MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 5972–5982 doi: 10.1111/mec.12524

FROM THE COVER

On the accumulation of deleterious mutations during range expansions

S. PEISCHL,*†‡ I. DUPANLOUP,*‡ M. KIRKPATRICK† and L. EXCOFFIER*‡

Institute of Ecology and Evolution, University of Berne, 3012 Berne, Switzerland, †Section of Integrative Biology, University of Texas, Austin, TX 78712, USA, ‡Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

Abstract

We investigate the effect of spatial range expansions on the evolution of fitness when beneficial and deleterious mutations cosegregate. We perform individual-based simulations of 1D and 2D range expansions and complement them with analytical approximations for the evolution of mean fitness at the edge of the expansion. We find that deleterious mutations accumulate steadily on the wave front during range expansions, thus creating an *expansion load*. Reduced fitness due to the expansion load is not restricted to the wave front, but occurs over a large proportion of newly colonized habitats. The expansion load can persist and represent a major fraction of the total mutation load for thousands of generations after the expansion. The phenomenon of expansion load may explain growing evidence that populations that have recently expanded, including humans, show an excess of deleterious mutations. To test the predictions of our model, we analyse functional genetic diversity in humans and find patterns that are consistent with our model.

Keywords: expansion load, fitness, gene surfing, mutation load, selective sweep

Received 31 July 2013; revision received 4 September 2013; accepted 5 September 2013

Introduction

The range of most species is fluctuating, a phenomenon that now occurs at an increasing rate owing to rapid climatic changes (e.g. Parmesan 2006; Yamano *et al.* 2011; Pateman *et al.* 2012). Many successful species have increased their range just after speciation (Bokma 2010), and climatic changes can trigger expansions, shifts and contractions for temperature sensitive species (Hewitt 2000; Walther *et al.* 2002; Parmesan & Yohe 2003). Understanding the theoretical and practical implications of dynamic range margins for the ecology, biology and genetics of species has become an important topic in evolutionary biology.

Range expansions can promote the spread of new and standing mutations that happen to be on the wave front of an expansion (Edmonds *et al.* 2004), a phenomenon called *gene surfing* (Klopfstein *et al.* 2006). By surfing on an expanding wave front, a neutral mutation can quickly fix and spread over large regions, mimicking the effect of adaptation to new environments. Surfing is due to enhanced genetic drift on the expansion wave front, where population density is low and growth rate is high. It can lead to clines of heterozygosity from the source to the edge of the expansion (Liu *et al.* 2006; Hallatschek & Nelson 2008; DeGiorgio *et al.* 2011; Slatkin & Excoffier 2012), which has been invoked to explain the decline of genetic diversity with distance from Africa in humans (Prugnolle *et al.* 2005; Ramachandran *et al.* 2005; DeGiorgio *et al.* 2009; Deshpande *et al.* 2009). Surfing is not restricted to neutral mutations: both beneficial and deleterious mutations can also surf (Travis *et al.* 2007; Hallatschek & Nelson 2010; Lehe *et al.* 2012).

Most previous theoretical and empirical studies on the consequences of range expansions have focused on single loci, and little is known about situations when multiple loci segregate. Simulations of a haploid two-locus model with epistasis showed that range shifts can substantially increase the chance of crossing a fitness valley (Burton & Travis 2008), but it remains unclear

Correspondence: Stephan Peischl, Fax: +41 31 631 48 88; E-mail: stephan.peischl@iee.unibe.ch and Laurent Excoffier, Fax: +41 31 631 48 88; E-mail: laurent.excoffier@iee.unibe.ch

how range expansions affect adaptation under a flux of beneficial and deleterious mutations that appear simultaneously throughout the genome.

In this study, we investigate the effect of range expansions on the evolution of fitness in different parts of a species range and on the mutation load. Our model includes both deleterious and advantageous mutations. We use both individual-based simulations and analytical approaches to describe this process. We then examine patterns of genetic variation in human populations and find that they are qualitatively consistent with our predictions.

Material and methods

Simulation models

We model a range expansion as the successive colonization of empty demes located on a one- or two-dimensional lattice. Generations are discrete and nonoverlapping. Migration is homogeneous and isotropic, except that the boundary is reflecting; that is, individuals cannot migrate out of the habitat. Adults migrate between adjacent demes with probability m, and mating within each deme is random. Demes grow logistically such that the expected number of juveniles in the next generation is given by $N^* = \exp(r)N/[1+ (\exp(r)-1)N/K]$, where r is the maximal growth rate, and K is the carrying capacity (Beverton & Holt 1957). Our main focus is on situations in which $m < r$, which is a reasonable assumption for most expanding species, where r is large, as is the case in humans (e.g. Cavalli-Sforza *et al.* 1994), many mammals (Hennemann 1983), insects (Pianka 2000; Frazier *et al.* 2006) and plant species (e.g. Apppendix in Franco & Silvertown 2004). Moreover, population growth rate is favoured during range expansions, such that individuals with high reproductive abilities are positively selected on wave fronts (Burton *et al.* 2010; Phillips *et al.* 2010). The actual number of offspring, N', is drawn from a Poisson distribution with mean N^*. Mating pairs are formed by randomly drawing individuals (with replacement) according to their fitness. Each mating pair produces a single offspring, and the process is repeated N' times, leading to an approximately Poisson-distributed number of offspring per individual.

Individuals are monoecious and diploid. Each gamete carries k_d new deleterious and k_b new beneficial mutations per generation, where k_d and k_b are drawn from Poisson distributions with mean u_d and u_b, respectively. Ignoring neutral mutations, we denote the genomewide mutation rate $u = u_d + u_b$. Mutations are randomly distributed over n independently segregating regions. Within these regions, sites are completely linked, and each new mutation falls on a unique site (infinite-site model). We denote by $\phi_d = u_d/u$, the probability that a new mutation is deleterious, and by $\phi_b = 1-\phi_d$, the probability that it is beneficial. Fitness effects are multiplicative, such that the fitness of an individual is given by $w = \prod_i(1 + s_i)$, where s_i is the selection coefficient associated with the ith mutation at any locus of the focus individual; that is, there is no dominance or epistasis.

