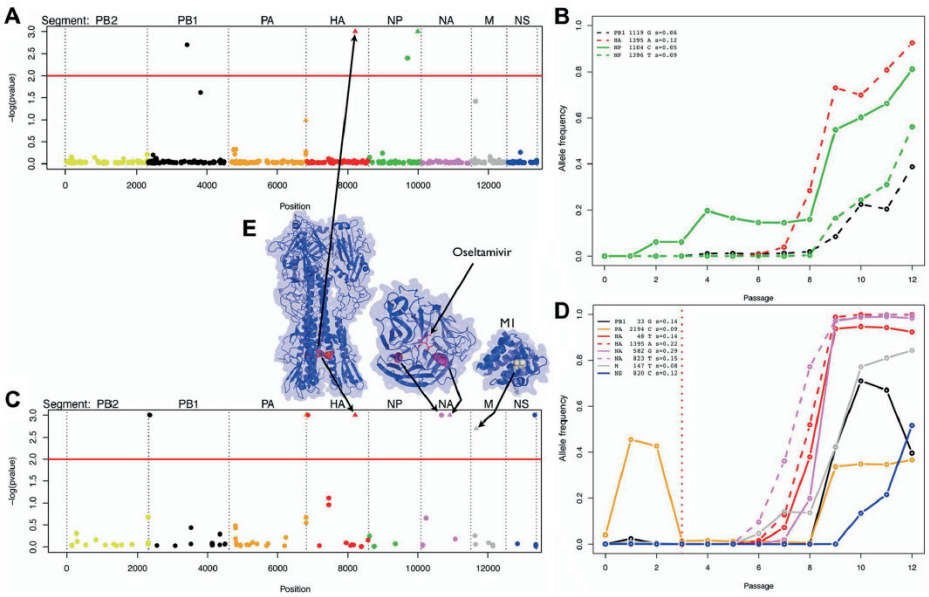


Fig. 3. Histogram of fitness effects (reproduced from Foll *et al.* (2014) in the A) absence and B) presence of oseltamivir. A generalized Pareto distribution (GPD) was fitted to the histogram to quantify the differences between the two distributions.

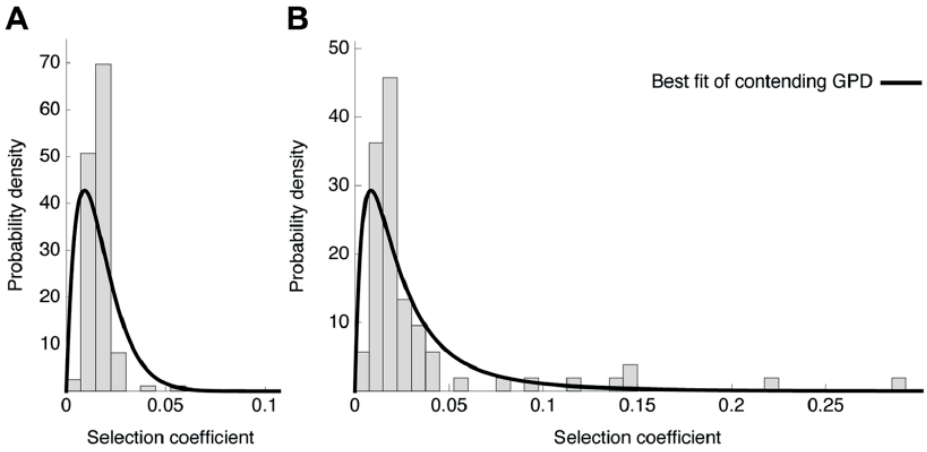


Fig. 4 Experimentally observed pattern of mutation accumulation and effective population size compared to theoretical expectations (reproduced from Bank *et al.* (2016))

Qualitative comparison of accumulation of mutations in the favipiravir population (blue dots) with pattern of mutation accumulation redrawn from Figure 1 of Lynch *et al.* (1993). The horizontal grey dashed line represents the census population size in the original model, which is overlaid by estimates of N_e between passages (grey dots). The dashed vertical line in green represents the transition to meltdown phase.

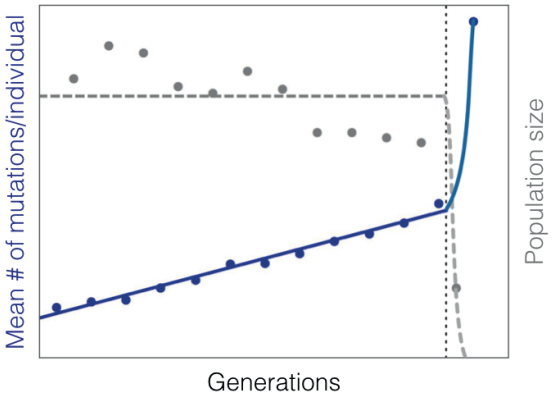


Fig. 5. Changes in absolute and relative growth rates of IAV (reproduced from Bank *et al.* (2016)). Panel A: Absolute growth rate showed a strong negative correlation with imposed drug concentrations, providing further evidence of the effectiveness of the treatment. Panel B: No correlation was observed between the initial population size in each passage and the absolute growth rate. Panel C: In the favipiravir (favi1) treated population, relative growth rates (compared with the parallel no drug control) were consistently negative. Panels D-F: Relative growth rates of additional treatment strategies across passages. Although the growth rate decreased in the first part of each experiment (passages 4-9) as drug concentration (blue line) increased, one of the populations under constant drug concentrations from passage 9 (constA in panel D) and one under withdrawal of treatment from passage 9 (withdrawal in panel E) showed signs of recovery upon the change of treatment.

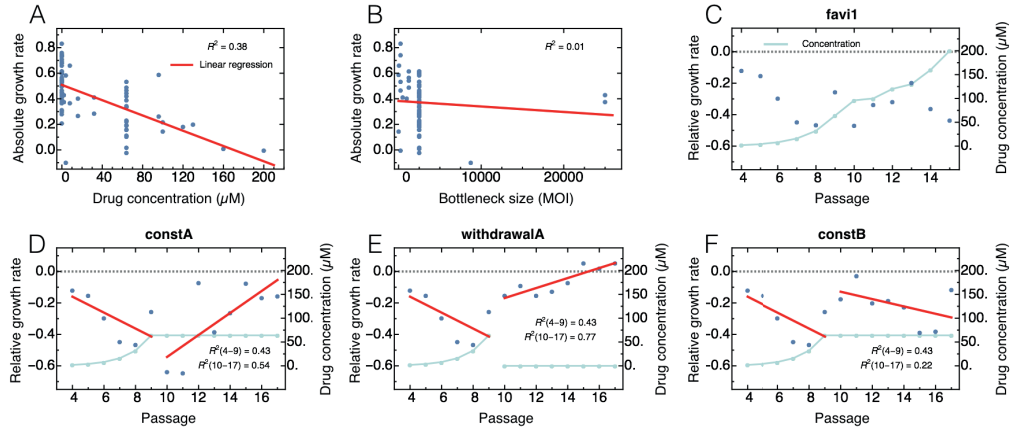


Fig. 6. Experimental set-up

In each of three replicates, influenza A virus was serially passaged in MDCK cells and exposed to increasing concentrations of either oseltamivir only, or of oseltamivir and favipiravir combined, from passage 4 onwards. The multiplicity of infection (MOI) used to seed each passage is shown on the right hand side. ED_{50} represents a 50% effective dose for drug-naïve virus. The MOIs are valid for all replicates except combined drug and oseltamivir replicates 1 passage 10, where an MOI of 0.005 was applied. The passaging was extended to an eleventh passage for oseltamivir replicate 1 only.

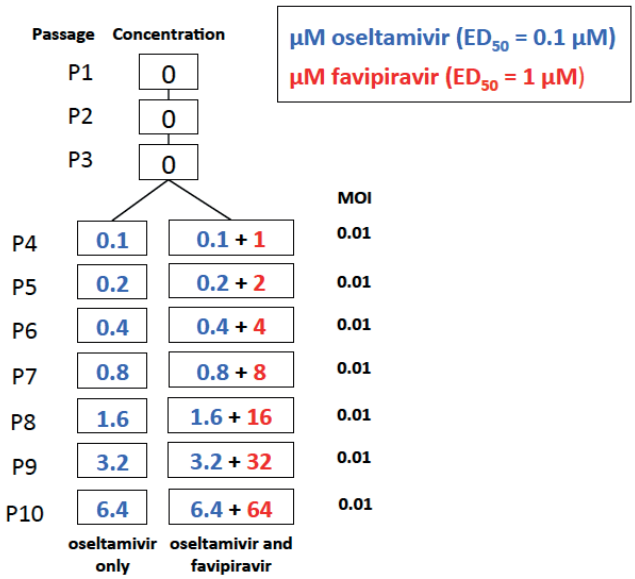


Fig. 7. Genetic diversity

Genetic diversity calculated as the average expected heterozygosity in passages 4 to 10 for the combined drug replicates (red) and the oseltamivir replicates (blue), and in passages 3-15 for the favipiravir (yellow) and no drug (grey) populations.

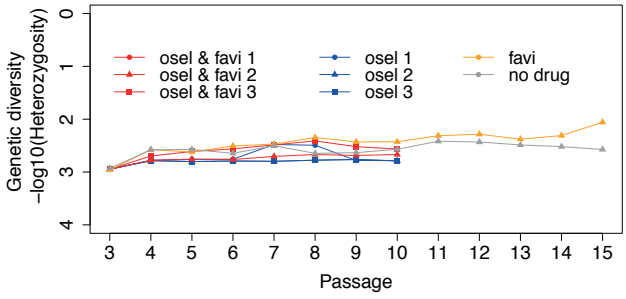


Fig. 8 Site Frequency Spectrum

For each site, counts were randomly (hypergeometrically) downsampled to 100 after filtering (based on frequency > 1% in one of the passages to reduce sequencing error estimated at 1%). The site frequency spectra are shown for combined drug replicate 1 at passage 4 (in blue) and passage 8 (in red). There is a left bias to the SFS reflecting the strength of purifying selection throughout all passages. Counts in the low frequency bins have increased by passage 8 because of the mutagenic impact of favipiravir on the number of segregating mutations.

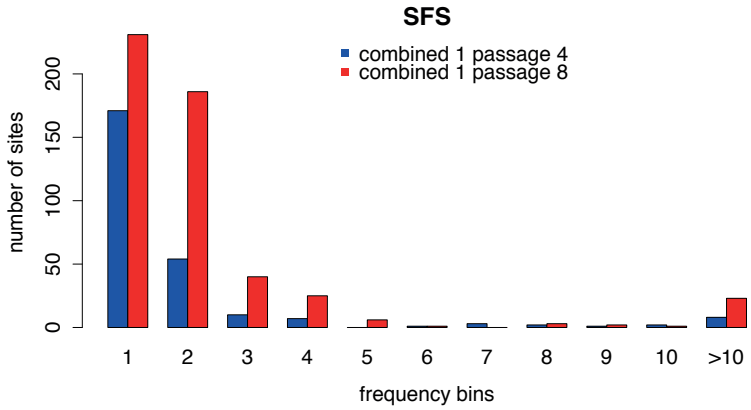
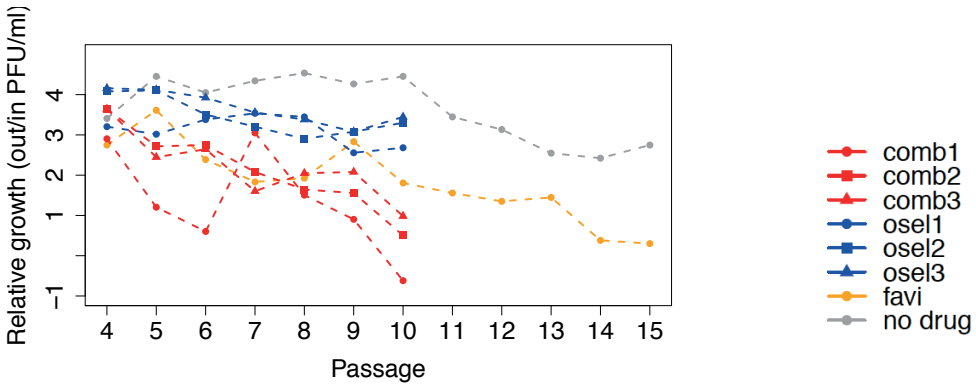


Fig. 9 Relative growth and total number of segregating sites in IAV treated with two antiviral agents. A) Relative growth (a measure of viral fitness) is calculated as $\log_{10}$(output/input PFU) for each passage. A more rapid decline in relative growth was observed in the combined drug replicates (red), than in the favipiravir population (yellow). Relative growth for the oseltamivir replicates (blue) remains relatively stable, although below the level exhibited by the control population (grey). B) The number of segregating mutations represents all sites segregating at more than 1% DAF at each passage. Despite the more rapid decline in relative growth, the number of segregating sites for the combined drug replicates (red) is lower than for the favipiravir population (orange), indicating a lower segregating mutation load. The number of segregating sites observed in the oseltamivir replicates is low, likely owing to selective sweeps around the oseltamivir resistance mutations, as well as to the absence of favipiravir's mutagenic effect.

A.



B.

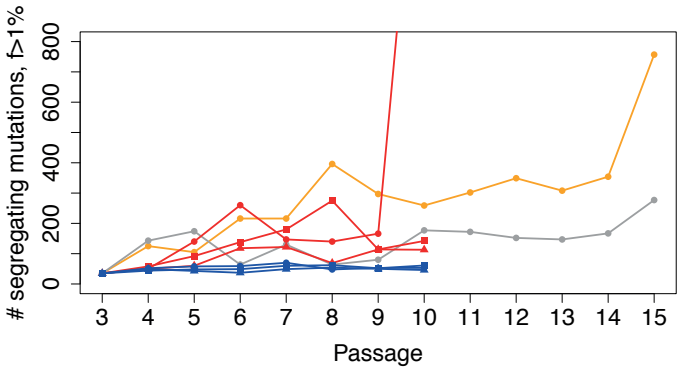


Fig. 10 Effective population sizes and mutation load per individual for combined drug replicates. The mean number of mutations per individual is calculated by dividing the total number of mutations segregating with a DAF >1% in the population by the effective population size. The effective population size is calculated in A-C using the method of (Jorde & Ryman 2007) , whereas in D-F it is calculated as the harmonic mean of population size per generation.

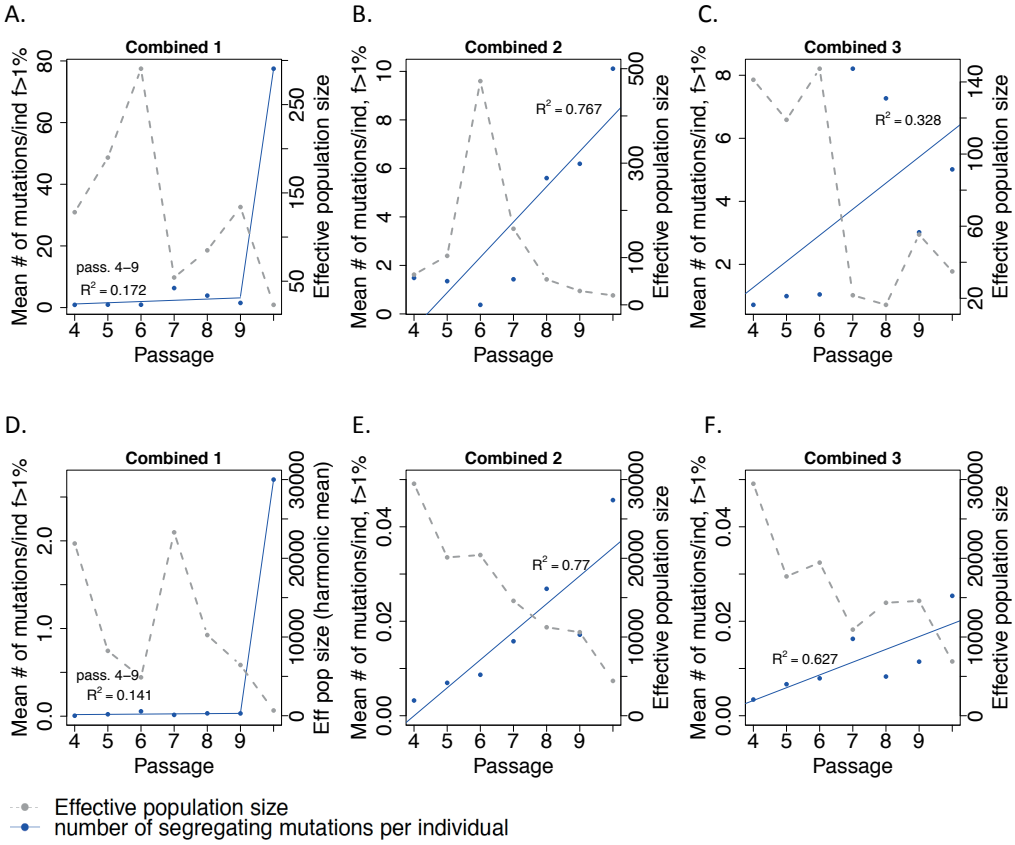


Fig. 11 WF-ABC Global Effective Population sizes. These are calculated from the variance in allele frequencies across 8 time points.

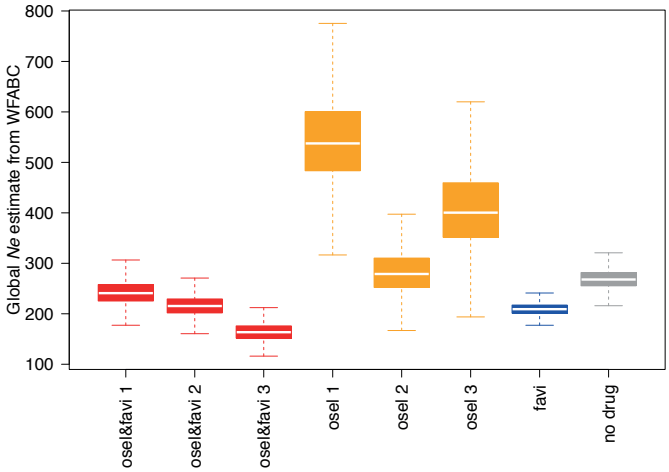


Fig. 12 Changes in relative growth rates for combined drug replicates. The relative growth rate for combined drug replicates per passage was calculated by subtracting the absolute growth rate per generation for oseltamivir replicates (as a control) from the absolute growth rate per generation for the combined drug replicates (shown as blue points with the linear regression in red, left axis). The escalating concentration of the combined drug treatment (in µM favipiravir and 0.1µM oseltamivir) is shown in light blue (right axis). The negative slope of the linear regression indicates a declining rate of relative growth, with no evidence of recovery or adaptation by combined drug populations. The relative growth is significantly different from 0 for replicate 2 (at the 5% level, p=0.0278).

