npg

ORIGINAL ARTICLE

Fitness effects of derived deleterious mutations in four closely related wild tomato species with spatial structure

A Tellier[1], I Fischer[1], C Merino[1], H Xia[1,2], L Camus-Kulandaivelu[1,3], T Städler[4] and W Stephan[1]

[1]Section of Evolutionary Biology, Department Biology II, University of Munich (LMU), Planegg-Martinsried, Germany; [2]College of Horticulture, Northwest A&F University, Shaanxi, China; [3]CIRAD, Montpellier, France and [4]Institute of Integrative Biology, Plant Ecological Genetics, ETH Zurich, Zurich, Switzerland

A key issue in evolutionary biology is an improved understanding of the genetic mechanisms by which species adapt to various environments. Using DNA sequence data, it is possible to quantify the number of adaptive and deleterious mutations, and the distribution of fitness effects of new mutations (its mean and variance) by simultaneously taking into account the demography of a given species. We investigated how selection functions at eight housekeeping genes of four closely related, outcrossing species of wild tomatoes that are native to diverse environments in western South America (*Solanum arcanum*, *S. chilense*, *S. habrochaites* and *S. peruvianum*). We found little evidence for adaptive mutations but pervasive evidence for strong purifying selection in coding regions of the four species. In contrast, the strength of purifying selection seems to vary among the four species in non-coding (NC) regions (introns). Using F_{ST}-based measures of fixation in subdivided populations, we suggest that weak purifying selection has affected the NC regions of *S. habrochaites*, *S. chilense* and *S. peruvianum*. In contrast, NC regions in *S. arcanum* show a distribution of fitness effects with mutations being either nearly neutral or very strongly deleterious. These results suggest that closely related species with similar genetic backgrounds but experiencing contrasting environments differ in the variance of deleterious fitness effects.
Heredity (2011) **107,** 189–199; doi:10.1038/hdy.2010.175; published online 19 January 2011

Introduction

Mutations are the raw material of evolution. To understand the nature of quantitative variation, and thus the potential and speed of adaptation of species to various environments, it is important to determine the positive or negative fitness effects of new mutations. The distribution of fitness effects of new mutations (henceforth denoted DFE) specifies the probability of a new mutation having a given fitness effect. Quantifying the DFE is a key predictor of the potential for adaptation of a population due to mutations with positive fitness effects, as well as a predictor of the decrease in fitness following the accumulation of deleterious mutations (Eyre-Walker and Keightley, 2007). Purifying selection against deleterious mutations increases the proportion of low-frequency alleles, and it reduces the effective population size and, thus, the levels of neutral heterozygosity at linked loci or sites (Charlesworth *et al.*, 1993). Measuring the strength of purifying selection is particularly important for designing strategies to conserve species with small population sizes, and to understand the appearance

and maintenance of low-frequency genetic diseases in humans.

The strength of purifying selection can be measured by the two parameters of the DFE: the mean $E(s)$ and the variance $V(s)$ of the distribution of the selection coefficient s for each new mutation (Eyre-Walker and Keightley, 2007). Using mutation accumulation lines, it has been shown that $E(s)$ differs among species of bacteria and *Drosophila*, as well as between the yeast *Saccharomyces cerevisiae*, the nematode worm *Caenorhabditis elegans*, and the model plant *Arabidopsis thaliana*, taxa characterized by very different genome sizes and genome organization (Martin and Lenormand, 2006b). Moreover, mutations tend to be more deleterious (higher $E(s)$) and less variable (small $V(s)$) in more 'complex' organisms, as defined by their genome size (Martin and Lenormand, 2006b). Interestingly, for a given species, a newly arising mutation will exhibit different fitness effects depending on the environmental (biotic and abiotic) conditions. In other words, the mean mutation effect $E(s)$ is fixed for the species, whereas $V(s)$ may vary across environments. For example, higher variance of the DFE is observed in more stressful environments (Martin and Lenormand, 2006a). These results may also suggest that recently diverged species that occupy different environments might show similar DFE means but different variances of selection coefficients. A key evolutionary question is whether these theoretical expectations are general for all animal

Correspondence: Dr A Tellier, Section of Evolutionary Biology, Department Biology II, University of Munich (LMU), Planegg-Martinsried 82152, Germany.
E-mail: tellier@biologie.uni-muenchen.de

and plant species, as well as for coding and non-coding (NC) regions.