Before the onset of the expansion, the five leftmost demes (or five leftmost columns in 2D linear expansions) are at carrying capacity, and all other demes are empty. The expansion starts after a burn-in phase of 10 K generations during which individuals migrate between already colonized demes, but not into new territories. Mean fitness at the edge of the colonized area is then normalized to 1 at the onset of the expansion.

To check the validity of our simulations, we calculated the speed of the wave of advance and found that it was indeed linearly related to $\sqrt{mr}$ (Skellam 1951) (Fig. S17, Supporting information).

Load after a bottleneck

We performed simulations of a bottleneck in a single panmictic population of constant size N. We first let the population evolve for 1000 generations such that mutation load reaches equilibrium. The population size then changed instantaneously to N_B and remained constant for T generations. After that, population size instantaneously changed back to N individuals. Let M_0 denote the expected number of deleterious mutations in a population that did not experience a bottleneck, and M_B is the number of deleterious mutations in a population experiencing a bottleneck. Figure S16 (Supporting information) shows $(M_B-M_0)/M_B$, that is, the relative excess (or deficiency) of deleterious mutations after the bottleneck. Negative values indicate that the total load decreased during the bottleneck.

Human genomic data

Excess of deleterious mutations in non-African populations

Human exomic diversity was inferred from the whole genome of 54 unrelated individuals from 11 populations sequenced by Complete Genomics at a depth of 51-89X coverage per genome (Drmanac *et al.* 2010). We mapped SNPs to 504 378 coding exons of 19 086 autosomal human genes [Ensembl, version 64, September 2011 (Birney *et al.* 2004)]. The derived and ancestral states of the SNPs were inferred from the comparison with the chimpanzee and orangutan genomes, using the syntenic net alignments between hg19 and panTro2 and

between hg19 and ponAbe2, both available on the UCSC platform (Karolchik *et al.* 2012). We then kept the SNPs found to be polymorphic in 42 individuals from three African populations (four Luhya from Webuye, Kenya; four Maasai from Kinyawa, Kenya; nine Yoruba from Ibadan, Nigeria) and from five non-African populations (nine Utah residents with Northern and Western European ancestry from the CEPH collection; four Han Chinese from Beijing; four Gujarati Indians from Houston, Texas, USA; four Japanese from Tokyo; four Toscans from Italy). We predicted the functional consequences of SNP mutations using PolyPhen-2 (Adzhubei *et al.* 2010), and we classified them as being either synonymous, nonsynonymous neutral or nonsynonymous deleterious.

Results

Fitness decreases during range expansions

We first performed individual-based simulations of a range expansion along an unbounded linear habitat. Individuals are diploid, and for simplicity, their genome is structured into n nonrecombining segments that segregate freely. Unless stated otherwise, each segment is subject to both advantageous and deleterious mutations of fixed effect sizes s and $-s$, respectively. Fitness effects are multiplicative such that each mutant allele carried by an individual increases fitness by a factor $1+s$ or decreases fitness by $(1-s)$; that is, there is no dominance or epistasis. We assume in the following that 90% of all non-neutral mutations are deleterious, which seems conservative (Eyre-Walker & Keightley 2007), but our results can be generalized for other distributions of fitness effects (see Figs S1 and S2, Supporting information). Figure 1A shows a representative example of the evolution of the mean fitness during a linear range expansion. We see a very strong difference in the rate of adaptation between core and peripheral populations. Deleterious mutations accumulate faster than beneficial mutations on the wave front, leading to a steady decrease in mean fitness during the expansion. In contrast, recurrent selective sweeps increase the fitness of core populations. This difference between the periphery and core generates a cline in fitness whose steepness increases over time. By tracking the spatial origin of mutations, we find that a large fraction of the total load is due to deleterious mutations that originated on the wave front (Fig. 1B and Fig. S3, Supporting information). We call this fraction of the mutation load 'the expansion load'.

We performed simulations for a variety of parameter combinations to identify conditions under which expansion load occurs. Expansion load is only prevented when deleterious mutations have a large effect, but not necessarily by increasing migration rates when mutations have a small effect (Fig. S4, Supporting information). Mean fitness at the wave front decreases fastest when migration is small and selection is intermediate (Fig. S4, Supporting information). As expected, the rate at which mean fitness declines for a given combination of parameters decreases with increasing local carrying capacities, which make selection more operational (cf. Fig. S4C,D, Supporting information). Nevertheless, expansion load can still occur for parameter combinations where $Kms > 1$ (as e.g. in Fig. S5, Supporting