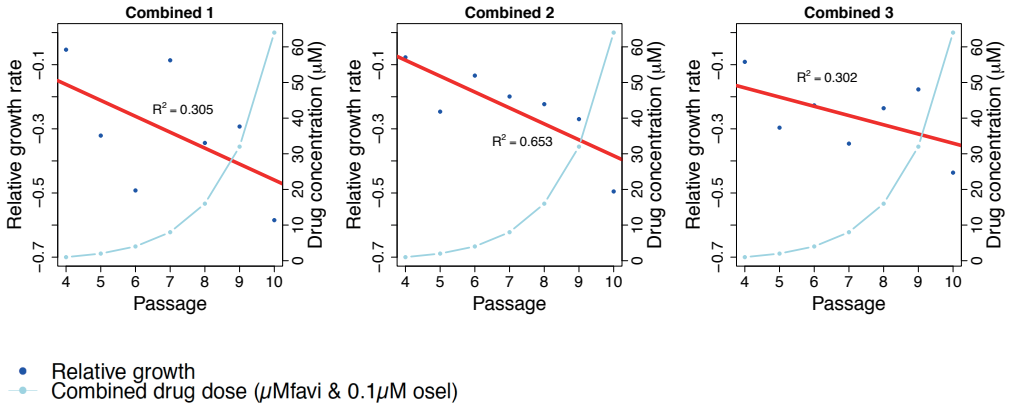


Fig. 13 Putatively beneficial mutations. The contending beneficial mutations (Table 1) are tracked in the oseltamivir replicates (**A-C**), in the combined drug replicates (**D-F**) and in the favipiravir and no drug populations (**G-H**). The key to these mutations is given below, with the NA mutations in red. Clustering is observed amongst these beneficial mutations in combined drug replicates. All other mutations (*i.e.*, arising in only one replicate) are assumed to be neutral or deleterious and are plotted in grey. In combined drug replicates, this class of mutations also shows evidence of genetic hitchhiking either with the resistance mutations or with HA E78G (a ubiquitous beneficial mutation in all oseltamivir and combined drug replicates except for combined drug replicate 2)(**D-F**). In **G-H** the beneficial mutations are tracked in the favipiravir and no drug control populations, with longer trajectories (weaker selection coefficients) and less evidence of clustering.

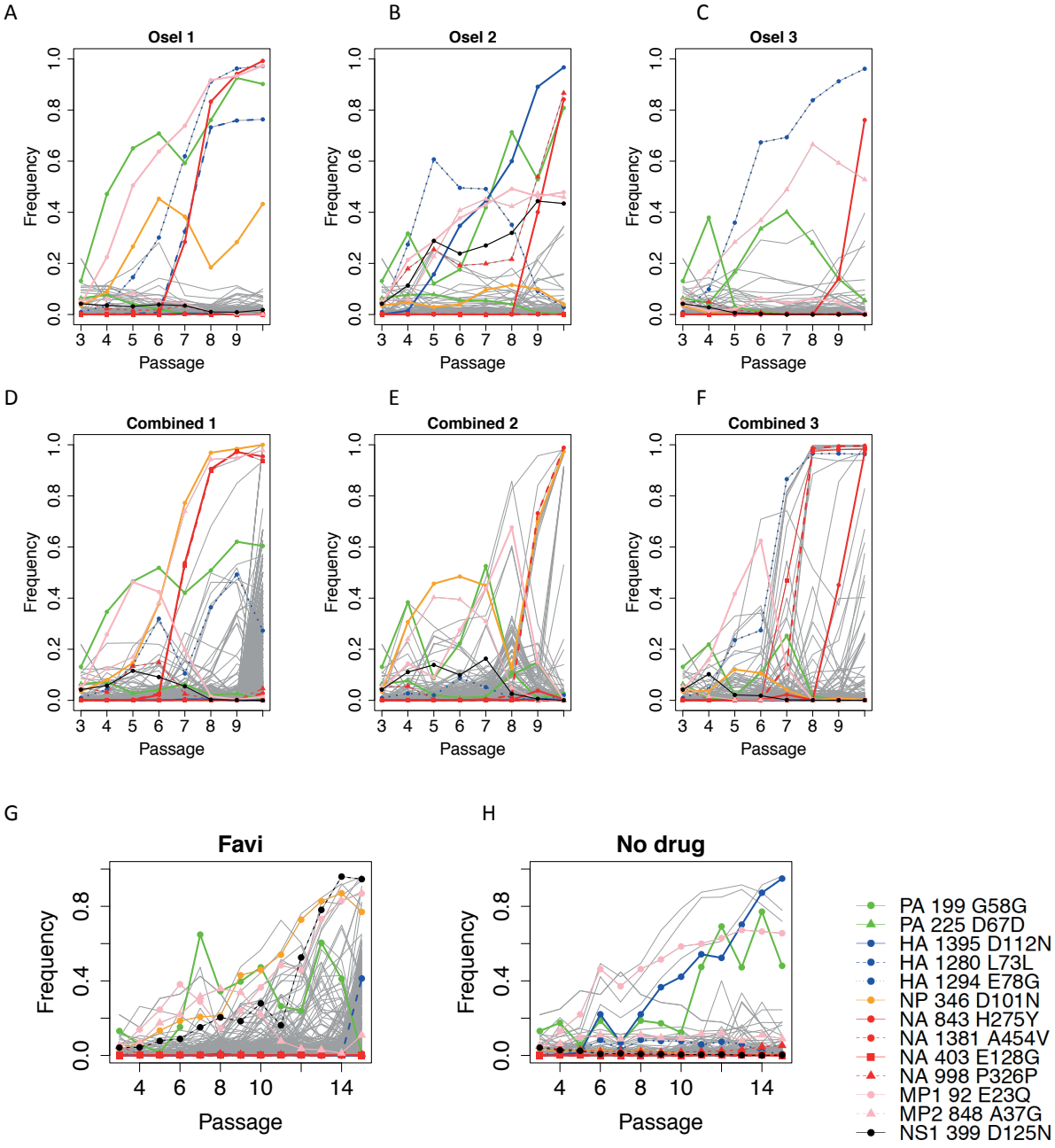
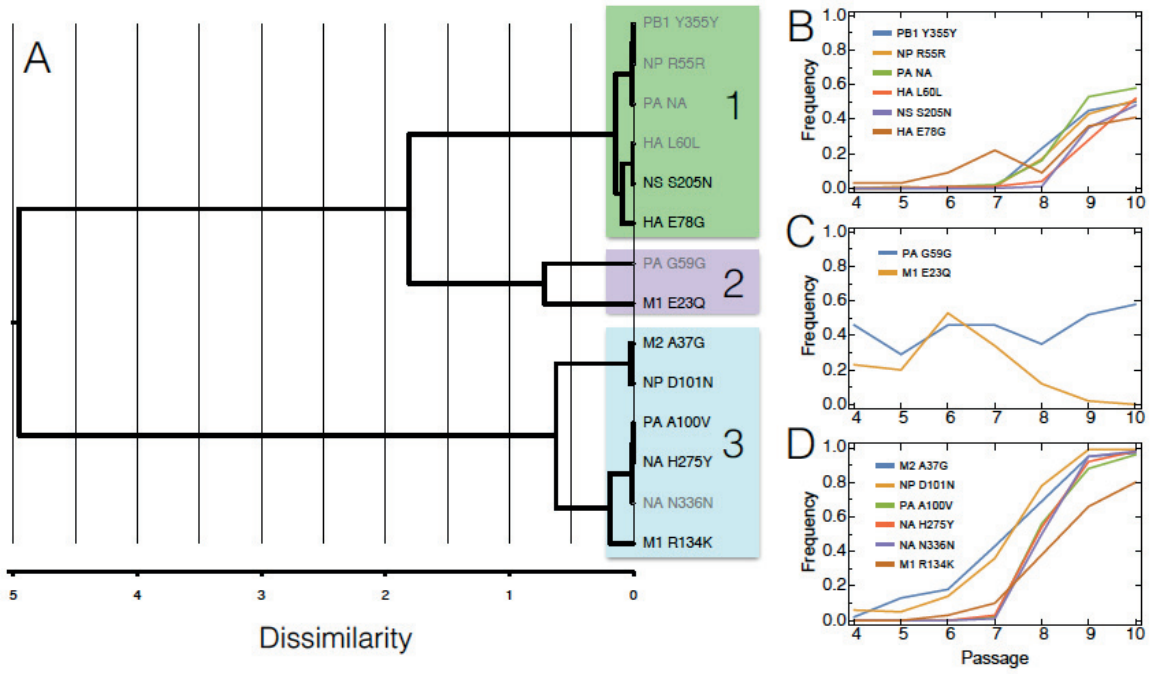
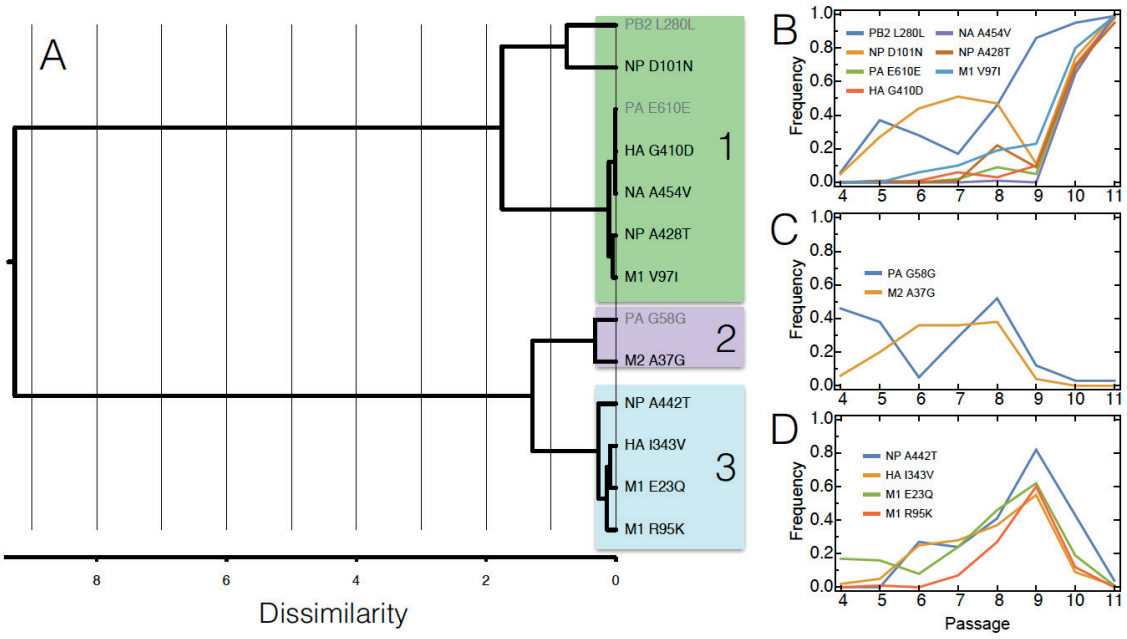


Fig. 15 Hierarchical cluster analysis. Ward's minimum variance criterion (Ward Jr 1963) was used to cluster allele-frequency trajectories. The dissimilarity distances are shown in panels A and the details of the clusters in panels B-D for replicates 1 and 2, and in panels B-G for replicate 3. We observe hitchhiking patterns suggesting that either NA mutations (H275Y, A454V or E128G) or HA E78G sweep other beneficial and neutral/deleterious mutations to fixation. Other cluster groups containing contending beneficials (Table 3) do not fix. On panel A, synonymous mutations are shown in grey font and non-synonymous mutations in black. Mutations marked NA did not have a known amino acid equivalent using H3 numbering.

Replicate 1



Replicate 2



Replicate 3

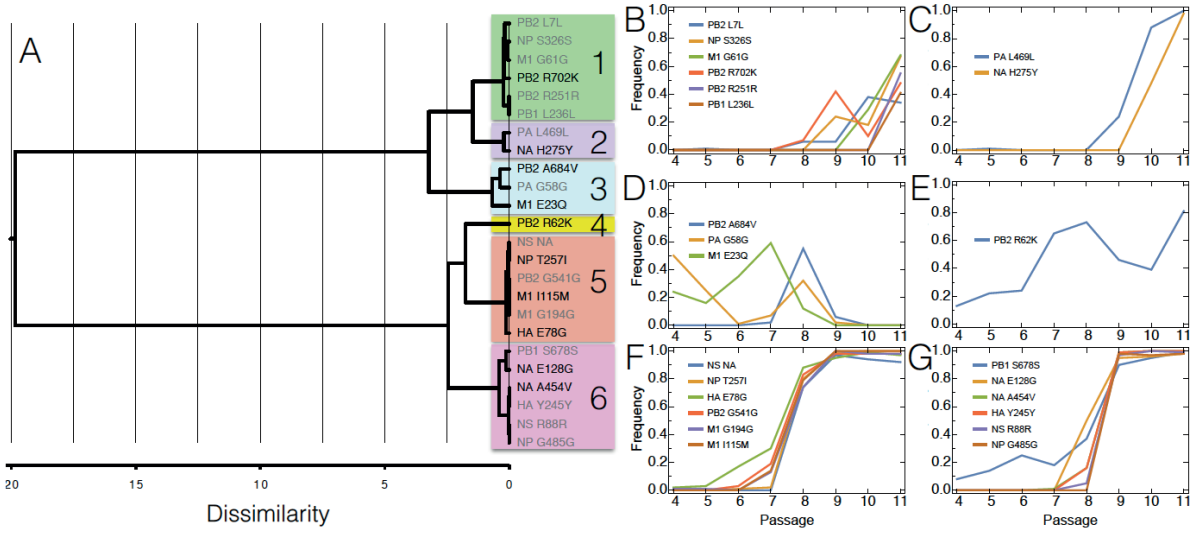
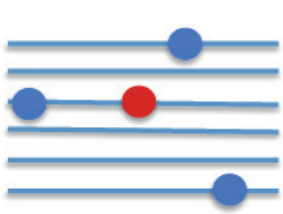
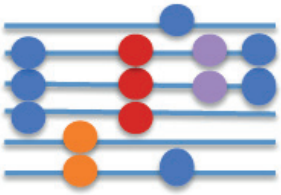


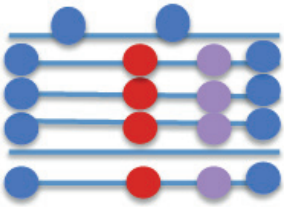
Fig. 15 Schematic of evolutionary dynamics in combined drug populations



- The starting point is mutation-selection-drift balance (strong purifying selection removes new deleterious mutations)
- Favipiravir increases the influx of new neutral and deleterious mutations (blue)
- New strongly selected oseltamivir resistance mutation (red) arises



- Beneficial resistance mutation increases rapidly in frequency
- Some beneficial mutations (purple) are generated *de novo* by mutation or reshuffled by segment reassortment onto the same genetic background as the strongly selected resistance mutation, and hitchhike to fixation
- Deleterious mutations also hitchhike to fixation with the strongly selected resistance mutation, before purifying selection can remove them
- Competing weakly selected benefical mutations (orange) arise on different haplotype backgrounds



- The resistance mutation, and the hitchhiked beneficials and deleterious mutations fix
- Competing weakly selected beneficial mutations (orange) on haplotypes not carrying the resistance mutation, are lost (Hill Robertson interference)
- Limited scope for WSHRI because of presence of the driver resistance mutation

References

Alexander HK, Martin G, Martin OY, Bonhoeffer S (2014) Evolutionary rescue: linking theory for conservation and medicine. *Evol Appl* **7**, 1161-1179.

Bank C, Hietpas RT, Wong A, Bolon DN, Jensen JD (2014) A bayesian MCMC approach to assess the complete distribution of fitness effects of new mutations: uncovering the potential for adaptive walks in challenging environments. *Genetics* **196**, 841-852.

Bank C, Renzette N, Liu P*, et al.* (2016) An experimental evaluation of drug-induced mutational meltdown as an antiviral treatment strategy. *Evolution*.

Bao Y, Bolotov P, Dernovoy D*, et al.* (2008) The influenza virus resource at the National Center for Biotechnology Information. *J Virol* **82**, 596-601.

Baranovich T, Wong SS, Armstrong J*, et al.* (2013) T-705 (favipiravir) induces lethal mutagenesis in influenza A H1N1 viruses in vitro. *J Virol* **87**, 3741-3751.

Barton NH (2010) Genetic linkage and natural selection. *Philos Trans R Soc Lond B Biol Sci* **365**, 2559-2569.

Bergstrom CT, Pritchard J (1998) Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. *Genetics* **149**, 2135-2146.

Bloom JD, Gong LI, Baltimore D (2010) Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science* **328**, 1272-1275.

Bouvier NM, Rahmat S, Pica N (2012) Enhanced mammalian transmissibility of seasonal influenza A/H1N1 viruses encoding an oseltamivir-resistant neuraminidase. *J Virol* **86**, 7268-7279.

Bull JJ, Sanjuán R, Wilke CO (2007) Theory of lethal mutagenesis for viruses. *J Virol* **81**, 2930-2939.

Butler J, Hooper KA, Petrie S*, et al.* (2014) Estimating the fitness advantage conferred by permissive neuraminidase mutations in recent oseltamivir-resistant A(H1N1)pdm09 influenza viruses. *PLoS Pathog* **10**, e1004065.

Charlesworth B (2009) Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**, 195-205.

Charlesworth B, Charlesworth D (1997) Rapid fixation of deleterious alleles can be caused by Muller's ratchet. *Genet Res* **70**, 63-73.

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289-1303.

Cheung PP, Watson SJ, Choy KT*, et al.* (2014) Generation and characterization of influenza A viruses with altered polymerase fidelity. *Nat Commun* **5**, 4794.

Collins PJ, Haire LF, Lin YP*, et al.* (2008) Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants. *Nature* **453**, 1258-1261.

Daniels RS, Downie JC, Hay AJ*, et al.* (1985) Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell* **40**, 431-439.

Domingo E, Mas A, Yuste E*, et al.* (2001) Virus population dynamics, fitness variations and the control of viral disease: an update. *Prog Drug Res* **57**, 77-115.

Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465-523.

Eigen M (2002) Error catastrophe and antiviral strategy. *Proc Natl Acad Sci U S A* **99**, 13374-13376.

Ewens WJ (1967) The probability of survival of a mutant. *Heredity (Edinb)* **22**, 307-310.

Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. *Nat Rev Genet* **8**, 610-618.

Felsenstein J (1974) The evolutionary advantage of recombination. *Genetics* **78**, 737-756.

Fisher RA. The genetical theory of natural selection. Oxford, UK: Oxford University Press; 1930.

Foll M, Poh YP, Renzette N*, et al.* (2014) Influenza virus drug resistance: a time-sampled population genetics perspective. *PLoS Genet* **10**, e1004185.

Furuta Y, Gowen BB, Takahashi K*, et al.* (2013) Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antiviral Res* **100**, 446-454.

Ghedin E, Holmes EC, DePasse JV*, et al.* (2012) Presence of oseltamivir-resistant pandemic A/H1N1 minor variants before drug therapy with subsequent selection and transmission. *J Infect Dis* **206**, 1504-1511.
Ginting TE, Shinya K, Kyan Y*, et al.* (2012) Amino acid changes in hemagglutinin contribute to the replication of oseltamivir-resistant H1N1 influenza viruses. *J Virol* **86**, 121-127.
Gordo I, Charlesworth B (2000a) On the speed of Muller's ratchet. *Genetics* **156**, 2137-2140.
Gordo I, Charlesworth B (2000b) The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* **154**, 1379-1387.
Gordo I, Charlesworth B (2001) The speed of Muller's ratchet with background selection, and the degeneration of Y chromosomes. *Genet Res* **78**, 149-161.
Grande-Pérez A, Sierra S, Castro MG, Domingo E, Lowenstein PR (2002) Molecular indetermination in the transition to error catastrophe: systematic elimination of lymphocytic choriomeningitis virus through mutagenesis does not correlate linearly with large increases in mutant spectrum complexity. *Proc Natl Acad Sci U S A* **99**, 12938-12943.
Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. Science. 2004;303(5656):327-32.
Haigh J (1978) The accumulation of deleterious genes in a population--Muller's Ratchet. *Theor Popul Biol* **14**, 251-267.
Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. *Genet Res* **8**, 269-294.
Holmes EC (2003) Error thresholds and the constraints to RNA virus evolution. *Trends Microbiol* **11**, 543-546.
Irwin KK, Laurent S, Matuszewski S*, et al.* (2016) On the importance of skewed offspring distributions and background selection in virus population genetics. *Heredity (Edinb)*.
Ives JA, Carr JA, Mendel DB*, et al.* (2002) The H274Y mutation in the influenza A/H1N1 neuraminidase active site following oseltamivir phosphate treatment leave virus severely compromised both in vitro and in vivo. *Antiviral Res* **55**, 307-317.
Jin H, Leser GP, Zhang J, Lamb RA (1997) Influenza virus hemagglutinin and neuraminidase cytoplasmic tails control particle shape. *EMBO J* **16**, 1236-1247.
Jorde PE, Ryman N (2007) Unbiased estimator for genetic drift and effective population size. *Genetics* **177**, 927-935.
Loeb LA, Essigmann JM, Kazazi F*, et al.* (1999) Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc Natl Acad Sci U S A* **96**, 1492-1497.
Loeb LA, Mullins JI (2000) Lethal mutagenesis of HIV by mutagenic ribonucleoside analogs. *AIDS Res Hum Retroviruses* **16**, 1-3.
Lynch M, Bürger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual populations. *J Hered* **84**, 339-344.
McVean GA, Charlesworth B (2000) The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* **155**, 929-944.
Meijer A, Rebelo-de-Andrade H, Correia V*, et al.* (2014) Global update on the susceptibility of human influenza viruses to neuraminidase inhibitors, 2012-2013. *Antiviral Res* **110**, 31-41.
Mitchison DA (2012) Prevention of drug resistance by combined drug treatment of tuberculosis. *Handb Exp Pharmacol*, 87-98.
Moscona A (2005) Oseltamivir resistance--disabling our influenza defenses. *N Engl J Med* **353**, 2633-2636.
Moscona A (2009) Global transmission of oseltamivir-resistant influenza. *N Engl J Med* **360**, 953-956.
Muller HJ. Some genetic aspects of sex. Am Nat. 1932;66:118-38.
Muller HJ (1964) The relation of recombination to mutational advance. *Mutat Res* **106**, 2-9.
Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A* **70**, 3321-3323.
Noton SL, Medcalf E, Fisher D*, et al.* (2007) Identification of the domains of the influenza A virus M1 matrix protein required for NP binding, oligomerization and incorporation into virions. *J Gen Virol* **88**, 2280-2290.