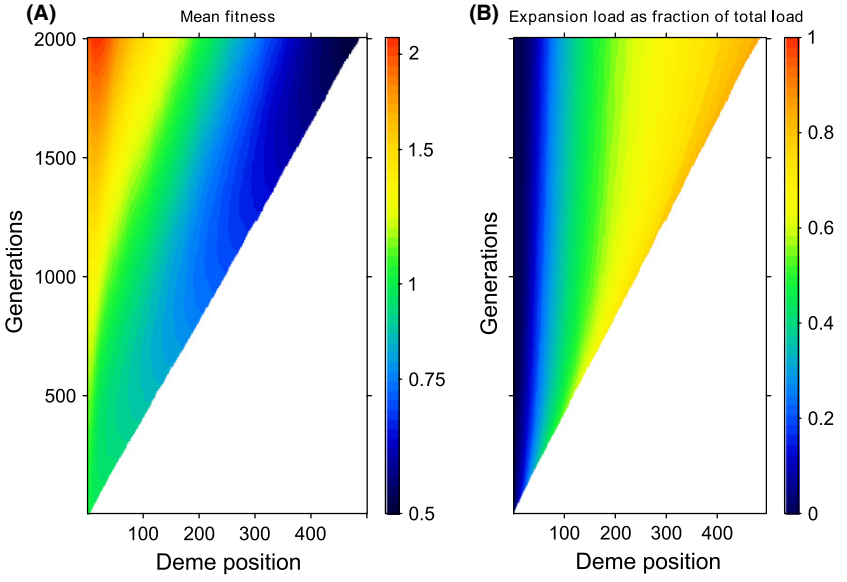

Fig. 1 Evolution of population mean fitness (A) and expansion load (B) during a range expansion. (A) Mean population fitness, normalized to 1 at the onset of the expansion. (B) Fraction of the total load due to mutations originating either directly on the wave front or one deme away from it. Plots are average over 50 simulations. Results are for diploid individuals with $n = 20$ freely recombining regions. The effects of mutations are $s = \pm 0.005$, demes have a carrying capacity of $K = 100$ diploid individuals, populations grow logistically at rate $r = \log(2)$, individuals migrate at rate $m = 0.05$ between adjacent demes, the genome-wide mutation rate is $u = 0.05$, and the fraction of mutations that are deleterious is $\phi_d = 0.9$.

information). While increasing the rate of migration has almost no influence on the build-up of expansion load if selection is weak, it decreases the rate at which fitness declines if selection is intermediate or strong (Fig. S4, Supporting information). In summary, expansion load is expected to occur for selection coefficients of up to a few per cent and local carrying capacities of several hundred to thousands of individuals per deme (see Figs S4 and S5, Supporting information).

Expansion load affects a large part of the species range

While the expanding wave front represents at any time only a tiny part of a species' entire range, most populations of an expanding species descend from ancestors who reproduced at the front (Moreau *et al.* 2011). Consequently, expansion load can represent a substantial part of the total load for most of the demes of an expanding species. For instance, in the case presented in Fig. 1B, more than 50% of the total load is due to mutations that originated at the wave front in territories colonized 300 generations after the onset of the expansion.

We can better appreciate this phenomenon by looking separately at the distribution of deleterious and beneficial mutations in space. Figure 2 shows that the number of deleterious mutations per deme increases almost linearly along the expansion axis, whereas the number of beneficial mutations decreases with distance from the origin of the expansion. This pattern is quite stable over time as similar gradients in deleterious and beneficial mutations are observed 1000 and 2000 generations after the onset of the expansion (Fig. 2). The slow purging of deleterious mutations suggests that they are often at high frequency within demes.

Although these results pertain to expansions in an unbounded range, a very similar pattern occurs on a finite habitat that is 200 demes wide colonized in about 800 generations (Fig. S6, Supporting information). The fitness clines visible for 1000 and 2000 generations after the onset of the expansion are very similar to those in an unbounded habitat, and the clines even persist 5000 generations after the onset of the expansion (Fig. S6, Supporting information). The main difference with an unbounded expansion is in the presence of an edge effect on the right-hand side of the habitat because the influx of beneficial mutations into the rightmost demes is reduced.

Analytical approximations

To better understand how expansion load builds up, we derived an analytical approximation for the expected change of mean fitness at the wave front in a simplified expansion model (see *Supporting Information*). Our approximation rests on the assumption that the migration rate m is small relative to the growth rate r, which as noted above seems reasonable for most expanding species. Using standard diffusion approximations (e.g. Kimura 1964), we find that the probability of fixation of a new mutation at the wave front is

$$p \approx \frac{\exp(-2s_e) - 1}{\exp(-4s_eK) - 1}, \qquad (1)$$

where $s_e = sm\ \log(2/m)/(2r)$ is an effective selection coefficient. Note that eqn (1) is equivalent to the probability of fixation of a mutation with effect s_e in a single panmictic population of constant size K (Kimura 1964) (see Fig. S7, Supporting information). The fact that $|s_e| < |s|$ (Fig. S8, Supporting information) shows that selection is always less efficient in expanding than in stationary populations, which is because the expansion process increases the strength of drift at the wave front. Deleterious mutations will thus become more readily

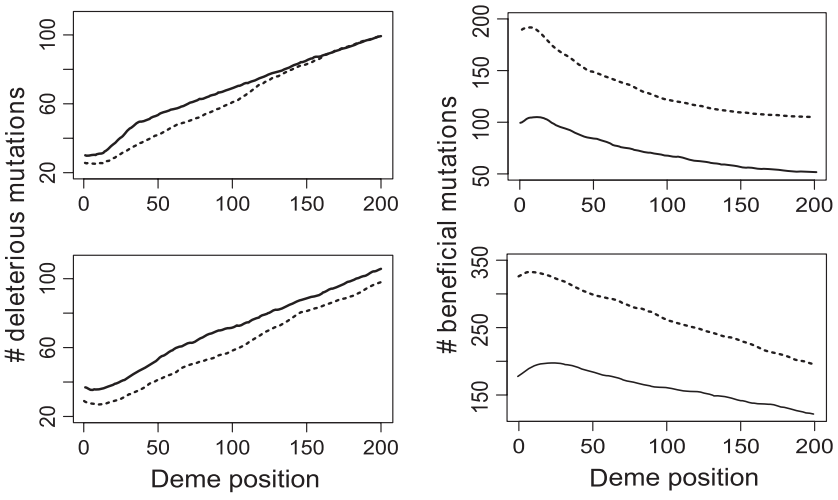

Fig. 2 Spatial distribution of deleterious (left) and beneficial (right) mutations during linear range expansions. Solid lines show the average number of deleterious and beneficial mutations per individual over the first 200 demes of a 1D species range (cf. Fig. 1) that has been colonized from the five leftmost deme either 1000 generations (top) or 2000 generations (bottom) after the onset of the expansion. Dashed lines show the mean number of mutations during 2D linear expansions in a 10 × 200 deme habitat. Parameter values are as in Fig. 1. Note the different scales in the top right and lower right panel.

established at the wave front, and beneficial mutations will sweep less often than in the core, contributing to the building up of a mutation load at the front. It also follows from eqn (1) that the efficiency of selection at the wave front decreases with increasing growth rate and increases with increasing migration rate (Fig. S8, Supporting information). Higher growth rates accelerate the expansion and hence the strength of drift at the wave front (Hallatschek & Nelson 2008). Increasing migration rates decrease the severity of founder effects and of drift during the colonization of new demes.