Noton SL, Simpson-Holley M, Medcalf E*, et al.* (2009) Studies of an influenza A virus temperature-sensitive mutant identify a late role for NP in the formation of infectious virions. *J Virol* **83**, 562-571.

Pariente N, Sierra S, Lowenstein PR, Domingo E (2001) Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. *J Virol* **75**, 9723-9730.

Reed ML, Yen HL, DuBois RM*, et al.* (2009) Amino acid residues in the fusion peptide pocket regulate the pH of activation of the H5N1 influenza virus hemagglutinin protein. *J Virol* **83**, 3568-3580.

Rossman JS, Lamb RA (2011) Influenza virus assembly and budding. *Virology* **411**, 229-236.

Sierra S, Dávila M, Lowenstein PR, Domingo E (2000) Response of foot-and-mouth disease virus to increased mutagenesis: influence of viral load and fitness in loss of infectivity. *J Virol* **74**, 8316-8323.

Smee DF, Tarbet EB, Furuta Y, Morrey JD, Barnard DL (2013) Synergistic combinations of favipiravir and oseltamivir against wild-type pandemic and oseltamivir-resistant influenza A virus infections in mice. *Future Virol* **8**, 1085-1094.

Smith JM, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genet Res* **23**, 23-35.

Takashita E, Meijer A, Lackenby A*, et al.* (2015) Global update on the susceptibility of human influenza viruses to neuraminidase inhibitors, 2013-2014. *Antiviral Res* **117**, 27-38.

Thoennes S, Li ZN, Lee BJ*, et al.* (2008) Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion. *Virology* **370**, 403-414.

Varghese JN, Smith PW, Sollis SL*, et al.* (1998) Drug design against a shifting target: a structural basis for resistance to inhibitors in a variant of influenza virus neuraminidase. *Structure* **6**, 735-746.

Ward Jr J (1963) Hierarchical Grouping to Optimize an Objective Function. *J.Am. Stat. Assoc.* **58: 236–244.**

Whitlock MC, Bourguet D (2000) Factors affecting the genetic load in Drosophila: synergistic epistasis and correlations among fitness components. *Evolution* **54**, 1654-1660.

Wright S (1931) Evolution in Mendelian Populations. *Genetics* **16**, 97-159.

Wylie CS, Shakhnovich EI (2012) Mutation induced extinction in finite populations: lethal mutagenesis and lethal isolation. *PLoS Comput Biol* **8**, e1002609.

References to the paper : Two sides of the same coin : A population genetics perspective on lethal mutagenesis and mutational meltdown

Achaz, G., A. Rodriguez-Verdugo, B.S. Gaut, and O. Tenaillon, 2014. The reproducibility of adaptation in the light of experimental evolution with whole genome sequencing. In Ecological Genomics, ed. C. R. Landry, N. Aubin-Horth, pp. 211–31. Dordrect, Neth.: Springer.

Airaksinen, A., N. Pariente, L. Menendez-Arias and E. Domingo, 2003. Curing of foot-and-mouth disease virus from persistently infected cells by ribavirin involves enhanced mutagenesis. Virology 311:339-349.

Allen, J. M., J. E. Light, M. A. Perotti, H. R. Braig, and D. L. Reed, 2009. Mutational meltdown in primary endosymbionts: Selection limits Muller's ratchet. PLoS ONE 4:1–9.

Anderson, J. P., R. Daifuku, and L. A. Loeb, 2004. Viral error catastrophe by mutagenic nucleosides. Annual Reviews of Microbiology 58:183–205.

Antoneli, F., F. Bosco, D. Castro, and L. M. Janini, 2013. Virus replication as a phenotypic version of polynucleotide evolution. Bulletin of mathematical biology, 75:602--628.

Bank, C., R. T. Hietpas, A. Wong, D. N. Bolon, and J. D. Jensen, 2014. A Bayesian MCMC approach to assess the complete distribution of fitness effects of new mutations: Uncovering the potential for adaptive walks in challenging environments. Genetics 196:841–852.

Bank, C., N. Renzette, P. Liu, S. Matuszewski, H. Shim, M. Foll, D. N. A. Bolon, K. B. Zeldovich, T. F. Kowalik, R. W. Finberg, J. P. Wang, and J. D. Jensen, 2016. An experimental evaluation of drug-induced mutational meltdown as an antiviral treatment strategy. Evolution 70: 2470-84.

Biebricher, C. K. and M. Eigen, 2005. The error threshold. Virus Research 107:117 – 127.

Bull, J. J., R. Sanjuán, and C. O. Wilke, 2007. Theory of lethal mutagenesis for viruses. Journal of Virology 81:2930–2939.

Bonhoeffer, S. and P. F. Stadler, 1993. Error thresholds on correlated fitness landscapes. Journal of Theoretical Biology 164:359–372.

Bull, J. J., P. Joyce, E. Gladstone, and I. J. Molineux, 2013. Empirical complexities in the genetic foundations of lethal mutagenesis. Genetics 195:541--552.

Bürger, R., 1989. Linkage and the maintenance of heritable variation by mutation-selection balance. Genetics 121:175-184.

Bürger, R., 1998. Mathematical properties of mutation-selection models. Genetica 102:279.

Charlesworth, B., 2009. Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics 10:195–205.

Chen, P. and E. I. Shakhnovich, 2009. Lethal mutagenesis in viruses and bacteria. Genetics 183:639–650.

Cooper, T. F., E. A. Ostrowski, and M. Travisano, 2007. A negative relationship between mutation pleiotropy and fitness effect in yeast. Evolution 61: 1495–1499.

Crotty, S., C. E. Cameron and R. Andino, 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. Proceedings of the National Academy of Sciences USA 98:6895-6900.

Crow, J. F., 1954. Statistics and Mathematics in Biology. Iowa State Univ. Press, Ames, Iowa, USA.

Eigen, M., 1971. Self-organization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523.

Eigen, M., 2002. Error catastrophe and antiviral strategy. Proceedings of the National Academy of Scences USA 99:13374–13376.

Eigen, M. and P. Schuster, 1979. The Hypercycle: A Principle of Natural Self-Organization. Springer-Verlag, Berlin.

Eyre-Walker, A. and P. D. Keightley, 2007. The distribution of fitness effects of new mutations. Nature Reviews Genetics 8:610–618.

Felsenstein, J., 1974. The evolutionary advantage of recombination. Genetics 78:737–756.

Fisher, R., 1930. The genetical theory of natural selection. Clarendon Press, Oxford, UK.

Frankham, R., 1995. Effective population size/adult population size ratios in wildlife: a review. Geneti-

cal Research 66:95–107.
Gabriel, W. and R. Bürger, 1992. Survival of small populations under demographic stochasticity. Theoretical Population Biology 41:44 – 71.
Gabriel, W. and R. Bürger, 1994. Extinction risk by mutational meltdown: synergistic effects between population regulation and genetic drift. Pp. 69–84, *in* Conservation Genetics. Birkhäuser Verlag, Basel.
Gabriel, W., M. Lynch, and R. Bürger, 1993. Muller's ratchet and mutational meltdowns. Evolution 47:1744–1757.
Grande-Pérez, A., G. Gómez-Mariano, P. R. Lowenstein, and E. Domingo, 2005. Mutagenesis-induced, large fitness variations with an invariant arenavirus consensus genomic nucleotide sequence. Journal of Virology 79: 10451-10459.
Gerrish, P. J., A. Colato, and P. D. Sniegowski, 2013 Genomic mutation rates that neutralize adaptive evolution and natural selection. J. R. Soc. Interface 10(85): 20130329.
Grenfell, B. T., O. G. Pybus, J. R. Gog, J. L. Wood, J. M. Daly, J. A. Mumford, and E. C. Holmes, 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303:327–332.
Haldane, J. B. S., 1932. The causes of evolution. Harper & Row, New York
Haldane, J. B. S., 1937. The effect of variation on fitness. The American Naturalist 71:337– 349.
Hermisson, J., O. Redner, H. Wagner and E. Baake, 2002. Mutation-Selection Balance: Ancestry, Load, and Maximum Principle. Theoretical Population Biology 62: 9-46.
Hietpas, R. T., C. Bank, J. D. Jensen, and D. N. A. Bolon, 2013. Shifting fitness landscapes in response to altered environments. Evolution 67:3512–3522.
Hietpas, R. T., J. D. Jensen, and D. N. A. Bolon, 2011. Experimental illumination of a fitness landscape. Proceedings of the National Academy of Sciences USA 108:7896–7901.
Hill, W. G. and A. Robertson, 1966. The effect of linkage on limits to artificial selection. Genetical Research 8:269–294.
Irwin, K. K., S. Laurent, S. Matuszewski, S. Vuilleumier, L. Ormond, H. Shim, C. Bank, and J. D. Jensen, 2016. On the importance of skewed offspring distributions and background selection in virus population genetics. Heredity 117:393--399
Kimura, M. and J. F. Crow, 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725-738.
Kimura, M. and T. Maruyama, 1966. The mutational load with epistatic gene interactions in fitness. Genetics 54:1337–1351.
Lande, R., 1994. Risk of population extinction from fixation of new deleterious mutations. Evolution 48:1460–1469.
Lande, R., 1998. Risk of population extinction from fixation of deleterious and reverse mutations. Genetica 102:21–27.
Loeb, L. A., J. M. Essigmann, F. Kazazi, J. Zhang, K. D. Rose, and J. I. Mullins, 1999. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proceedings of the National Academy of Sciences USA 96:1492–1497.
Lynch, M., R. Bürger, D. Butcher, and W. Gabriel, 1993. The mutational meltdown in asexual populations. Journal of Heredity 84:339–344.
Lynch, M., J. Conery, and R. Burger, 1995a. Mutation accumulation and the extinction of small populations. American Naturalist 146(4):489–518.
Lynch, M., J. Conery, and R. Burger , 1995b. Mutational meltdowns in sexual populations. Evolution 49(6):1067–1080.
Lynch, M. and W. Gabriel, 1990. Mutation load and the survival of small populations. Evolution 44:1725–1737.
Manrubia, S. C., E. Domingo, and E. Lázaro, 2010. Pathways to extinction: beyond the error threshold. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365: 1943-1952.
Martin, G. and S. Gandon, 2010. Lethal mutagenesis and evolutionary epidemiology. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365:1953–1963.

Martin, G., and T. Lenormand, 2006a, A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60: 893–907.
Maynard Smith, J., 1970. Natural selection and the concept of a protein space. Nature 225:563–564.
McVean, G. A. T. and B. Charlesworth, 2000. The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics 155:929–944.
Muller, H., 1964. The relation of recombination to mutational advance. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 1:2 – 9.
Neher, R. A. and O. Hallatschek, 2013. Genealogies of rapidly adapting populations. Proceedings of the National Academy of Sciences USA 110:437–442.
Nowak, M. and R. May, 2000. Virus Dynamics: Mathematical Principles of Immunology and Virology: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford, UK.
Ochoa, G. and I. Harvey, 1998. Recombination and error thresholds in finite populations. Foundations of Genetic Algorithms 5:245–264.
Pariente, N., S. Sierra, P. R. Lowenstein and E. Domingo, 2001. Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. Journal of Virology 75:9723-9730.
Pennings, P. S., S. Kryazhimskiy, and J. Wakeley, 2014. Loss and recovery of genetic diversity in adapting populations of HIV. PLoS Genetics 10:e1004000.
Rowe, G. and T. J. C. Beebee, 2003. Population on the verge of a mutational meltdown? Fitness costs of genetic load for an amphibian in the wild. Evolution 57:177–181.
Shoubridge, E. A. and T. Wai, 2008. Sidestepping mutational meltdown. Science 319:914– 915.
Sniegowski P. D., P. J. Gerrish and R. E. Lenski, 1997. Evolution of high mutation rates in experimental populations of *Escherichia coli*. Nature 387: 703–705.
Springman, R., T. Keller, I. J. Molineux, and J. J. Bull, 2010 Evolution at a high imposed mutation rate: adaptation obscures the load in phage T7. Genetics 184: 221–232.
Stich, M., E. Lázaro and S. C. Manrubia, 2010. Phenotypic effect of mutations in evolving populations of RNA molecules. BMC Evolutionary Biology 10:46.
Swetina, J. and P. Schuster, 1982. Self-replication with errors: A model for polynucleotide replication. Biophysical Chemistry 16:329–345.
Tarazona, P., 1992. Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin-glass models. Physical Review A 45:6038–6050.
Tenaillon, O., 2014. The utility of Fisher's geometric model in evolutionary genetics. Annual review of ecology, evolution, and systematics 45:179-201.
Whitlock, M. C. and D. Bourguet, 2000. Factors affecting the genetic load in drosophila: synergistic epistasis and correlations among fitness components. Evolution 54:1654–1660.
Willi, Y., 2013. Mutational meltdown in selfing Arabidopsis lyrata. Evolution 67:806–815.
Woodruff, R.C., 2013. An extreme test of mutational meltdown shows mutational firm up instead. Genetica 141: 185-88.
Wright, S., 1931. Evolution in Mendelian populations. Genetics 16:97–159.
Wylie, C. S. and E. I. Shakhnovich, 2011. A biophysical protein folding model accounts for most mutational fitness effects in viruses. Proceedings of the National Academy of Sciences USA 108:9916–9921.
Wylie, C. S. and E. I. Shakhnovich , 2012. Mutation induced extinction in finite populations: Lethal mutagenesis and lethal isolation. PLoS Computational Biology 8:1–10.
Zeyl, C., M. Mizesko, and J. De Visser, 2001. Mutational meltdown in laboratory yeast populations. Evolution 55:909–917.

Chapter 5 Conclusion

In the first chapter we developed an ABC approach to estimate allele age, selection strength, and starting frequency for fixed mutations using single population, single time-point datasets. Several haplotype-based estimators indeed exist for inferring the age of low or intermediate segregating beneficial mutations, but not fixed beneficial mutations. We show that we can differentiate between different orders of magnitude of the selection coefficient *s*, between old and young sweeps, and between *de novo/*rare and common starting frequencies.

Our method fits a useful niche relevant to environmental studies for understanding the mode and tempo of adaptation in relation to rapid environmental change. These methods can be used to estimate the age of variants underpinning cryptic coloration resulting from environmental changes since the last ice age, for example. In chapter 2, we apply our method to make such estimation in mice sampled from the Nebraska Sand Hills, and estimate a recent, strong selective sweep that is consistent with their formation during the last Ice Age. In chapter 3, the evolution of blanched coloration in White Sands lizards is examined more generally. We incorporate the differing demographic histories and dominance effects of two species into our method to estimate the age and strength of selected beneficial mutations. In both examples, our results suggest that in the populations considered here, selection on *de novo* mutations that post-date environmental shifts is the dominant mode of adaptation – and that adaptation is thus mutation-limited, without a reservoir of standing variation for natural selection to act on.

In chapter 4, we extend this work to a clinical setting, examining the combined effect *in vitro* of two drugs with different mechanisms of action, favipiravir and oseltmamivir, on experimental influenza virus populations. Thus, in this instance, the selective pressures is man-made – though in many ways the application of drug is analagous to the shift in selective pressure experienced by the populations discussed above post-glaciation. However, in this case the viral populations are unable to adapt to the new pressure. The results fit a classic model of mutational meltdown, with a rapidly escalating mutational load until a sharp point of extinction, as previously described in favipiravir treated populations. However, we observe a more rapid decline in these combined drug populations than in populations treated with favipiravir alone, despite the rapid fixation of oseltamivir resistance mutations and a lower segregating mutation load than in favipiravir populations. The processes behind the synergistic combined effect of favipiravir and oseltamivir were previously unknown. Here we describe a complex interplay of mutation, genetic drift and selection. We observe that the mutagenic effect of favipiravir enables the virus to rapidly explore sequence space and generate new oseltamivir resistance mutations, in addition to the widespread H275Y mutation, and acts to cluster some beneficial mutations on the same haplotype. We find evidence of small effective population sizes supporting the action of Muller's ratchet in addition to the impact of the escalating mutation rate on deleterious mutation load.

Importantly, we show that strong selective sweeps around the oseltamivir resistance mutations hitchhike not only beneficial mutations but also deleterious mutations to fixation. Thus our results usefully inform clinical strategies of both the risks (increased transmission of new and established oseltamivir resistance mutations) and benefits (faster extinction) of using a combined drug protocol. Further studies are underway to explore the effects of using a combined drug protocol with a high and low dosage of favipiravir, compared to the same protocol with favipiravir alone, in order to explore these mechanisms further and identify the minimum dosage required to induce extinction.

In this dissertation, I develop population genetic methods to estimate allele age and selection parameters that serve to address fundamental questions in evolutionary biology. Whether in an empirical or a clinical context, similar questions regarding the mode and tempo of adaptation arise. The focus of my PhD has been on developing tools and approaches to help answer some of these questions. In the case of rapid environmental shifts, estimating allele age and selection parameters allows us to understand the nature of the adaptive process, which will become increasingly relevant in an era of intensifying climate change. In a clinical context, where drug resistance is a pressing public health issue, understanding the evolutionary dynamics of H1N1 virus populations will help to manage and optimise the risks and benefits of different drug protocols. Thus, this work fits specific, previously-void niches and contributes towards the goal of population genetics of studying evolutionary change.

Curriculum Vitae

LOUISE ORMOND

Ecole Polytechnique Federale de Lausanne

louise.ormond@epfl.ch

http://jensenlab.epfl.ch/

ADDRESS Rue Louis-de-Savoie 90, Morges 1110, Switzerland

TEL +41-788168684

NATIONALITY British

SUMMARY I am a PhD student in population genetics in the laboratory of Prof Jeffrey Jensen at the Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland. My recent work focuses on the development of single time point Approximate Bayesian Computation (ABC) methods for estimating background selection from polymorphism data. In parallel, I am analysing time series experimental evolution data on the effect of combining two drugs (oseltamivir and favipiravir) in treating influenza virus.

My background includes 10 years' professional experience spanning the private and the non-profit sectors. In the non-profit sector, I have worked on large-scale public health programmes in Africa and Asia. This experience originally spurred my interest in population genetics in relation to disease and host-pathogen co-evolution.

EDUCATION

ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE, Switzerland 2012-
PhD in Population Genetics (completion expected Feb 2017)
Adviser: Prof Jeffrey D. Jensen

- Development of ABC methods for the inference of background selection and of the shape of the Distribution of Fitness Effects (DFE), and application of these methods to polymorphism data from a range of organisms (from chimpanzees to human cytomegalovirus (HCMV))
- Analysis of time series data exploring the effects of combining oseltamivir and favipiravir in treating influenza virus
- Development of ABC methods for co-estimating beneficial allele age and selection with application to cryptic coloration data sets (published)

LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE 2006-2008
MSc Immunology of Infectious Diseases

- Classes in immunology, parasitology, immunization and molecular cell biology
- Practical classes on techniques including PCR, ELISA and SDS-PAGE.
- Written projects on the immunology of malaria, tuberculosis, leprosy and schistosomiasis.
- MSc. practical projects: 1) the impact of priming the innate immune system using TLR ligands and 2) the effect of modulating T cell co-stimulation in malaria murine models of blood stage *Plasmodium yoelii* and *Plasmodium berghei* infection.

UNIVERSITY OF CAMBRIDGE (ST EDMUND'S COLLEGE) 2002-2006
BA Hons Natural Sciences (2.1)

- Second and third year specialisation in physics (quantum mechanics, electrodynamics and relativity, statistical and thermal physics, nuclear and particle physics, astrophysics and biophysics)
- First year courses: molecular cell biology, physiology, physics and mathematics
- Written projects on Nuclear Magnetic Resonance (NMR) and quantum mechanics in semi-conductors

UNIVERSITY OF CAMBRIDGE (NEW HALL COLLEGE) 1989-1992
BA Hons Social and Political Sciences

CONFERENCES & WORKSHOPS

Evolutionary genomics and systems biology: bringing together theoretical and experimental approaches (Roscoff, France, to attend in October 2016) 2016
Mind the Gap 5 (Vienna Graduate School of Population Genetics, Austria, to attend in October 2016)
Workshop "Bridging theoretical and experimental evolution" (La Fouly, Switzerland) 2015
Systems Genetics and Evolution of non-human (model) organisms (Ascona, Switzerland) 2014
Society for Molecular Biology and Evolution (SMBE Puerto Rico)
Society for Molecular Biology and Evolution (SMBE Chicago) 2013
Workshop "Bridging theoretical and experimental evolution" (La Fouly, Switzerland)

COURSES & SCHOLARSHIPS

Scholarship awarded for the Summer Institute in Statistical Genetics (SISG) 2015
(University of Washington, Seattle, USA). Modules on probability and statistical inference, regression and analysis of variance, supervised and unsupervised methods for machine learning, MCMC for genetics and Bayesian statistics for genetics.

PhD courses attended at EPFL
Population Genetic Data Analysis, Doctoral Program in Ecology and Evolution (DPEE) 2014
Elements of R for Genetics and Bioinformatics (DPEE)
History of Evolutionary Thought
Introduction to R, Introduction to Python, Introduction to UNIX 2013
Bioinformatics and Genomics (Masters course) 2012

TEACHING EXPERIENCE

Teaching assistant on 14 week introductory third year BA course in Statistical Genetics run by Prof Jensen at EPFL. 2014 & 2013

EMPLOYMENT

International Health Consultant 2010-2012
- Clients included: the UK Department for International Development (DFID), WHO, Comic Relief UK, AusAID, the Innovative Vector Control Consortium (IVCC) and the Liverpool School of Tropical Medicine (LSTM).
- Project examples: DFID-funded programmes for the control of malaria and neglected diseases in several countries (eg. South Sudan, Nigeria), global trachoma elimination programme, evaluation of WHO TDR and non-communicable diseases programmes, economic assessment of the feasibility of a range of malarial control tools.

- Funding range of projects: up to UKP 40m for large scale, national level programmes.

Vice President, Private Equity, Stephens Inc (London office) 2008-2010

- Assessment of potential new investment opportunities in the pharmaceutical, biotechnology & medical device sectors
- Project management of due diligence work
- Industry and market analysis and modelling

Consultant, TechnoServe Tanzania 2004-2005

- Set up a Public Private Partnership (PPP) in Arusha, Tanzania, for a cultivation program for *Artemisia annua* (the plant from which artemisinin is derived), securing $1m USAID/WHO grant for the project.
- Prepared report for IFC (World Bank) into investment opportunities along the supply chain for the production of artemisinin-based combination therapies (ACTs), and for the manufacture and distribution of bednets and indoor residual spraying (IRS)

Associate Director, Schroder Investment Management 1996-2002
(London & Hong Kong)

- Industry and company analysis, focusing initially on the UK leisure sector and subsequently on the Hong Kong and China markets.

PUBLICATIONS

Ormond L, Matuszewski S, Laurent S, Pfeifer S, Renzette N, Jensen JD. Inferring the extent of background selection using an approximate Bayesian approach. 2016. Manuscript in preparation.

Ormond L, Liu P, Matuszewski S, Renzette N, Bank C, Zeldovich K, Bolon DN, Kowalik TF, Finberg RW, Jensen JD, and Wang JP. The combined effect of oseltamivir and favipiravir on influenza A virus evolution. 2017. Manuscript in preparation.

Matuszewski S, Ormond L, Bank C, Jensen JD. Two sides of the same coin: a population genetics perspective on lethal mutagenesis and mutational meltdown. 2017. *Virus Evolution* (accepted).

Irwin KK, Laurent S, Matuszewski S, Vuilleumier S, Ormond L, Shim H, Bank C, Jensen JD. On the importance of skewed offspring distributions and background selection in viral population genetics. 2016. *Heredity* 117(6): 393-399.

Ormond L, Foll M, Ewing GB, Pfeifer SP, Jensen JD. Inferring the age of a fixed beneficial allele. 2016. *Molecular Ecology* 25(1):157-69.

Laurent S, Pfeifer SP, Settles ML, Hunter SS, Hardwick KM, Ormond L, Sousa VC, Jensen JD, Rosenblum EB. The population genomics of rapid adaptation: disentangling signatures of selection and demography in white sands lizards. 2016. *Molecular Ecology* 25(1):306-23.

LANGUAGES Fluent in English and French (written & spoken)

REFERENCE

Prof. Jeffrey D. Jensen
Address: EPFL SV IBI, AAB 048, CH-1015, Switzerland

Supporting information : Estimating the age of a fixed beneficial allele

Supp. Tables

Supp. Table 1. Summary statistics calculated in msstats

(https://github.com/molpopgen/msstats)

Nbr	Statistic	Definition	Type	Reference
1	*S*	Number of segregating sites	Diversity	(Watterson 1975)
2	*n1*	Number of singletons (mutations occurring once in the sample), i.e. at frequency 1 and *n*-1	SFS	
3	*next*	Number of external mutations, corresponding to the number of derived singletons	SFS	(Fu & Li 1993)
4	θ	Watterson's estimate of θ	SFS	(Watterson 1975)
5	π	Sum of site heterozygosity = 2*pq* over *S* sites (where p is the frequency of one allele at each site, and q is the frequency of the alternate allele at that site)	SFS	(Tajima 1983)
6	θ_H	Fay & Wu's estimator of θ where $H= \pi - \theta_H$	SFS	(Fay & Wu 2000)
7	*H'*	Zeng et al's normalized Fay & Wu's *H*	SFS	(Zeng *et al.* 2006)
8	*D*	Tajima's *D* = difference between the number of segregating sites S and the average number of nucleotide differences estimated from pairwise comparison	SFS	(Tajima 1989)
9	*F*	Fu & Li's *F*	SFS	(Fu & Li 1993)
10	*D*	Fu & Li's *D*	SFS	
11	*F**	Fu & Li's *F**	SFS	
12	*D**	Fu & Li's *D**	SFS	
13	*rm*	Hudson & Kaplan's R_{min} = the minimum number of recombination events	LD	(Hudson & Kaplan 1985)
14	*rmmg*	Myers & Griffiths lowest bound on R_{min}	LD	(Myers & Griffiths 2003)
15	*nhap*	Number of haplotypes	Haplotype	(Depaulis & Veuille 1998)
16	*hdiv*	Haplotype diversity	Haplotype	(Depaulis & Veuille 1998)
17	*wallB*	JD Wall's B statistic	LD	(Wall 1999)
18	*wallQ*	JD Wall's Q statistic	LD	
19	*rosarf*	Rf statistic for detecting population growth	SFS	(Ramos-Onsins & Rozas 2002)
20	*rosaru*	Ru statistic for detecting population growth	SFS	
21	*Zns*	Average pairwise LD measured by r^2	LD	(Kelly 1997)

Supp. Table 2. Relative bias and RMSE estimates for inference of *T* alone under strong and moderate selection (*s*=0.1 and *s*=0.01), for young and old sweeps (*L*=20kb, *x*=0.5)

Pseudo observable true values	***T*=0.1**	***T*=0.01**	***T*=0.001**
Strong selection *s*=0.1			
Msstats-ABC			
Relative bias	-0.0573	-0.2352	-0.4969
RMSE	0.0437	0.0006	0.0009
ω_{max}-ABC			
Relative bias	-0.0501	-0.4457	-0.3943
RMSE	0.0410	0.0067	0.0063
Weak selection *s*=0.01			
Msstats-ABC	-0.2897	-0.2451	3.7335
Relative bias	0.0529	0.0147	0.0121
RMSE			
ω_{max}-ABC	-0.3117	-0.8921	0.0384
Relative bias	0.0556	0.0089	0.0005
RMSE			

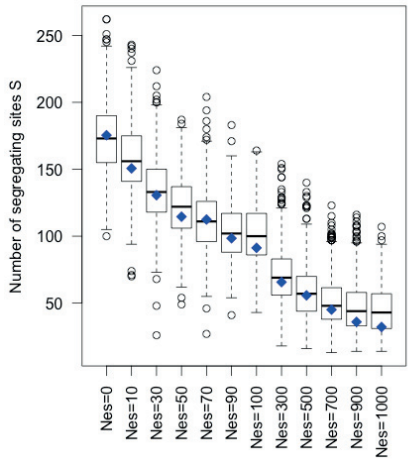
Supp. Table 3. Relative bias and RMSE estimates for joint inference of *s* and *T* for young and old sweeps (*L*=10kb *x*=0.5)

Recent sweeps **Pseudo observable true values**	***T*=0.01** ***s*=0.1**	***T*=0.01** ***s*=0.01**	***T*=0.01** ***s*=0.001**
msstats-ABC			
Relative bias in *s*	-0.0553	0.6208	83.4751
RMSE in *s*	0.0556	0.0195	0.1117
Relative bias in *T*	-0.2803	1.4785	5.1630
RMSE in *T*	0.0072	0.0361	0.0737
ω_{max}-ABC			
Relative bias in *s*	-0.0906	0.2472	6.4747
RMSE in *s*	0.0567	0.0085	0.0082
Relative bias in *T*	-0.2064	0.3972	3.0285
RMSE in *T*	0.0074	0.0215	0.0525
Old sweeps **Pseudo observable true values**	***T*=0.1** ***s*=0.1**	***T*=0.1** **s=0.01**	***T*=0.1** ***s*=0.001**
msstats-ABC			
Relative bias in *s*	0.0784	1.5859	55.9636
RMSE in *s*	0.0514	0.0399	0.0804
Relative bias in *T*	-0.0327	-0.0108	0.3254
RMSE in *T*	0.0340	0.0637	0.0648
ω_{max}-ABC			
Relative bias in *s*	-0.3330	0.6524	4.7463
RMSE in *s*	0.0646	0.0214	0.0113
Relative bias in *T*	-0.2272	-0.0656	0.1842
RMSE in *T*	0.0491	0.0602	0.0626

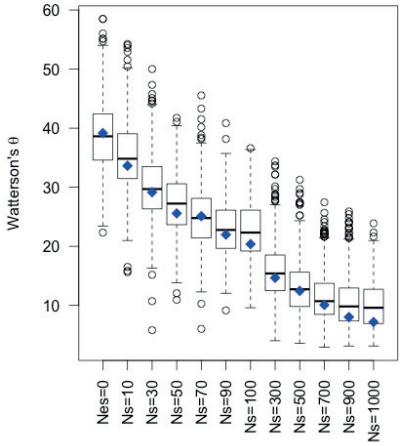
Supp. Figures

Supp. Fig.1 Variation of selected summary statistics with population scaled selection coefficient N_es. Boxplots represent the results of 100 simulations over different orders of magnitude, including one neutral scenario (N_es=0). Sweeps are young (0.01x4N_e generations). Blue points represent the modes of distributions, lines represent the medians of distributions. Plots are shown spanning 3 orders of magnitude from N_es=0 to N_es=1000. A. Number of segregating sites *S,* B. Watterson's θ, C. π, D. Tajima's *D,* E. Number of haplotypes, F. Haplotype diversity.

A.



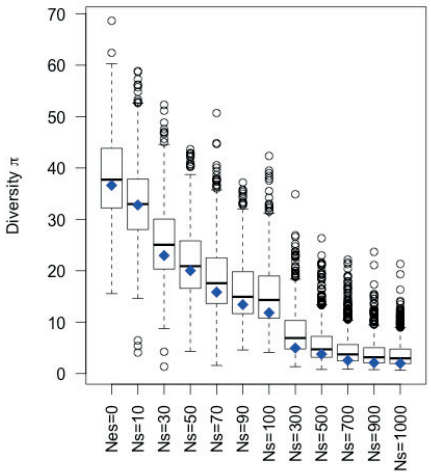
B.



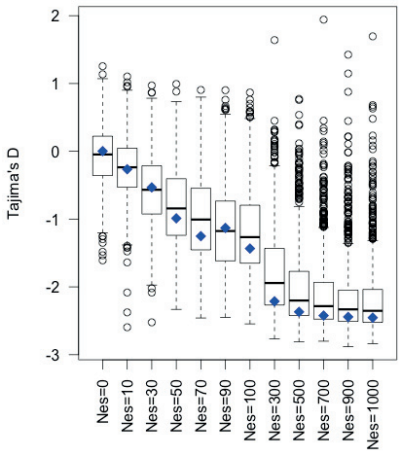
C.

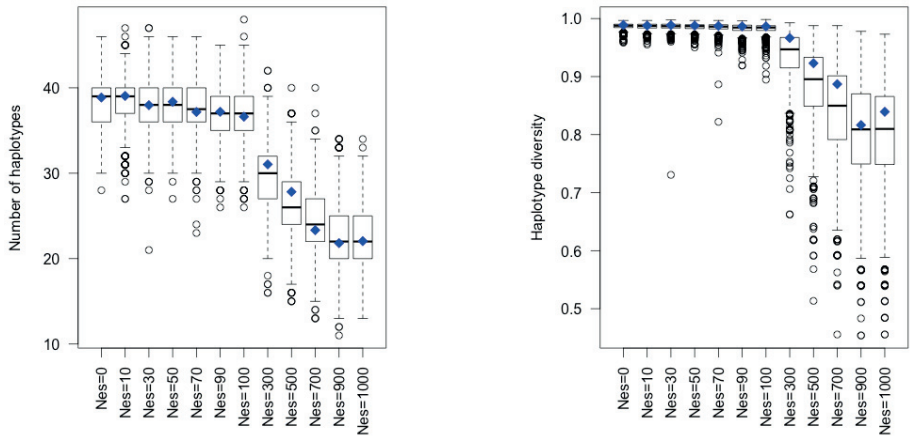
D.

E.



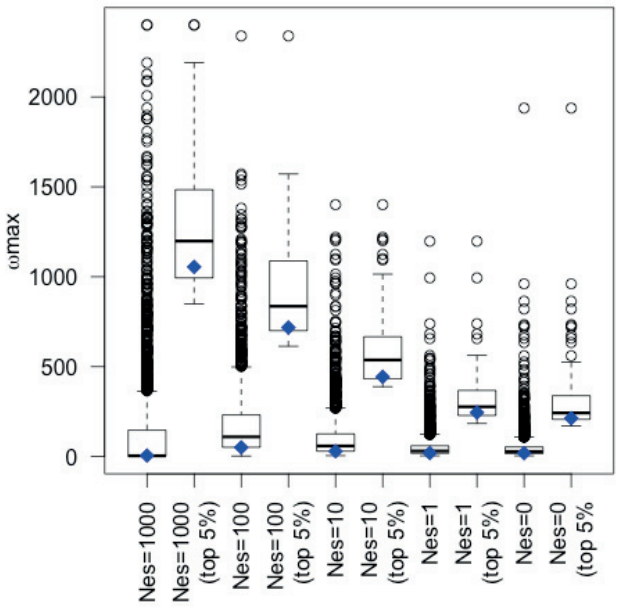
F.



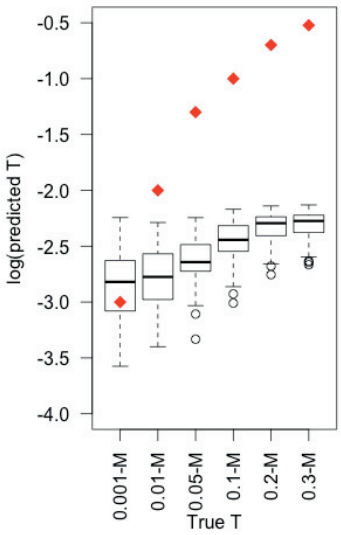


Supp Fig. 2 Variation of ω_{max} with selection

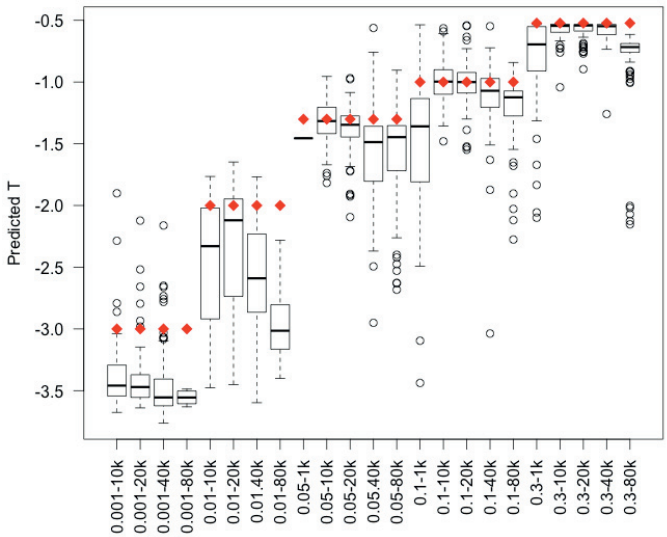
The values of ω_{max} are shown for 100 data points. Values not corrected for ascertainment are shown in the left hand boxes within each pair (100 simulations for each case of s=0.1, 0.01,0.001 and 0 and T=0.01x4N_e). Values corrected for ascertainment (top 5% of 2000 simulations, which gives 100 data points) are shown in right boxes. L=50kb. $x_{\omega max}$= 0.5 (25kb).