Equation (1) can also be used to infer the change in mean fitness at the wave front, by assuming that it is due to the serial fixation of independent mutations. Unless selection is strong (i.e. $Ks >> 1$), our analytical approximation is in excellent agreement with individual-based simulations carried out under a more complex expansion model in one dimension (relative error of <5%, see Figs S1, S2 and S9–S12, Supporting information), but also in 2D expansions (see Fig. 3 below). Overall, we find that our analytical approximation is conservative in the sense that it tends to underestimate the parameter region in which expansion load occurs (cf. Fig. S4A,C and B,D, Supporting information).

Expansion load in two-dimensional expansions

We find that deleterious mutations accumulate at the same average rate during linear or radial expansions in 2D habitats than in 1D expansions. Indeed, as shown on Fig. 3, the evolution of the mean fitness at the wave front in linear or radial 2D expansions is remarkably similar to the 1D case. This suggests that 1D analytical theory for the evolution of mean fitness at the wave front can be applied to 2D expansions. The variance in fitness at a given distance from the source of the expansion increases, however, with the width of the front (Fig. 4). Mean fitness can even increase in some parts of the wave front during early stages of the expansion before it eventually decreases over time (Fig. 4). We also observe that in 2D habitats, populations recover more quickly from expansion load (i.e. deleterious mutations are purged more quickly, compare solid and dashed lines in Fig. 2) than after a 1D expansion. This difference is probably due to several causes. First, different deleterious mutations accumulate in different parts of the 2D wave front and single deleterious mutations rarely fix on the whole wave front (Fig. S18, Supporting information), unlike in 1D expansions. Second, stationary 2D-structured populations have a larger effective size making selection more efficient after the expansion (Whitlock & Barton 1997). Indeed, because the total population size is larger in the 2D model, the influx of beneficial mutations and hence the total number of established beneficial mutations increases relatively to 1D expansions (see Fig. 2). Nevertheless, despite a faster

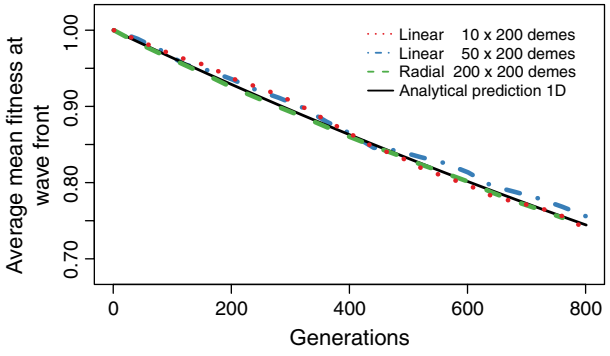

Fig. 3 Decline of mean fitness at the wave front for different expansion types. The figure shows the average mean fitness on the whole wave front (calculated from 20 simulation runs). Parameter values are $K = 100$, $r = \log(2)$, $u = 0.05$, $\phi_d = 0.9$, and $n = 20$.

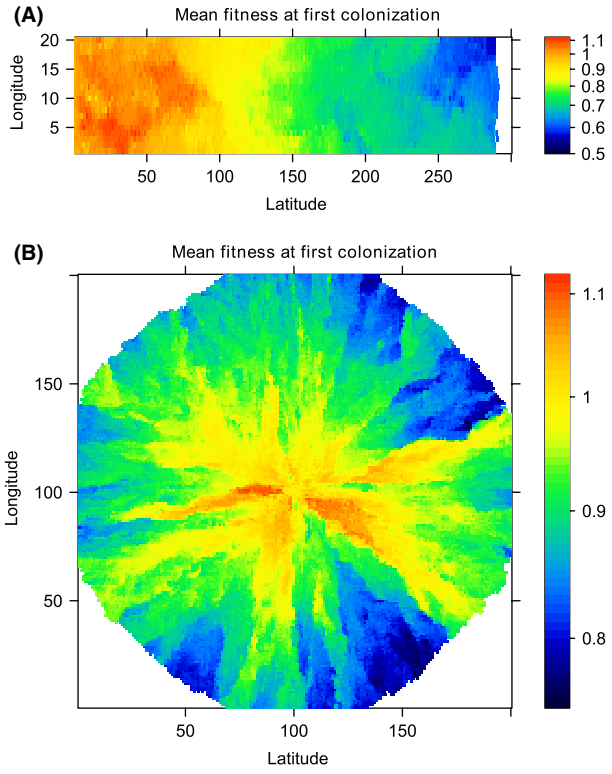

Fig. 4 Evolution of the mean fitness at the wave front during an expansion in a 2D habitat. Mean fitness is shown for the generation at which each deme is colonized. (A) Linear expansion starting from the five leftmost columns of demes, assumed to be at carrying capacity. (B) Radial expansion starting from 25 demes at carrying capacity arranged in a 5×5 square in the middle of the habitat. In both cases, expansion happens after a burn-in phase of 1000 generations; parameter values are as in Fig. 1 ($K = 100$, $r = \log(2)$, $u = 0.05$, $\phi_d = 0.9$, and $n = 20$).

recovery, expansion load can still be visible for thousands of generations after the expansion has stopped (see Figs S13B and S14, Supporting information).