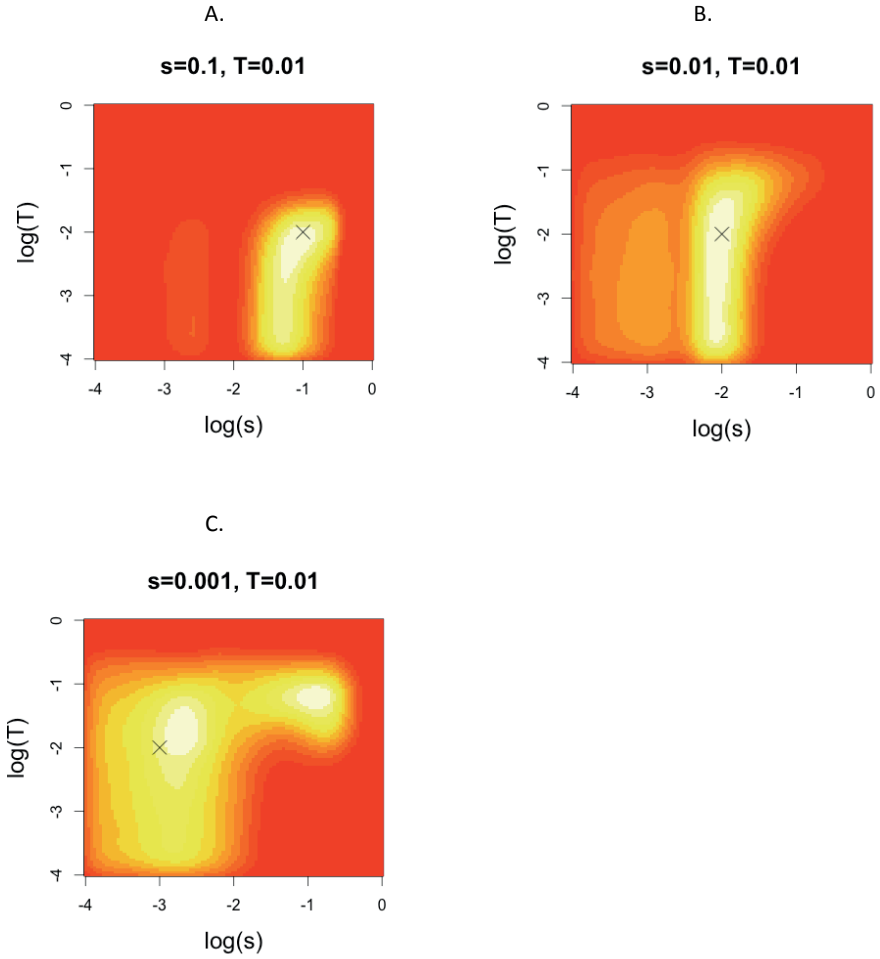
Supp Fig. 3 Boxplot of the modes of posterior distributions of allele age *T* under weak selection (*s*=0.001). The boxplots show results from msstats-ABC (marked M). *T* is drawn from a log uniform priors: $log_{10}(T)$ ~ U(-4, -0.5). Boxplots represent the modes of posterior distributions for 100 pseudo-observables. L = 2 x 10^4 bps, x = 0.5, ρ = μ = 10^{-7} per base pair per generation, N = 10,000, $\alpha = 2Ns$ = 20. Red diamonds mark the true values of pseudo-observables(T = 0.001,0.01,0.05, 0.1,0.2,0.3 x $4N_e$ generations).

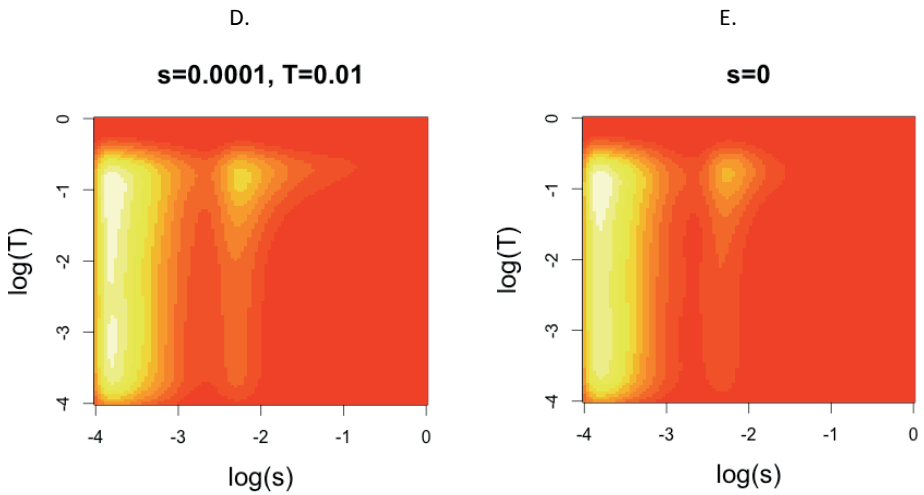


Supp Fig. 4 Boxplot comparing the impact of different window lengths on inference power for allele age *T*, under strong selection (*s*=0.1). Boxplots compare results from msstats-ABC for inference using length of 1kb, 10kb, 20kb, 40kb and 80kb. Parameters are as described in methods. *T* is drawn from a log uniform prior: $log_{10}(T)$ ~ U(-4, -0.5). Red diamonds mark the true values for respective inferred values (T = 0.001,0.05,0.01,0.1,0.2,0.3 x $4N_e$ generations).

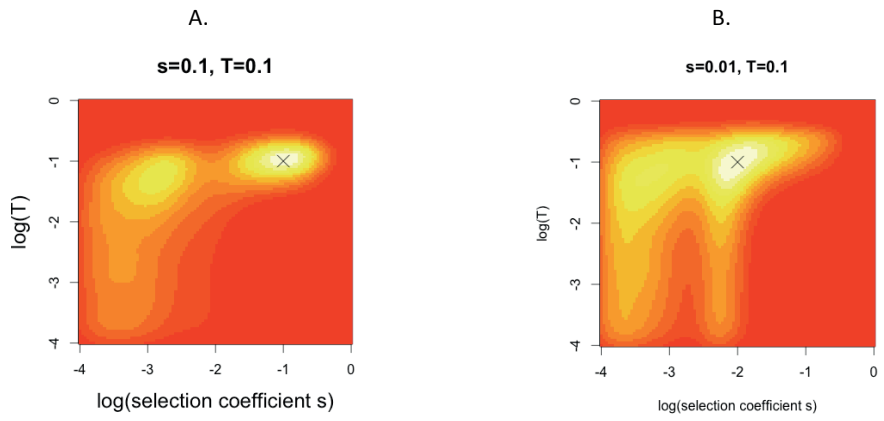


Supp. Fig. 5. Joint inference of *s* and *T* in equilibrium populations for young sweeps (*T*=0.01) (msstats-ABC). Figures show the cumulative joint posterior density plots for 100 pseudo-observable simulations over different orders of magnitude of the selection coefficient *s*, for young sweeps (*T*=0.01) and A. *s* = 0.1 ($\alpha=N_e s=10^3$); B. *s*=0.01($\alpha=N_e s=10^2$); C. *s*= 0.001($\alpha=N_e s=10$);. The bottom two panels represent neutral scenarios with D. *s*=0.0001 ($\alpha=N_e s=1$); and E. *s*=0. The white, yellow and red colors mark areas of high, moderate and low joint density respectively. Black crosses indicate the true values of pseudo-observables. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods.

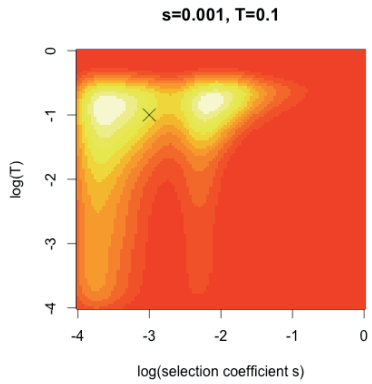




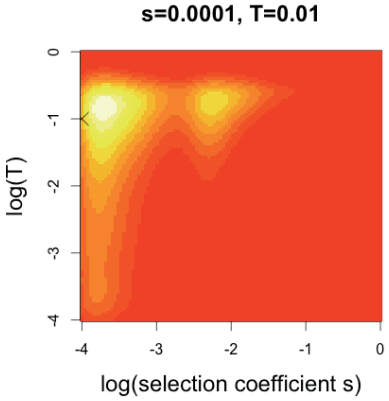
Supp Fig. 6 Joint inference of *s* and *T* in equilibrium populations for old sweeps (*T*=0.01) (ω_{max} –ABC). Figures show the cumulative joint posterior density plots for 100 pseudo-observable simulations over different orders of magnitude of the selection coefficient *s,* for old sweeps (*T*=0.1) and A. *s* = 0.1 ($\alpha=N_e s=10^3$); B. *s*=0.01($\alpha=N_e s=10^2$); C. *s*= 0.001($\alpha=N_e s=10$). The bottom two panels represent neutral scenarios with D. *s*=0.0001 ($\alpha=N_e s=1$); and E. *s*=0. Pseudo-observables were set up by running $2x10^3$ simulations for selected values of *s* and *T,* and retaining the top 5% by values of ω_{max}, generating 100 test values. The white, yellow and red colors mark areas of high, moderate and low joint density respectively. Black crosses indicate the true values of pseudo-observables. Other parameters are as in previous figures. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods.



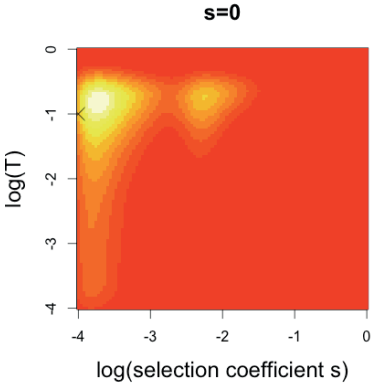
C.


s=0.001, T=0.1
log(T)
0
-1
-2
-3
-4
-4
-3
-2
-1
0
log(selection coefficient s)

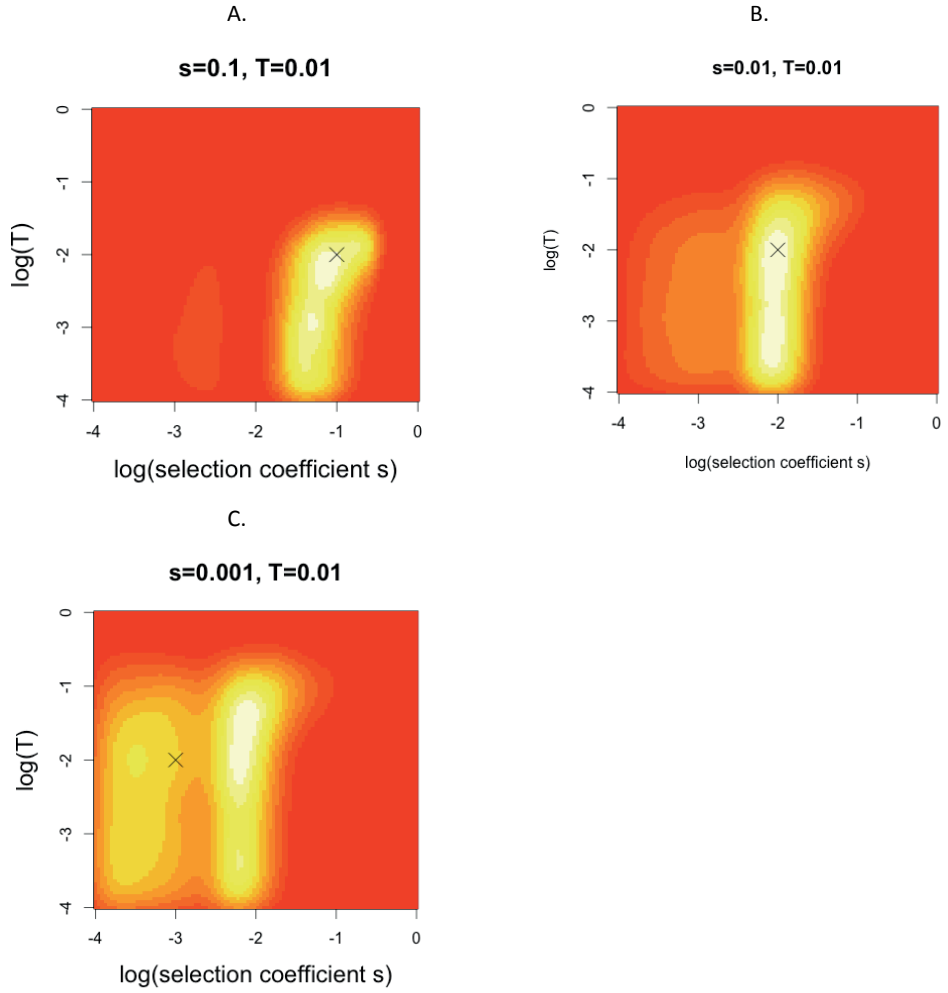
D.

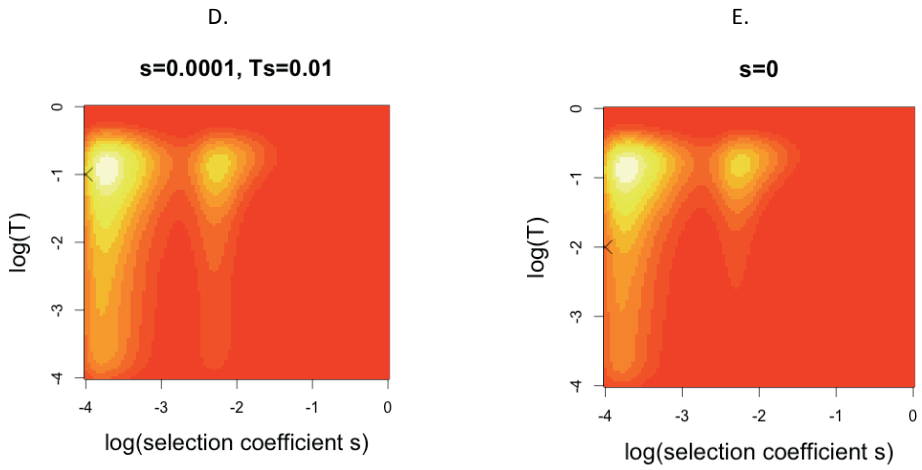

s=0.0001, T=0.01
log(T)
0
-1
-2
-3
-4
-4
-3
-2
-1
0
log(selection coefficient s)

E.

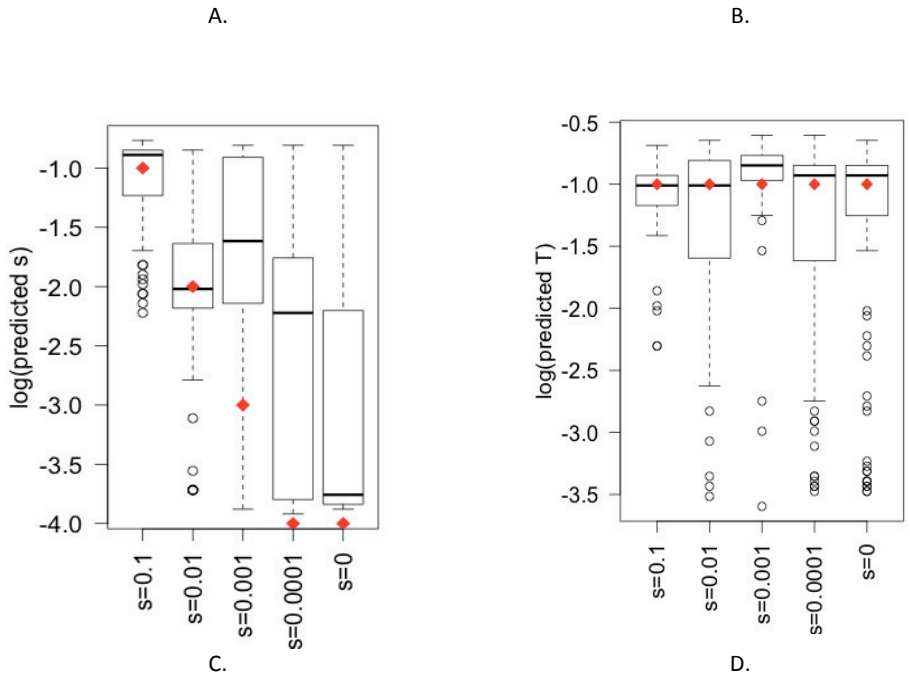

s=0
log(T)
0
-1
-2
-3
-4
-4
-3
-2
-1
0
log(selection coefficient s)

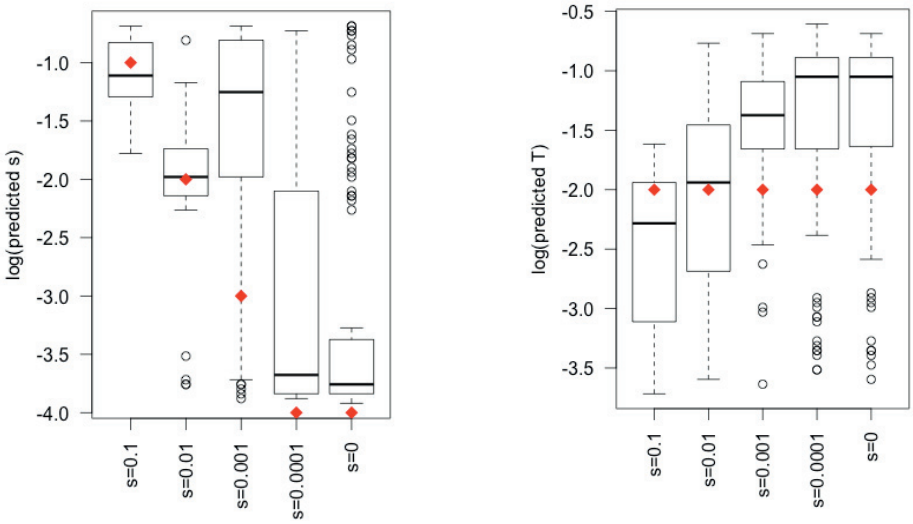
Supp Fig. 7 Joint inference of *s* and *T* in equilibrium populations for young sweeps (*T*=0.01) (ω_{max} –ABC). Figures show the cumulative joint posterior density plots for 100 pseudo-observable simulations over different orders of magnitude of the selection coefficient *s,* for young sweeps (*T*=0.01) and A. *s* = 0.1 ($\alpha=N_es=10^3$); B. *s*=0.01($\alpha=N_es=10^2$); C. *s*= 0.001($\alpha=N_es=10$). The bottom two panels represent neutral scenarios with D. *s*=0.0001 ($\alpha=N_es=1$); and E. *s*=0. Pseudo-observables were set up by running $2x10^3$ simulations for selected values of *s* and *T,* and retaining the top 5% by values of ω_{max}, generating 100 test values. The white, yellow and red colors mark areas of high, moderate and low joint density respectively. Black crosses indicate the true values of pseudo-observables. Other parameters are as in previous figures. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods.





Supp Fig. 8 Boxplots of predicted values from joint inference of *s* and *T* in equilibrium populations (*msstats*-ABC). Boxplots show the predicted values for 100 pseudo-observables over different magnitudes of selection coefficient *s* (*s* = 0.1, 0.01,0.001, 0.0001 and 0), and *T* = 0.1 (A and B) or *T* = 0.01 (C and D). Red diamonds mark the known values of pseudo-observables. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods.

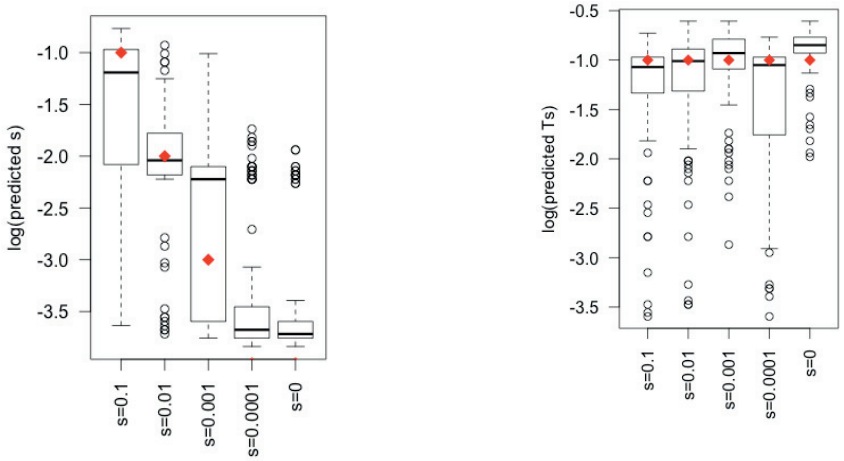




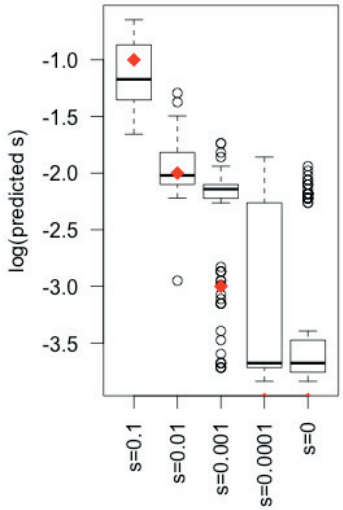
Supp Fig. 9 Boxplots of predicted values from joint inference of *s* and *T* in equilibrium populations (*w_max*-ABC). Boxplots show the predicted values for 100 pseudo-observables over different magnitudes of selection coefficient *s* (*s* = 0.1, 0.01,0.001, 0.0001 and 0), and *T* = 0.1 (A and B) or *T* = 0.1 (C and D). Red diamonds mark the known values of pseudo-observables. *s* and *T* are drawn from log uniform priors: $log_{10}(s)$ ~ U(-4, -0.5) and $log_{10}(T)$ ~ U(-4, -0.5). Other parameters are as described in methods

A.

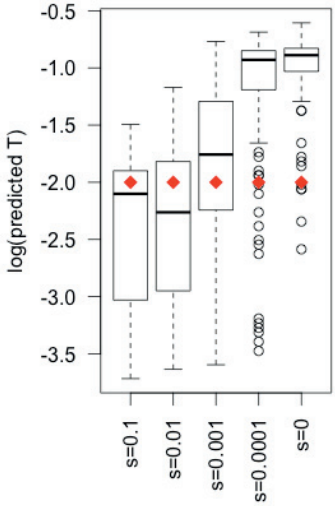
B.



C.

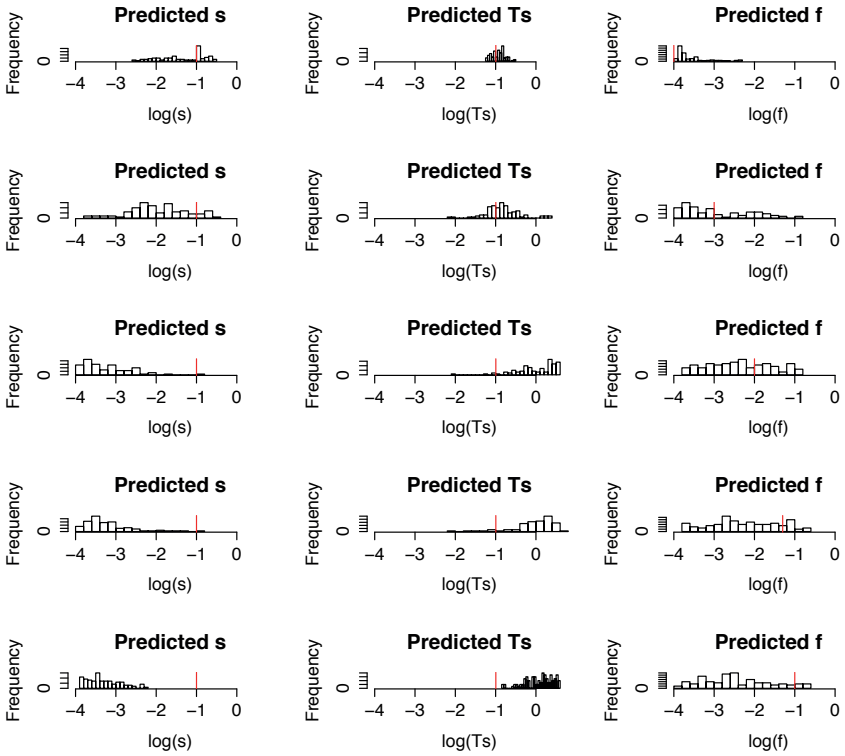

log(predicted s)
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
s=0.1
s=0.01
s=0.001
s=0.0001
s=0

D.

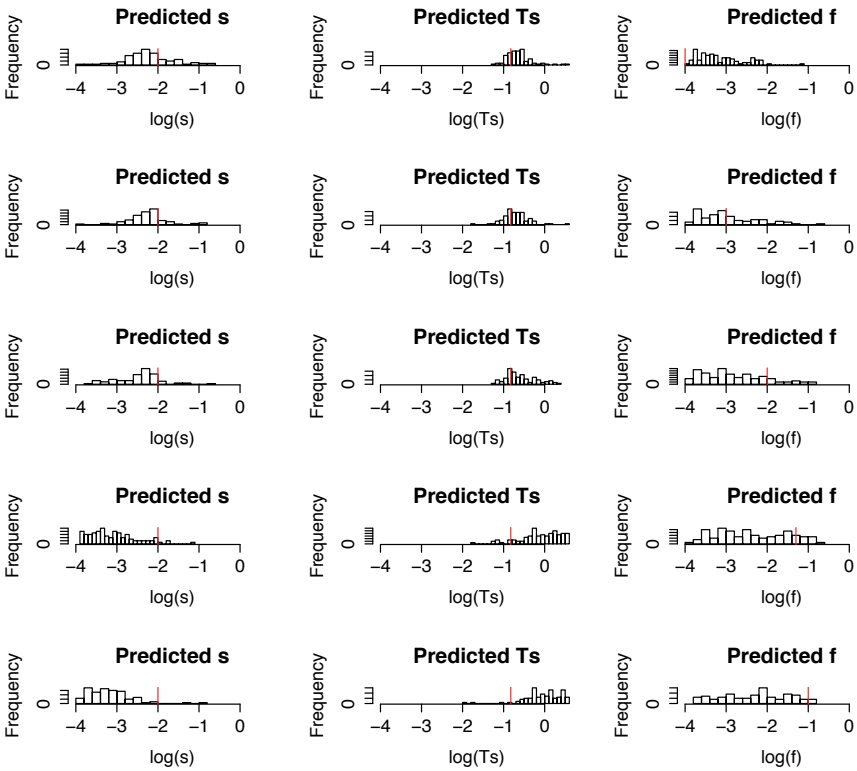

log(predicted T)
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
s=0.1
s=0.01
s=0.001
s=0.0001
s=0

Supp. Fig 10 Joint inference of *s, T_s* and *f* in equilibrium populations. Figures show the predicted values for 100 pseudo-observables for five cases of starting frequency *f*=0.0001, 0.001, 0.01, 0.05, 0.1. Estimates of *s, T_s* and *f* were obtained from the mode of the joint posterior density. Red lines indicate the true values of the pseudo-observables. T_s represents the time since selection began acting on the allele (time to fixation plus sojourn time). A. *s*=0.01, T_s=0.105 B. *s*=0.01, T_s=0.150. C. *s*=0.1, T_s =0.015.D *s*=0.001, T_s=0.505 E. *s*=0.001, T_s=0.006.

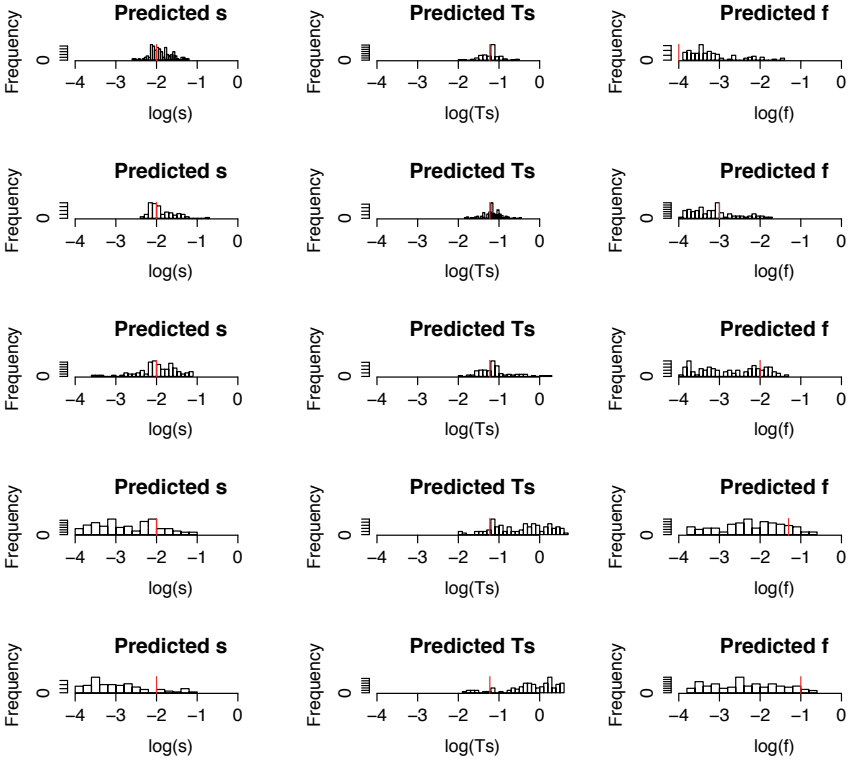
A. *s*=0.1, T_s=0.105



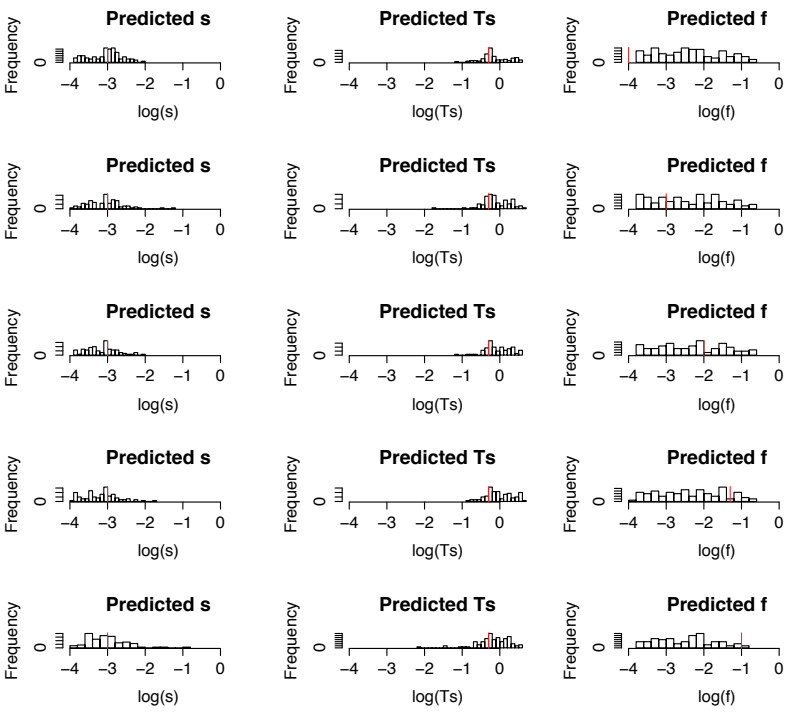
B. s=0.01, T_s=0.150



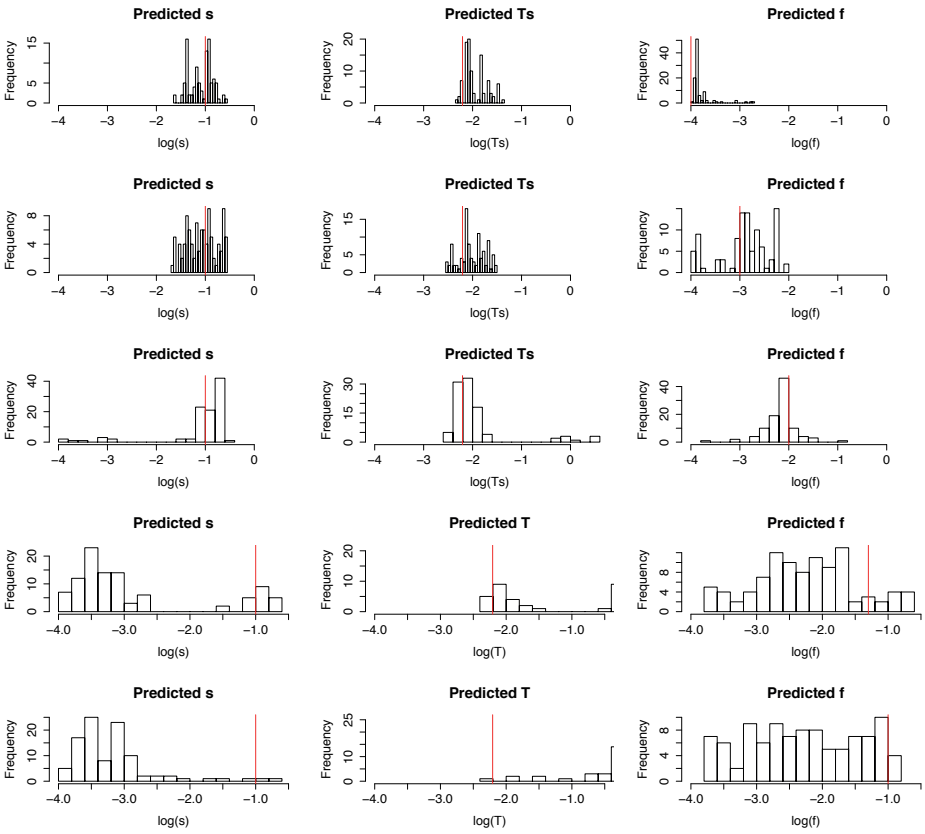
C. s=0.1, T_s=0.015



D. s=0.001, T_s=0.505

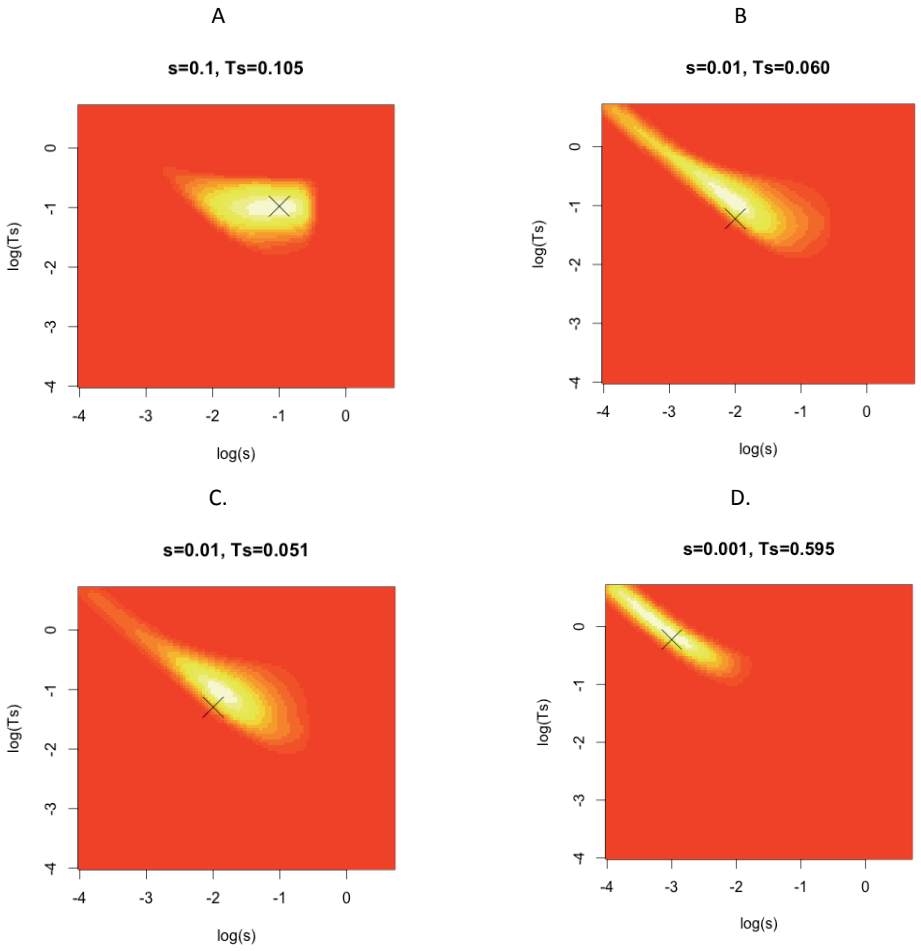


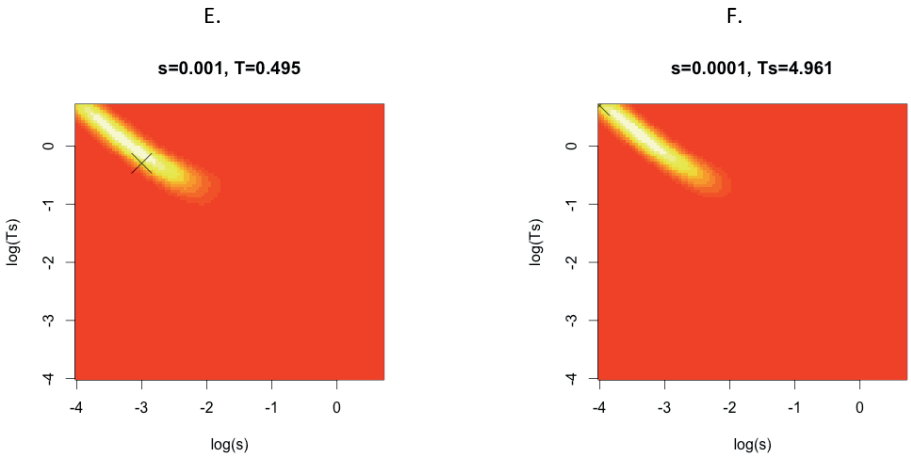
E. s=0.1, T_s=0.006



Supp. Fig. 11 Joint inference of *s* and T_s under the demographic scenario of exponential growth following a sharp bottleneck using *msstats*–ABC (demographic scenario 3 described in methods). Exponential growth is modeled with parameter α=460.5, following a 1% bottleneck at 0.01 x $4N_e$ generations. Figures show the cumulative joint posterior density plots for 100 pseudo-observable simulations. The white, yellow and red colors mark areas of high, moderate and low joint density respectively. Black crosses indicate the true values of pseudo-observables. *S* and T_s are drawn from a log uniform prior for s: $log_{10}(s)$ ~ U(-4, -0.5) and adjusted log uniform prior for T_s: $log_{10}(T_s)$ ~ U($\log_{10}(T_{soj})$, $\log_{10}(0.3+T_{soj})$). Other parameters are as in previous figures.