Evidence for expansion load in humans

Non-African populations show an excess of deleterious mutations. There is a mounting evidence that human populations that expanded out of Africa carry an excess of rare deleterious mutations (e.g. Lohmueller *et al.* 2008; Breen & Kondrashov 2010; Li *et al.* 2010; Subramanian 2012). We ask here whether this pattern might result from an expansion load as predicted by our model. To address this question, we analysed the autosomal exomic diversity in 17 Africans and in 25 non-Africans sequenced at high coverage (>50X (Drmanac *et al.* 2010), see *Material and Methods* for details). In agreement with a previous analysis of a smaller representation of functional human diversity (Lohmueller *et al.* 2008), we find that the proportion of deleterious mutations in coding-regions is significantly higher in non-Africans (Table 1). This excess is particularly strong when focusing on private derived alleles, which are likely to have arisen recently. More than 27% of private alleles in non-African populations are predicted to have deleterious effects, as compared to only 21% in Africans ($P < 0.001$, based on a permutation test). These results are in agreement with our model, which suggests that many of the deleterious mutations in Europeans arose during the range expansions out of Africa.

An alternative explanation for the observed pattern would be an excess of rare genetic variants caused by a recent explosive human population growth (Fu *et al.* 2012; Keinan & Clark 2012; Nelson *et al.* 2012). However, the excess of deleterious mutations observed in large samples was mainly due to very rare or even private deleterious mutations, whereas we find that the excess of deleterious mutations in non-Africans is not restricted to rare variants (Table 1). Another explanation for the larger proportion of deleterious alleles outside Africa would be the occurrence of a bottleneck during the exit out of Africa (Lohmueller *et al.* 2008). Our simulations of a single bottleneck in an unsubdivided population confirm that it can increase the proportion of deleterious mutations (Fig. S16, Supporting information), but this bottleneck needs to be unrealistically severe to show levels of mutation load that are commonly obtained after a relatively short range expansion (compare Fig. 1B and Fig. S3 with Fig. S16, Supporting information). This is in keeping with the results of Lohmueller *et al.* (2008) who showed that the excess of deleterious mutations observed in non-Africans can only be obtained with a bottleneck lasting more than 7500 generations (>150 ky), which appears quite unrealistic. Moreover, we note that the observed decrease in heterozygosity with distance from Africa is incompatible with a single bottleneck, but rather in line with a range expansion out of Africa (Prugnolle *et al.* 2005; Ramachandran *et al.* 2005; DeGiorgio *et al.* 2009; Deshpande *et al.* 2009).

Discussion

Evolution during range expansions is often viewed as a process in which species evolve adaptively as they encounter new environments (Barrett & Schluter 2008; Pritchard *et al.* 2010). While adaptation certainly plays a role, there is evidence that deleterious mutations are also established (Lohmueller *et al.* 2008; Hernandez *et al.* 2011; Charlesworth 2012). We show here that genetic drift on the front of range expansions can lead

Table 1 Proportion of exomic sites with deleterious mutations in humans

Mutation frequencies	All sites Africans		All sites non-Africans		Private sites Africans		Private sites non-Africans		Shared sites		All sites Africans vs. non-Africans	Private vs. all sites	
		Sign.		Sign.		Sign.		Sign.		Sign.	Sign.	Africans	Non-Africans
All	0.166	***†	0.185	***‡	0.210	***†	0.274	***‡	0.104	***†	***	***§	***¶
Rare (≤1/34)	0.231	**†	0.282	***‡	0.238	***†	0.298	***‡	0.171	***†	***	***§	NS
Common (>1/34)	0.131	***‡	0.127	NS	0.177	***†	0.218	*‡	0.093	***†	***	***§	**¶

NS, not significant.
Significance levels are obtained by 1000 permutations of individuals between African and non-African groups: ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.
†Significantly smaller proportion of deleterious mutations than in a randomly mixed group of Africans and non-Africans.
‡Significantly larger proportion of deleterious mutations than in a randomly mixed group of Africans and non-Africans.
§Smaller excess of deleterious mutations at private sites than expected by chance in a randomly mixed group of Africans and non-Africans.
¶Larger excess of deleterious mutations at private sites than expected by chance.

to a steady and long-lasting accumulation of deleterious mutations over most of a species range, a phenomenon we call 'expansion load'. This load develops under quite general conditions, such as large local carrying capacities (>1000 individuals, Fig. S5, Supporting information), large selection coefficients (up to several per cent, Fig. S4, Supporting information), large migration rates (>25%, including cases where $Km > 1$, Fig. S4, Supporting information), long distance dispersal (Fig. S9, Supporting information) or alternative distributions of fitness effects (Fig. S1, Supporting information). Although the exact shape of the distribution of fitness effects clearly impacts the evolution of fitness at the wave front, the most important assumption of our model is that deleterious mutations occur frequently relative to beneficial mutations. For instance, we can use eqn (S4) in Supplementary Information to find the minimum proportion of selected mutations to be deleterious to still have expansion load. For the parameter values used in Fig. 1, we estimate that fitness will decrease along the expansion axis if more than 57% of selected mutations are deleterious (see Figs S2 and S15, Supporting information).

Stronger migration usually decreases the rate at which mean fitness declines because bottlenecks during colonization events are less severe, and peripheral populations are less isolated from the core. Note, however, that even very large migration rates between neighbouring demes (e.g. $m > 0.2$, Fig. S4, Supporting information) cannot completely prevent an expansion load from developing unless the effect of individual mutations is strong (e.g. $s > 0.02$, Fig. S4, Supporting information), and local deme carrying capacities are large ($K > 200$, Fig. S4, Supporting information). This implies that unless species have large local carrying capacities (as with some invertebrates and microbes) or beneficial mutations with large effects are common, they are very likely to be affected by expansion load, and their fitness will decrease during the expansion. This decline is fastest when mutations have intermediate fitness effects (s between 1% and 5%, depending on parameter values, Fig. S4, Supporting information). Although the expansion load is progressively eliminated by selection in the range core, this process can be quite slow. Consequently, expansion load can linger as a major component of the mutational total load for thousands of generations (Figs 1B and 2, Figs S3 and S6, Supporting information). Note that these predictions also hold for linear or radial 2D expansions (Fig. 2, Figs S13B and S14, Supporting information), despite the overall larger variance in fitness seen in 2D expansions (Fig. 4).