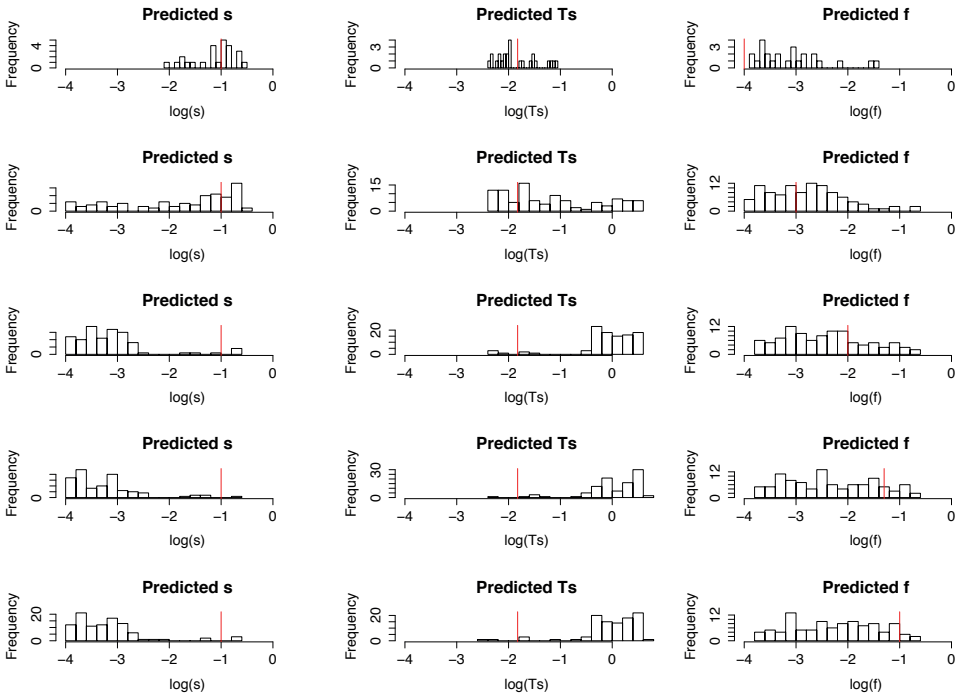
The figures below represent supplementary figures to Fig. 4. A. Inference of a strong, old sweep with pseudo-observable values s=0.1 and T_s=0.105 (calculated from $T_s=T+T_{soj}$ where T=0.1). B. Inference of a moderately strong, recent sweep with pseudo-observable values s=0.01 and T_s=0.06 (calculated from $T_s=T+T_{soj}$ where T=0.01). C. Inference of a moderately strong, very recent sweep with pseudo-observable values s=0.1 and T_s=0.051 (calculated from $T_s=T+T_{soj}$ where T=0.001). D. Inference of a weak, old sweep with pseudo-observable values s=0.001, T_s=0.1(calculated from $T_s=T+T_{soj}$ where T=0.1). E. Inference of a weak, recent sweep with pseudo-observable values s=0.01, T_s=0.01(calculated from $T_s=T+T_{soj}$ where T=0.01). F. Inference of a neutral mutation drifting to fixation with pseudo-observable values s=0.0001, T_s=4.961(calculated from $T_s=T+T_{soj}$ where T=0.01)



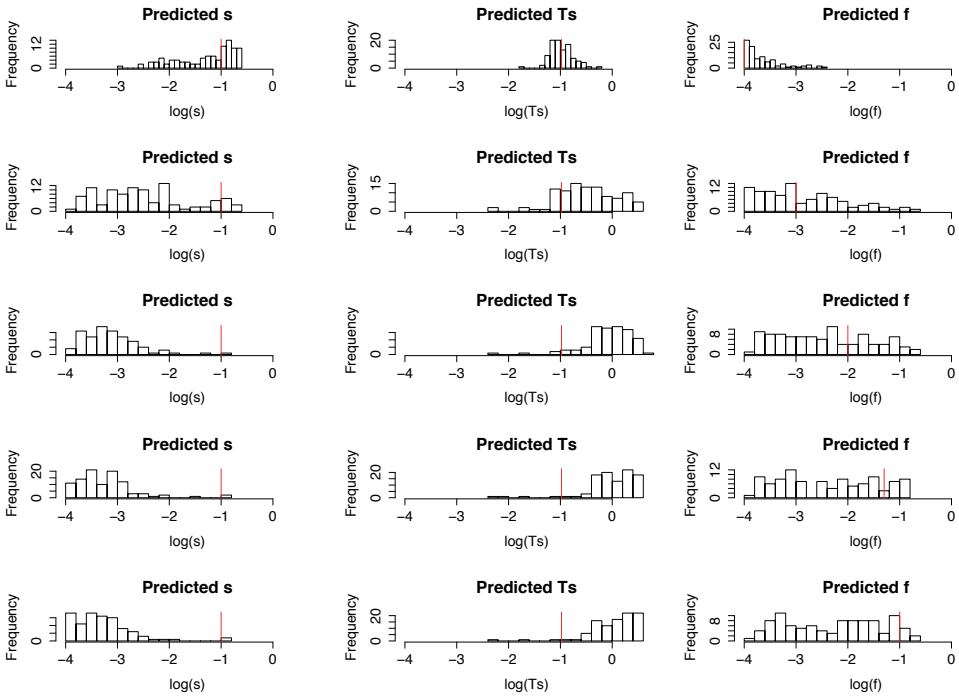


Supp. Fig 12 Joint inference of *s, T_s* and *f* in a population undergoing exponential growth after a sharp bottleneck (demographic scenario 3) Figures show the predicted values for 100 pseudo-observables for five cases of starting frequency *f*=0.0001, 0.001, 0.01, 0.05, 0.1. Estimates of *s, T_s* and *f* were obtained from the mode of the joint posterior density. Red lines indicate the true values of the pseudo-observables. T_s represents the time since selection began acting on the allele (time to fixation plus sojourn time). A. *s*=0.1, T_s=0.015. B. *s*=0.1, T_s=0.105. C. *s*=0.01, T_s=0.060

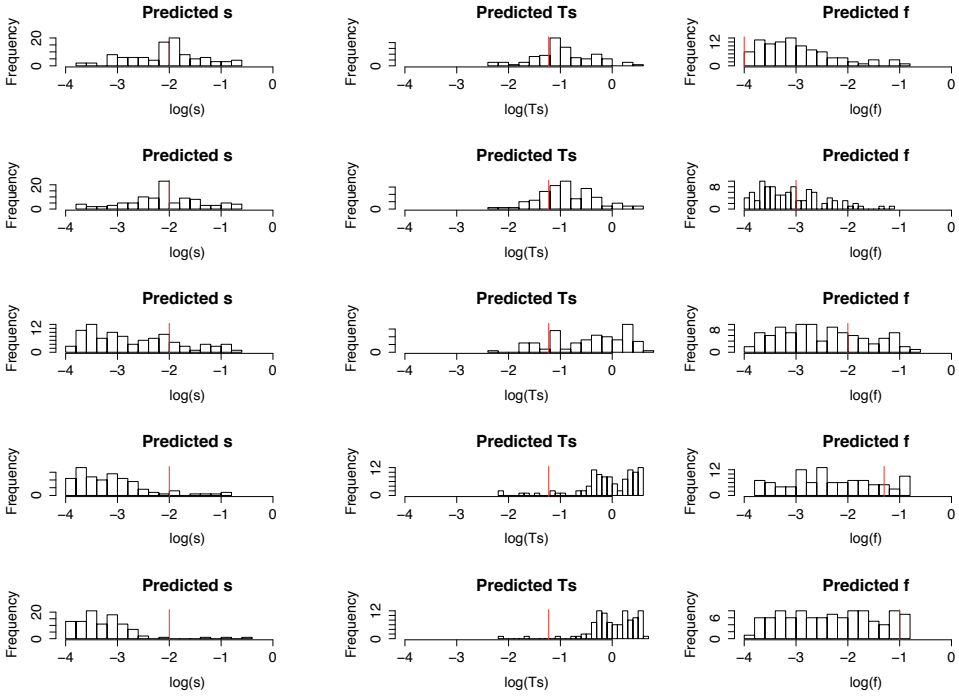
A. *s*=0.1, T_s=0.015



B. s=0.1, T_s=0.105



C. s=0.01, T_s=0.060



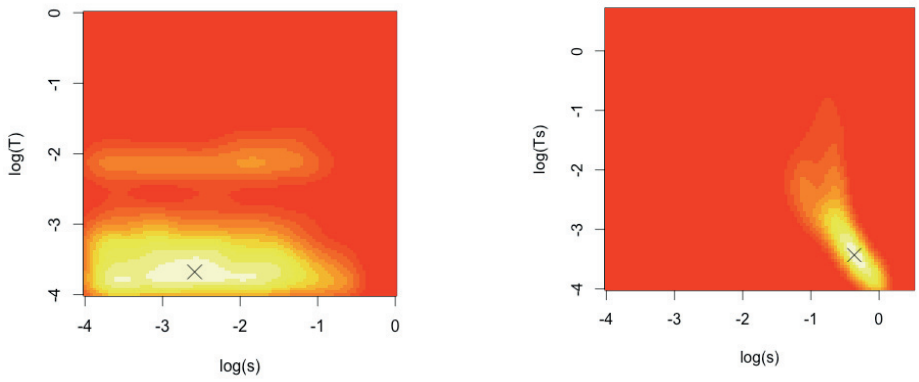
Supp Fig. 13 Joint inference of allele age and selection coefficient for *P. maniculatus*

The density plots represent the joint posterior distribution of *s* and either *T* (A) or T_s (B) for the serine deletion position 128,150 on exon 2. 500,000 simulations were run assuming N_e= 53,080, $\rho = 3.67 \times 10^{-8}$,$\mu = 0.62 \times 10^{-8}$ per base pair per generation. Simple rejection ABC was used with a tolerance level of 0.005. Density plots are shown with the mutation positioned centrally (x=0.5) on a length *L*=40kb : A. with equilibrium demography and B. with the demographic scenario included in simulations).

For A, the joint density maximum occurs at $s=2.6\times10^{-4}$ (1.2×10^{-3}-2.3×10^{-1}), $T=2.1\times10^{-4}$(1.1×10^{-4}-2.2×10^{-2}).

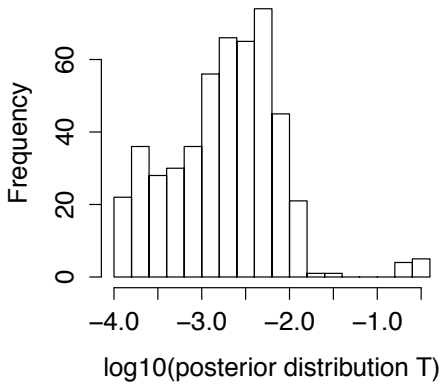
For B, the joint density maximum occurs at $s=4.3\times10^{-1}$ (3.5×10^{-2}-1.0), $T_s=6.3\times10^{-2}$(3.7×10^{-4}-1.8×10^{-1})

A. B.



Supp. Fig. 14 Inference of allele age *T* alone for *P. maniculatus*

The histogram represents the posterior distribution on a log10 scale of allele age *T* for the serine deletion at position 128,150 on exon 2. 40kb were used on either side of the serine deletion for a total length *L*=80kb. 100,000 simulations were run assuming N_e= 53,080, $\rho = 3.67 \times 10^{-8}$,$\mu = 0.62 \times 10^{-8}$ per base pair per generation and *s*=0.126 (Linnen *et al.* 2013). *T* was drawn from a log uniform prior: $log_{10}(T) \sim U(-4, -0.5)$.



References

Depaulis F, Veuille M (1998) Neutrality tests based on the distribution of haplotypes under an infinite-site model. *Mol Biol Evol* **15**, 1788-1790.

Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics* **155**, 1405-1413.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* **133**, 693-709.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**, 147-164.

Kelly J (1997) A test of neutrality based on interlocus associations. *Genetics* **146**, 1197-1206.

Linnen CR, Poh YP, Peterson BK*, et al.* (2013) Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312-1316.

Myers SR, Griffiths RC (2003) Bounds on the minimum number of recombination events in a sample history. *Genetics* **163**, 375-394.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**, 2092-2100.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437-460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.

Wall J (1999) Recombination and the power of statistical tests of neutrality. *Genetical Research Cambridge* **74**, 65-79.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* **7**, 256-276.

Zeng K, Fu YX, Shi S, Wu CI (2006) Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* **174**, 1431-1439.

Supporting information : The population genomics of rapid adaptation : disentangling the signatures of selection and demography in White Sands lizards

Table S1a

ID	Habitat	Coverage[1]
CP1	DS2	20.97
CP2	DS2	3.25
CP3	DS2	175.87
CP4	WS	216.56
CP5	WS	20.73
CP6	WS	12.36
CP13	WS	112.91
CP15	DS2	7.7
CP16	DS2	10.57
CP17	DS2	25.53
CP18	WS	222.69
CP19	WS	15.75
CP33	WS	147.80
CP34	WS	39.98
CP35	WS	129.65
CP36	WS	106.72
CP37	WS	190.66
CP38	WS	143.68
CP48	DS2	197.19
CP49	DS2	228.17
CP50	DS2	66.14
CP51	DS2	72.45
CP52	DS2	94.47
CP64	DS1	211.49
CP65	DS1	245.64
CP66	DS1	144.84
CP67	DS1	66.81
CP68	DS1	153.11
CP69	DS1	142.35
CP70	DS1	80.31
CP71	DS1	137.65
CP72	DS1	129.57
Mean		**111.67**

Table S1a: *A. inornata* samples and statistics of mapped coverage.

[1] After duplicate removal and only counting read pairs with appropriate orientation. The genome size was taken to be 3.2 Gb.

Table S1b

ID	**Habitat**	**Coverage[1]**
CP7	DS2	268.77
CP8	DS2	18.61
CP9	WS	15.77
CP10	WS	165.63
CP14	DS2	112.27
CP20	DS2	8.63
CP22	DS2	2.72
CP23	DS2	49.86
CP29	WS	149.17
CP30	WS	205.59
CP31	WS	148.46
CP41	WS	39.87
CP42	WS	68.99
CP43	WS	170.79
CP44	WS	133.45
CP45	WS	60.10
CP57	DS2	59.91
CP58	DS2	93.93
CP59	DS2	140.37
CP73	DS1	226.67
CP74	DS1	153.94
CP75	DS1	91.54
CP76	DS1	68.19
CP77	DS1	262.20
CP78	DS1	165.35
CP79	DS1	122.43
CP80	DS1	89.10
CP81	DS1	113.45
Mean		**114.49**

Table S1b: *S. cowlesi* samples and statistics of mapped coverage.

[1] After duplicate removal and only counting read pairs with appropriate orientation. The genome size was taken to be 3.3 Gb.

Table S2

A.inornata

mutation rate	gen time	Ne(W)	Ne(J)	Ne(P)	Ne(ref)	Ne(N)	T2 (years)	T1 (years)
1.50E-08	1	2'339	8'742	744	32'411	12'326	1'211	303
1.50E-09	1	23'392	87'418	7'443	324'114	123'264	12'112	3'025
1.50E-08	1.5	2'339	8'742	744	32'411	12'326	1'817	454
1.50E-09	1.5	23'392	87'418	7'443	324'114	123'264	18'168	4'538
1.50E-08	2	2'339	8'742	744	32'411	12'326	2'422	605
1.50E-09	2	23'392	87'418	7'443	324'114	123'264	24'224	6'051

S.cowlesi

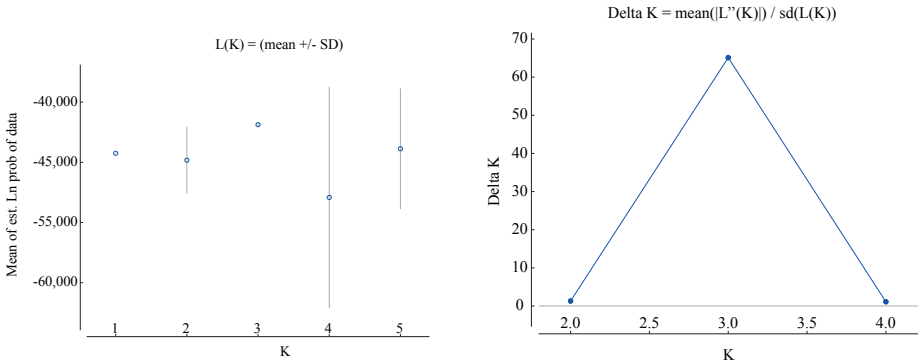
mutation rate	gen time	Ne(W)	Ne(J)	Ne(P)	Ne(ref)	Ne(N)	Ne(PS)	T2 (years)	T1 (years)	m(W->J)	m(W->P)	m(J->W)	m(P->W)	m(J->P)	m(P->J)
1.50E-08	1	2'024	3'207	1'523	42'068	151'557	107'677	16'977	493	0.000031	0.000175	0.000030	0.000295	0.000204	0.000275
1.50E-09	1	20'237	32'068	15'232	420'678	1'515'572	1'076'768	169'768	4'934	0.000003	0.000017	0.000003	0.000029	0.000020	0.000028
1.50E-08	1.5	2'024	3'207	1'523	42'068	151'557	107'677	25'465	740	0.000031	0.000175	0.000030	0.000295	0.000204	0.000275
1.50E-09	1.5	20'237	32'068	15'232	420'678	1'515'572	1'076'768	254'652	7'401	0.000003	0.000017	0.000003	0.000029	0.000020	0.000028
1.50E-08	2	2'024	3'207	1'523	42'068	151'557	107'677	33'954	986.7	0.000031	0.000175	0.000030	0.000295	0.000204	0.000275
1.50E-09	2	20'237	32'068	15'232	420'678	1'515'572	1'076'768	339'536	9'867	0.000003	0.000017	0.000003	0.000029	0.000020	0.000028

Table S2: An example of the effects of uncertainty in mutation rate (two rates given) and generation time (three times given) on estimated demographic parameters.

Supplementary Figures

Figure S1

A



B

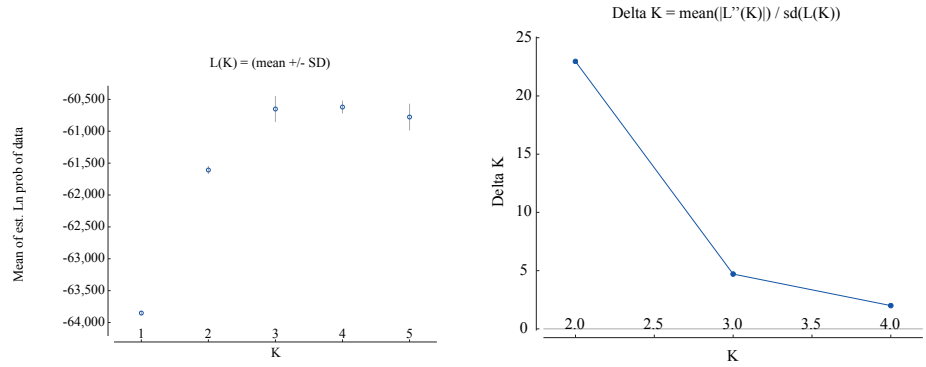


Figure S1: Left: Mean log posterior probability of the data, L(K), in (A) *A. inornata* and (B) *S. cowlesi*. Right: Measure of the rate of change in the log probability of the data between successive *K* values, Δ(K), in (A) *A. inornata* and (B) *S. cowlesi*.

Figure S2

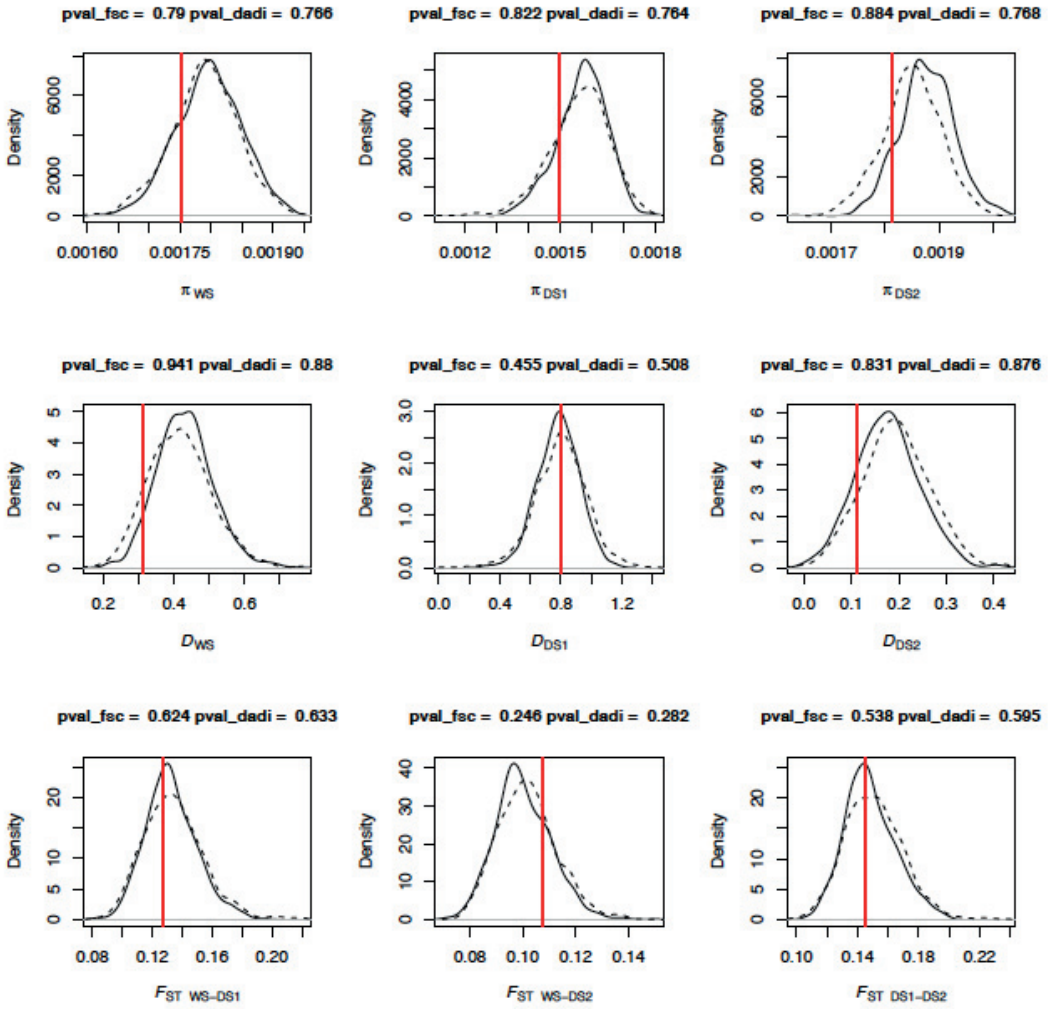


Figure S2: Predictive distributions of 1,000 simulated values of π, Tajima's D, and F_{st} under the best demographic model for *A. inornata* as inferred by fastsimcoal2 (solid line) and *δaδi* (dashed line). Red horizontal bars represent the observed values. P-values represent the probabilities that the simulated values are larger than the observed values.