We derived a simple analytical approximation for the probability of fixation of mutations at the wave front (eqn 1). This approximation is based on the assumption that demes evolve independently at the wave front, which is justified if $m \ll r$. We find, however, that the approximation is accurate (relative error <5%) for a wide range of migration rates ($m > 0.25$, $r = \log(2)$, Fig. S12, Supporting information). If $m > r$, however, our approximation will be inaccurate because the dynamics on the wave front is diffusive rather than discrete. In this case, we expect other approaches based on continuous space approximations to be more useful to study such cases analytically (as, e.g., in Barton *et al.* 2013; Brunet *et al.* 2007; Neher & Hallatschek 2013; Nullmeier & Hallatschek 2013). Although continuous space models are well understood, the derivation of closed forms solutions for fixation probabilities (like our eqn 1) is mathematically challenging. For instance, most results request exact estimates of the shape and speed of the wave of advance, which are known only for a few special cases (e.g. very large, but finite local carrying capacities, Mueller *et al.* 2011), which implies that these parameters need to be estimated from simulations in other cases (e.g. Hallatschek & Nelson 2008; Barton *et al.* 2013). In general, we still expect that expansion load occurs if $m > r$ because densities can become extremely low at the very tip of the wave front, and (rare) fluctuations (Brunet *et al.* 2007) may lead to the establishment of deleterious variants. If $r < s$, however, beneficial mutations (or wild-type alleles) that establish in the wake of the wave will expand and eventually catch up with the moving wave front (Hallatschek & Nelson 2008), which would prevent, or at least mitigate, the expansion load.

Evidence for expansion load

Even though the higher proportion of deleterious mutations observed outside Africa is in line with an expansion load in those populations, a better check of our predictions in human populations would consist in examining the correlation between the extent of mutation load and distance from Africa. However, this would require high-quality and nonascertained genomic data from a collection of well-chosen populations. The sequencing of HGDP samples (Cann *et al.* 2002) would seem ideal, as these populations are more numerous and less admixed and cover more continents than the current 1000 genome population samples (Abecasis *et al.* 2012). We note also that the genomic study of human populations having recently expanded and showing evidence for an increased rate of rare diseases [e.g. in Finland (Norio 2003) and Quebec (Laberge *et al.* 2005)] could also potentially reveal signals of expansion load.

Fewer genomic resources are available in other species than humans, which makes it harder to detect expansion load, but the solitary bee *Lasioglossum leucozonium* that recently expanded to North America might

be a potential example. Invasive populations of this species indeed carry mutations that reduce population growth (Zayed *et al.* 2007). Another potential case comes from the flowering plant *Mercurialis annua*. Populations that recently invaded the Iberian Peninsula from North Africa interestingly show no evidence of inbreeding depression, apparently because mildly deleterious mutations were fixed during the range expansion (Pujol *et al.* 2009). A third suggestive observation is that populations of several invasive species sometimes suddenly collapse without clear explanations [see, e.g., Simberloff & Gibbons (2004) for a review of 17 such cases]. Our results suggest that those extinctions could have a genetic basis, a hypothesis that could be tested.

What might prevent expansion load?

Many species have successfully colonized new areas, suggesting either that it is possible to escape the negative consequences of expansion load, that the reduction in fitness has not been severe enough to stop the expansion or that expansion load has been compensated by beneficial mutations. Several factors not taken into account in our analyses might mitigate the load. Our model assumes that mutations affect fitness multiplicatively. Negative epistatic interactions among deleterious mutations could increase the purging of deleterious mutations and so mitigate the expansion load (Maisnier-Patin *et al.* 2005). Similarly, nonrandom mating with respect to fitness has been shown to increase the efficiency of selection against deleterious mutations (Whitlock & Agrawal 2009). Allee effects (lower growth rate at low densities) have been shown to preserve genetic variation at the front of expanding populations (Roques *et al.* 2012) and thus increase effective population size on the front and lower the severity of expansion load. There are several examples of successful invasions resulting from multiple introductions, often from different source populations (Colautti *et al.* 2006; Simberloff 2009), suggesting that admixture could be a factor limiting the negative consequences of expansion load.

Our results show that fitness variance at the wave front increases with the width of the wave front in 2D expansions (see Fig. 4). Despite this spatial genetic variation at the wave front, it is remarkable that mean fitness at the wave front decreases at essentially the same rate as in 1D expansions (see Fig. 3). This quite unexpected result implies that individual demes on a 2D wave front behave essentially independently from their neighbours, such that the wave front dynamics is essentially one dimensional. Simulations of the dynamics of heterozygosity during 2D expansions show that this is the case, as heterozygosity is virtually nil within demes at the wave front (Fig. S18, Supporting information), implying that mutations are usually fixed within demes on the wave front, like in 1D expansions. However, as can be seen in Fig. S18 (Supporting information), where we show within deme heterozygosity and total wave front heterozygosity during 2D expansions, single mutations (deleterious or beneficial) are usually not fixed in all demes of the wave front. Therefore, due to migration between neighbouring populations in the wake of the wave front, heterozygosity is restored much more quickly in 2D than in 1D expansions (Fig. S18, Supporting information). In our model with multiplicative fitness over loci, the average fitness of the population is, however, little affected by this subsequent 2D gene flow. This is because the fitness of an individual depends on the absolute number of deleterious and beneficial mutations it carries (and their fitness effects), and the homozygote or heterozygote status of these mutations does not matter. This explains why the fitness dynamics of 2D expansions is very similar to that of 1D expansions despite the fact that deleterious alleles have little chance to establish on the wave front in 2D expansions. Note, however, that the fitness is more quickly recovered in the long run in 2D expansions (Fig. 2). Finally, even though we restricted our study to uniform environments, we would posit that the relatively large variance in fitness observed on 2D expanding fronts (Fig. 4) could be useful for invading populations that need to adapt to new environmental conditions.

Conclusions

The accumulation of deleterious mutations during range expansions has been largely unappreciated, perhaps because the main focus of most studies has been on adaptive processes. We show here that many species whose ranges have recently expanded are expected to suffer an increased mutation load. This excess of deleterious mutations is not restricted to the wave front, but affects all the newly colonized species range and can persist for thousands of generations after the end of the expansion. We find that an expansion load can explain a growing body of data from several species, including humans. However, because many species are successful invaders and often perform better in their new ranges (Parker *et al.* 2013), our results do not imply that expanding species are doomed, but would suggest that mechanisms preventing expansion load might have been overlooked and deserve further studies.

Acknowledgements

We thank Daniel Wegmann, Montgomery Slatkin, Sarah Otto, Mike Whitlock, Diogo Meyer, and Loren Rieseberg for their comments and useful suggestions. We also thank Nick Barton

and two anonymous reviewers for their critical assessment. SP was supported by a US NSF Grant DEB-0819901 to MK and by a Swiss NSF Grant 31003A-143393 to LE.

References

Abecasis GR, Auton A, Brooks LD *et al.* (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature*, **491**, 56–65.

Adzhubei IA, Schmidt S, Peshkin L *et al.* (2010) A method and server for predicting damaging missense mutations. *Nature methods*, **7**, 248–249.

Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38–44.

Barton NH, Etheridge AM, Kelleher J, Veber A (2013) Genetic hitchhiking in spatially extended populations. *Theoretical population biology*, **87**, 75–89.

Beverton R, Holt S (1957) *On the Dynamics of Exploited Fish Populations*. Fisheries Investigation Series 2 (19). Ministry of Agriculture, London, *Fisheries and Food*.

Birney E, Andrews TD, Bevan P *et al.* (2004) An overview of Ensembl. *Genome research*, **14**, 925–928.

Bokma F (2010) Speciation and patterns of diversity. *Systematic Biology*, **59**, 365–367.

Breen MS, Kondrashov FA (2010) Mitochondrial pathogenic mutations are population-specific. *Biology Direct*, **5**, 68.

Brunet E, Derrida B, Mueller AH, Munier S (2007) Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, **76**, 041104.

Burton OJ, Travis JMJ (2008) The frequency of fitness peak shifts is increased at expanding range margins due to mutation surfing. *Genetics*, **179**, 941–950.

Burton OJ, Phillips BL, Travis JMJ (2010) Trade-offs and the evolution of life-histories during range expansion. *Ecology Letters*, **13**, 1210–1220.

Cann HM, de Toma C, Cazes L *et al.* (2002) A human genome diversity cell line panel. *Science*, **296**, 261–262.

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The History and Geography of Human Genes*. Princeton University Press, Princeton, New Jersey.

Charlesworth B (2012) The role of background selection in shaping patterns of molecular evolution and variation: evidence from variability on the Drosophila X chromosome. *Genetics*, **191**, 233–246.

Colautti RI, Grigorovich IA, MacIsaac HJ (2006) Propagule pressure: a null model for biological invasions. *Biological Invasions*, **8**, 1023–1037.

DeGiorgio M, Jakobsson M, Rosenberg N (2009) Out of Africa: modern human origins special feature: explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 16057.

DeGiorgio M, Degnan JH, Rosenberg NA (2011) Coalescence-time distributions in a serial founder model of human evolutionary history. *Genetics*, **189**, 579–593.

Deshpande O, Batzoglou S, Feldman MW, Cavalli-Sforza LL (2009) A serial founder effect model for human settlement out of Africa. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 291–300.

Drmanac R, Sparks AB, Callow MJ *et al.* (2010) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, **327**, 78–81.

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

Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, **8**, 610–618.

Franco M, Silvertown J (2004) Comparative demography of plants based upon elasticities of vital rates. *Ecology*, **85**, 531–538.

Frazier MR, Huey RB, Berrigan D (2006) Thermodynamics constrains the evolution of insect population growth rates: "warmer is better". *The American naturalist*, **168**, 512–520.

Fu W, O'Connor TD, Jun G *et al.* (2012) Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature*, **493**, 216–220.

Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. *Theoretical population biology*, **73**, 158–170.

Hallatschek O, Nelson DR (2010) Life at the front of an expanding population. *Evolution; International Journal of Organic Evolution*, **64**, 193–206.

Hennemann WW (1983) Relationship among body-mass, metabolic-rate and the intrinsic rate of natural increase in mammals. *Oecologia*, **56**, 104–108.

Hernandez RD, Kelley JL, Elyashiv E *et al.* (2011) Classic selective sweeps were rare in recent human evolution. *Science*, **331**, 920–924.

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

Karolchik D, Hinrichs AS, Kent WJ (2012) The UCSC genome browser. *Current Protocols in Bioinformatics*, **Chapter 1**, Unit 1 4.

Keinan A, Clark AG (2012) Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science*, **336**, 740–743.

Kimura M (1964) Diffusion models in population genetics. *Journal of Applied Probability*, **1**, 177–232.

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

Laberge AM, Michaud J, Richter A *et al.* (2005) Population history and its impact on medical genetics in Quebec. *Clinical Genetics*, **68**, 287–301.

Lehe R, Hallatschek O, Peliti L (2012) The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS Computational Biology*, **8**, e1002447.

Li YR, Vinckenbosch N, Tian G *et al.* (2010) Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. *Nature Genetics*, **42**, 969–972.

Liu H, Prugnolle F, Manica A, Balloux F (2006) A geographically explicit genetic model of worldwide human-settlement history. *American Journal of Human Genetics*, **79**, 230–237.