Figure S3

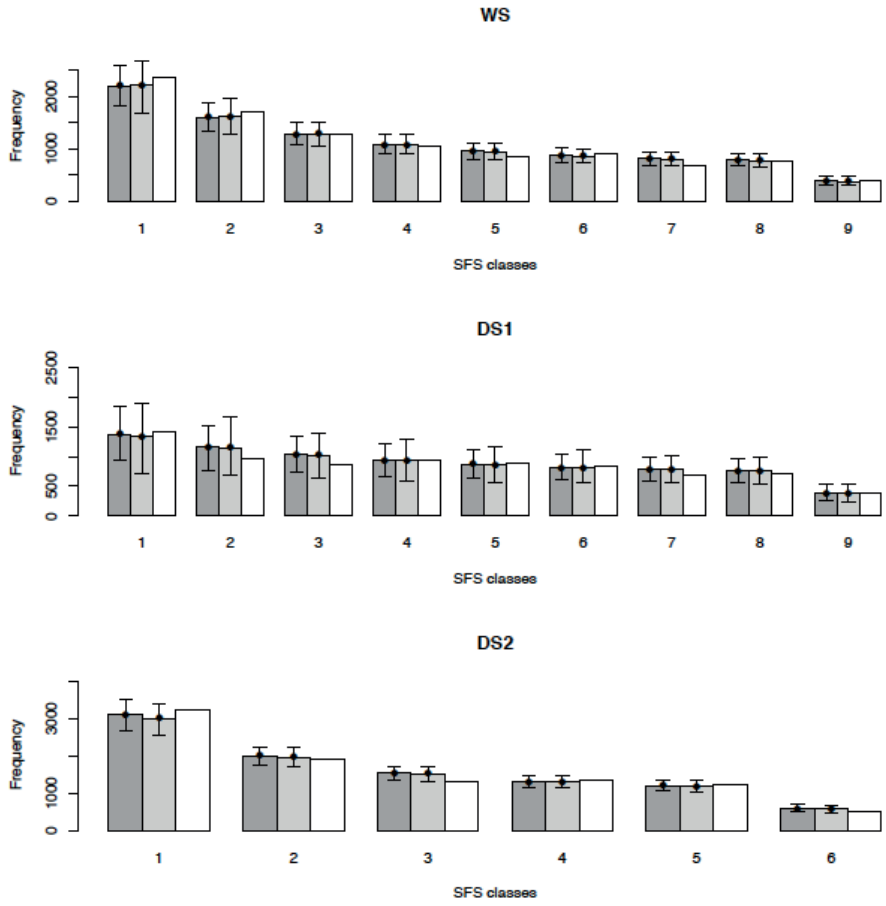


Figure S3: Predictive distributions of 1,000 simulated site frequency spectra under the best demographic model for *A. inornata* as inferred by fastsimcoal2 (dark gray) and *δaδi* (light gray). The error bars represent the 0.025^{th} and 0.975^{th} percentiles of the simulated distributions. The white bars represent the observed SFS.

Figure S4

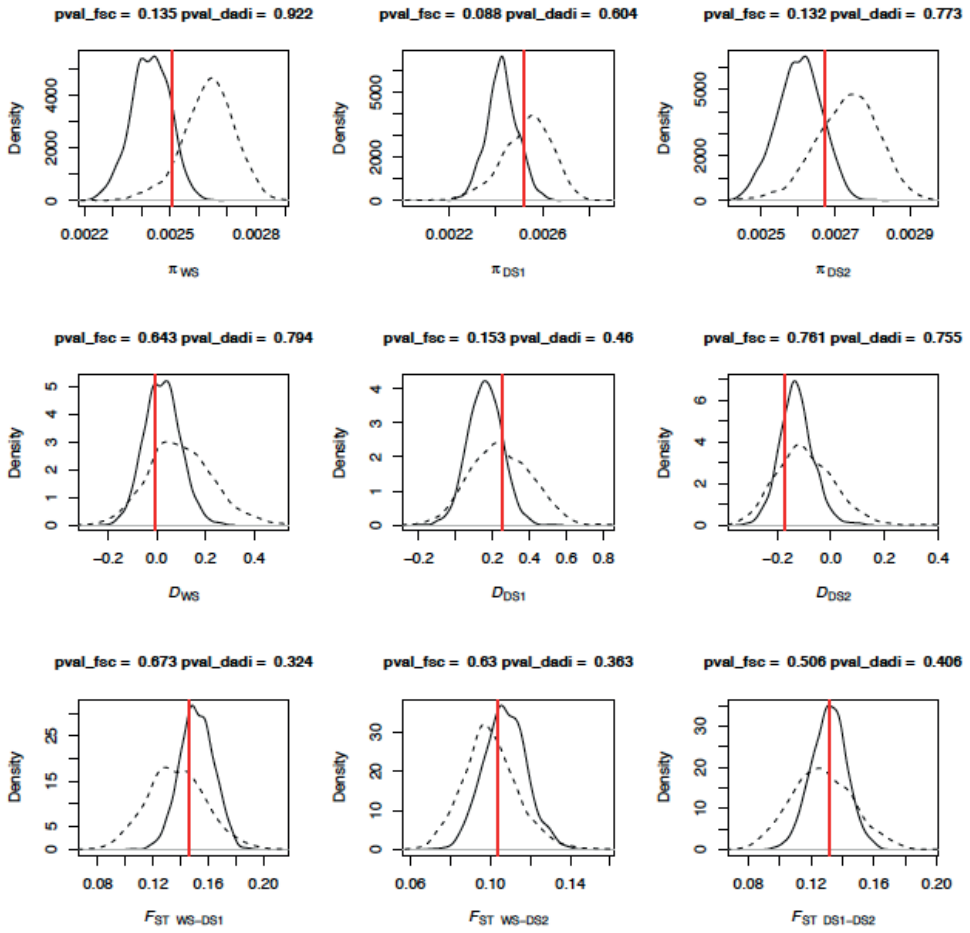


Figure S4: Predictive distributions of 1,000 simulated values of π, Tajima's D, and F_{st} under the best demographic model for *S. cowlesi* as inferred by fastsimcoal2 (solid line) and *δaδi* (dashed line). Red horizontal bars represent the observed values. P-values represent the probabilities that the simulated values are larger than the observed values.

Figure S5

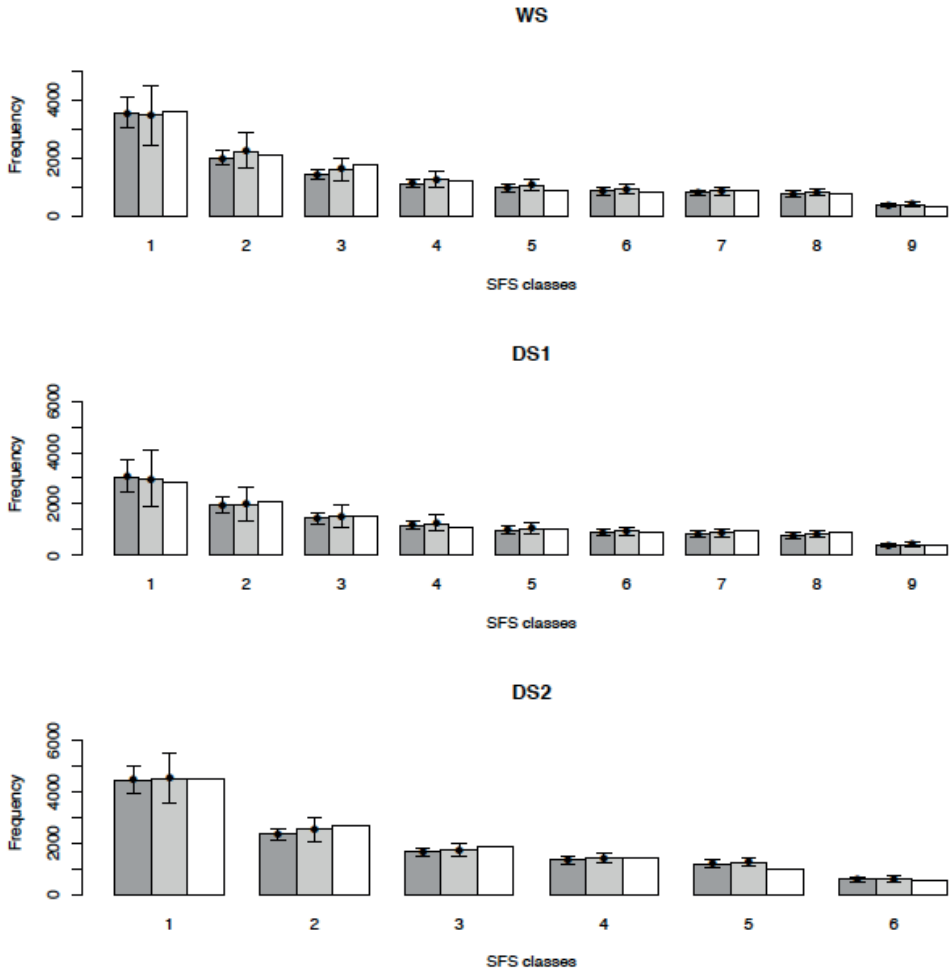


Figure S5: Predictive distributions of 1,000 simulated site frequency spectra under the best demographic model for *S. cowlesi* as inferred by fastsimcoal2 (dark gray) and *δaδi* (light gray). The error bars represent the 0.025th and 0.975th percentiles of the simulated distributions. The white bars represent the observed SFS.

Figure S6

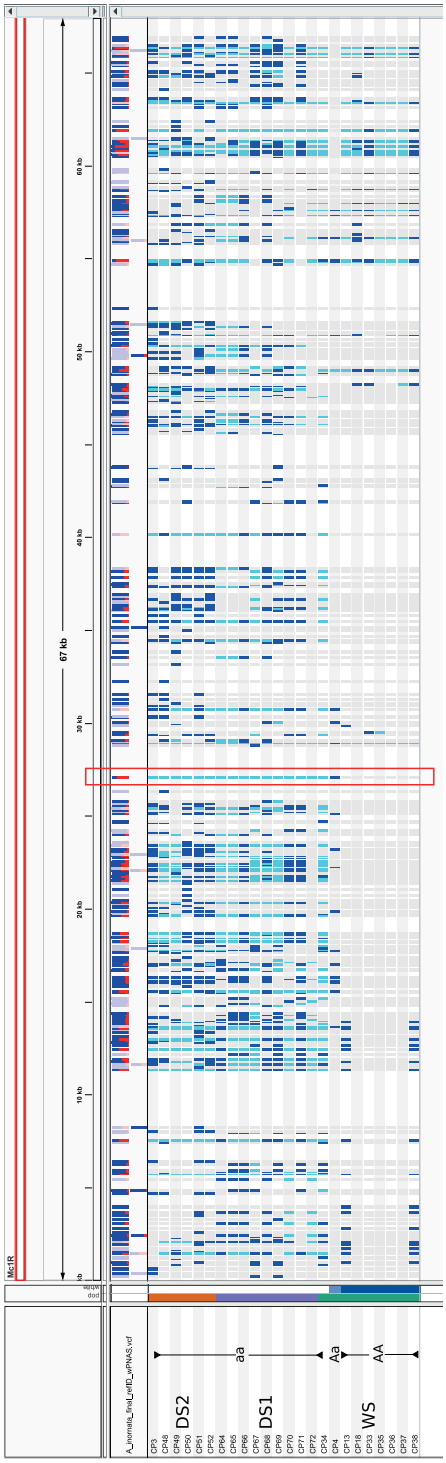


Figure S6: Graphical representation of the genetic variation in the contig encompassing the *Mc1r* gene in *A. inornata*. The colors represent the genotypes of every site that passed the quality filter criteria (light blue: homozygote for the alternate allele, dark blue: heterozygote, gray: homozygote for the reference allele). The reference sequence was constructed from a WS individual. Individuals were grouped according to their sampling locations (green: WS; purple: DS1; orange: DS2) and by the number of white alleles they carried at the candidate locus. (white: 0, light blue: 1, dark blue: 2; see left panel of the figure). The red box indicates the position of the candidate mutation reported by Rosenblum *et al.* (2010).

Figure S7

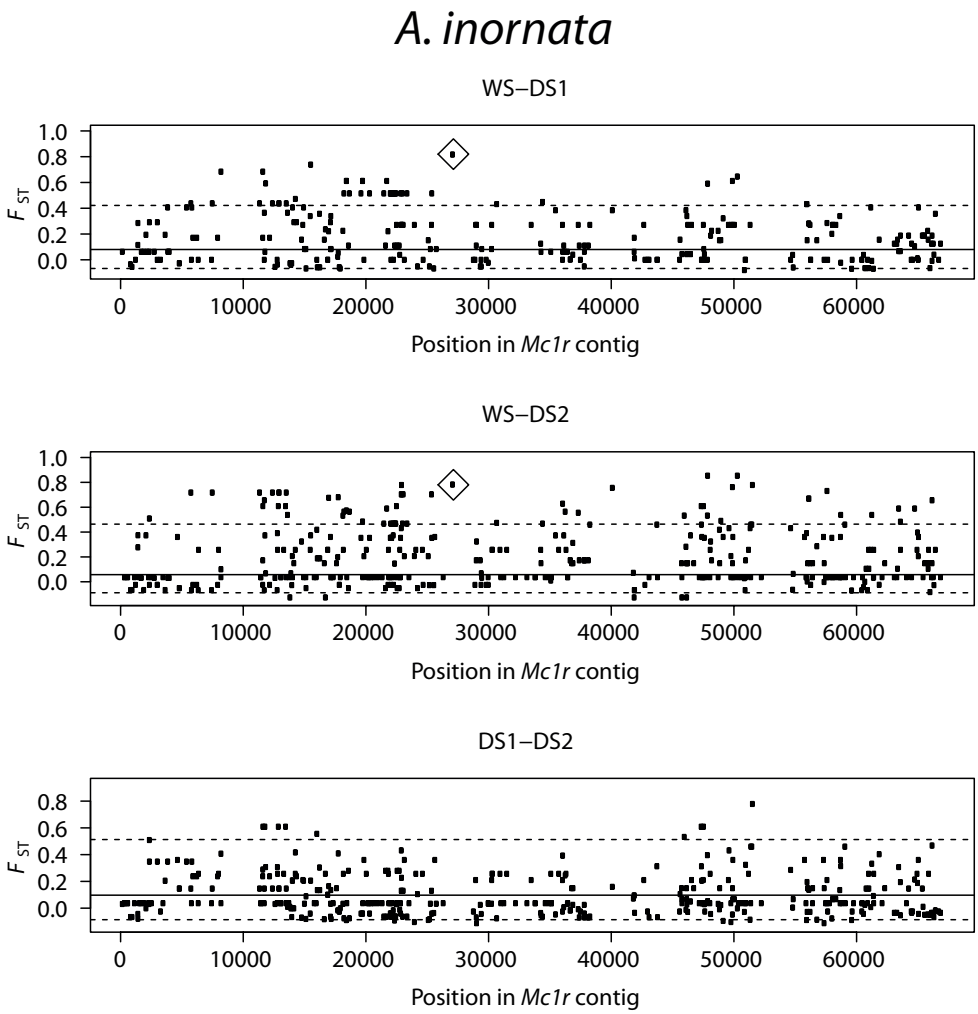


Figure S7: Per site weighted F_{st} values in the *Mc1r* region in *A. inornata* (Weir and Cockerham's F_{st}). The F_{st} value of the candidate mutation reported by Rosenblum *et al.* (2010) is highlighted by a diamond symbol. Figure S8

Figure S8

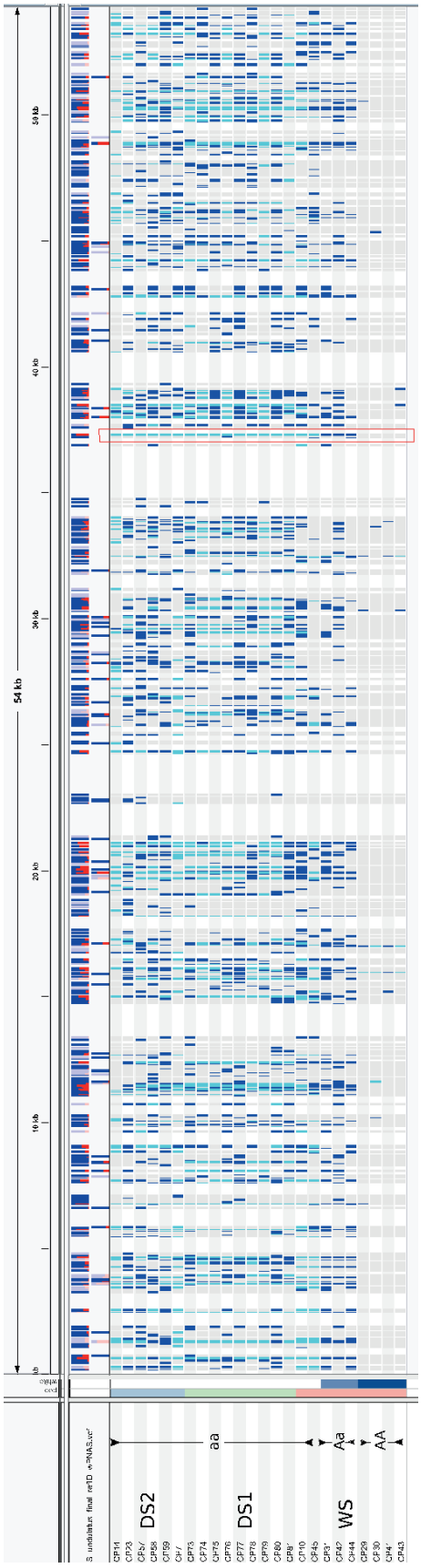


Figure S8: Graphical representation of the genetic variation in the contig encompassing the *Mc1r* gene in *S. cowlesi*. The colors represent the genotypes of every site that passed the quality filter criteria (light blue: homozygote for the alternate allele, dark blue: heterozygote, gray: homozygote for the reference allele). The reference sequence was constructed from a WS individual. Individuals were grouped according to their sampling locations (red: WS, green: DS1, blue: DS2) and to the number of white alleles they carried at the candidate locus. (white: 0, light blue: 1, dark blue: 2; see left panel of the figure). The red box indicates the position of the candidate mutation reported by Rosenblum *et al.* (2010).