Lohmueller KE, Indap AR, Schmidt S *et al.* (2008) Proportionally more deleterious genetic variation in European than in African populations. *Nature*, **451**, 994–997.

Maisnier-Patin S, Roth JR, Fredriksson A *et al.* (2005) Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nature Genetics*, **37**, 1376–1379.

Moreau C, Bherer C, Vezina H *et al.* (2011) Deep human genealogies reveal a selective advantage to be on an expanding wave front. *Science*, **334**, 1148–1150.

Mueller C, Mytnik L, Quastel J (2011) Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Inventiones Mathematicae*, **184**, 405–453.
Neher RA, Hallatschek O (2013) Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 437–442.
Nelson MR, Wegmann D, Ehm MG *et al.* (2012) An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science*, **337**, 100–104.
Norio R (2003) The Finnish Disease Heritage III: the individual diseases. *Human Genetics*, **112**, 470–526.
Nullmeier J, Hallatschek O (2013) The coalescent in boundary-limited range expansions. *Evolution; International Journal of Organic Evolution*, **67**, 1307–1320.
Parker JD, Torchin ME, Hufbauer RA *et al.* (2013) Do invasive species perform better in their new ranges? *Ecology*, **94**, 985–994.
Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics*, **37**, 637–669.
Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.
Pateman RM, Hill JK, Roy DB, Fox R, Thomas CD (2012) Temperature-dependent alterations in host use drive rapid range expansion in a butterfly. *Science*, **336**, 1028–1030.
Phillips BL, Brown GP, Shine R (2010) Life-history evolution in range-shifting populations. *Ecology*, **91**, 1617–1627.
Pianka ER (2000) *Evolutionary ecology*, 6th edn. Benjamin/Cummings, San Francisco, California.
Pritchard JK, Pickrell JK, Coop G (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, R208–R215.
Prugnolle F, Manica A, Balloux F (2005) Geography predicts neutral genetic diversity of human populations. *Current Biology*, **15**, R159–R160.
Pujol B, Zhou SR, Vilas JS, Pannell JR (2009) Reduced inbreeding depression after species range expansion. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 15379–15383.
Ramachandran S, Deshpande O, Roseman CC *et al.* (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15942–15947.
Roques L, Garnier J, Hamel F, Klein EK (2012) Allee effect promotes diversity in traveling waves of colonization. *Proceedings of the National Academy of Sciences*, **109**, 8828–8833.
Simberloff D (2009) The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 81–102.
Simberloff D, Gibbons L (2004) Now you see them, now you don't!–population crashes of established introduced species. *Biological Invasions*, **6**, 161–172.
Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika*, **38**, 196–218.
Slatkin M, Excoffier L (2012) Serial founder effects during range expansion: a spatial analog of genetic drift. *Genetics*, **191**, 171–181.
Subramanian S (2012) The abundance of deleterious polymorphisms in humans. *Genetics*, **190**, 1579–1583.
Travis JMJ, Munkemuller T, Burton OJ *et al.* (2007) Deleterious mutations can surf to high densities on the wave front of an expanding population. *Molecular Biology and Evolution*, **24**, 2334–2343.
Walther GR, Post E, Convey P *et al.* (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.
Whitlock MC, Agrawal AF (2009) Purging the genome with sexual selection: reducing mutation load through selection on males. *Evolution; International Journal of Organic Evolution*, **63**, 569–582.
Whitlock MC, Barton N (1997) The effective size of a subdivided population. *Genetics*, **146**, 427–441.
Yamano H, Sugihara K, Nomura K (2011) Rapid poleward range expansion of tropical reef corals in response to rising sea surface temperatures. *Geophysical Research Letters*, **38**, L04601.
Zayed A, Constantin SA, Packer L (2007) Successful biological invasion despite a severe genetic load. *PLoS ONE*, **2**, e868.

L.E., M.K. and S.P. conceived and designed the study. S.P. wrote the simulation code and analyzed the simulation output. L.E., M.K. and S.P. developed the analytical theory. I.D. analyzed the human data. I.D., L.E., M.K. and S.P. wrote the paper.

Data accessibility

The code used for the simulations and the human SNP data are available on GitHub: https://github.com/CMPG/ADMRE.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Changes in mean fitness on the wave front of an expanding population for different distribution of fitness effects.

Fig. S2 Changes in mean fitness on the wave front of an expanding population for different ratios of deleterious and beneficial mutations.

Fig. S3 Fraction of total load that originated from the wave front (expansion load) for a finite range.

Fig. S4 Change in mean fitness on the wave front.

Fig. S5 Evolution of population mean fitness with large local carrying capacities.

Fig. S6 Spatial distribution of mutations under after a range expansion in a linear habitat restricted to 200 demes.

Fig. S7 Probability of fixation of mutations at the wave front.

Fig. S8 Ratio of the effective selection coefficent at the wave front, s_e, and the actual selection coefficient, s.

Fig. S9 Changes in mean fitness on the wave front of an expanding population with long distance dispersal.

Fig. S10 Changes in mean fitness on the wave front of an expanding population.

Fig. S11 Changes in mean fitness on the wave front of an expanding population for different migration rates.

Fig. S12 Relative error of the analytical approximation [4] as a function of m and s for $K = 100$ (A) and $K = 250$ (B).

Fig. S13 Evolution of population mean fitness and expansion load during a range expansion in a two-dimensional habitat.

Fig. S14 Fraction of total load that originated from the wave front (expansion load) for a linear range expansion in a finite two-dimensional habitat.

Fig. S15 Critical value of as a function of ϕ_d $2Ks_e$.

Fig. S16 Fraction of the total load that is established during a single bottleneck.

Fig. S17 Speed (v) of a linear 1D expansion as a function of $\sqrt{mr}$ (Skellam 1951).

Fig. S18 Heterozygosity at deleterious sites (H_d) during a 2D range expansion on a 10×100 grid.