More recent methods to measure the DFE rely on using polymorphism data at synonymous (S) and non-synonymous (NS) sites summarized as the site-frequency spectrum, that is, the allele frequency distribution observed in a population sample. Quantifying directional selection (purifying and positive) using the site-frequency spectrum is a powerful approach, but its effectiveness relies on distinguishing the signature of selection in polymorphism data from that of demographic processes. Indeed, similar patterns of genetic diversity and site-frequency spectra, such as an excess of low-frequency polymorphisms, can occur because of the demographic events (population expansion) or because of purifying selection (Eyre-Walker and Keightley, 2007). Methods to measure selection, thus, attempt to estimate the past demography of species, usually based on S sites, and simultaneously or subsequently compute the effects of selection (positive or negative) on NS and NC sites (Eyre-Walker and Keightley, 2007). Such methods have so far been applied mainly to model organisms for which genome-wide polymorphism data are available (humans, *Drosophila*, *S. cerevisiae* and *Arabidopsis*; Wright and Andolfatto, 2008; Eyre-Walker and Keightley, 2009; Keightley and Eyre-Walker, 2010; but see Gossmann et al., 2010 and Slotte et al., 2010 for recent studies on other plant species).

As a first step, we use simulations to evaluate the robustness of the method of Eyre-Walker and Keightley (2009) for estimating the DFE parameters when only a limited set of 100–300 single-nucleotide polymorphisms (SNPs) and 40–50 sampled alleles are available (for rationale, see Materials and methods). The DFE parameters ($E(s)$ and $V(s)$) can not be accurately estimated with such a low number of SNPs (Keightley and Eyre-Walker, 2010). Using simulated datasets, we show, however, that statistically significant differences can be inferred between the shapes of the DFE for neutral mutations, weakly deleterious mutations and strongly deleterious mutations. We also show that the statistical differences between DFE shapes observed across species are robust under population expansion.

The populations of most (if not all) plant species are spatially sub-structured to some extent. Spatially structured populations with demes connected by migration can lead to patterns of nucleotide diversity dramatically different from those expected in a single panmictic population. For example, the efficacy of positive (negative) selection is affected in spatially structured populations, because drift and migration can counteract the rise (or decrease) in frequency of favorable (unfavorable) alleles. Specifically, the time required for deleterious mutations to be eliminated from a single panmictic population with a given effective population size N_e is shorter than in a structured population with similar N_e. This occurs when gene flow is low, because genetic drift counterbalances the effect of negative selection in demes with small effective size (Whitlock, 2003). Moreover, because selection prevents deleterious mutations from reaching high frequencies, such polymorphisms are mostly private to particular demes rather than shared among subpopulations (Fay et al., 2001; Whitlock, 2003). For a given level of migration and mutation rate, both real genetic differentiation (*sensu stricto* Jost, 2008) and

the traditional fixation index F_{ST}, thus, ought to be higher at sites under purifying selection compared with sites under neutral evolution, assuming linkage equilibrium between sites (Charlesworth et al., 1997). However, for cases in which the effective size per deme, migration among demes and recombination rates are small, the fixation index F_{ST} can be lower under purifying selection compared with neutral evolution (Pamilo et al., 1999). This suggests that (i) very strong purging of deleterious mutations in addition to small recombination rate, and (ii) absence of purifying selection can give similar patterns of F_{ST} because migrants possess a fitness advantage (Charlesworth et al., 1997; Pamilo et al., 1999).

The main empirical objective of this study is to investigate the strength of purifying selection in four closely related wild tomato species: *Solanum peruvianum*, *S. chilense*, *S. habrochaites* and *S. arcanum*. These species are native to western South America, their composite geographic ranges extending from central Ecuador to northern Chile. Collectively, they occupy diverse ecological habitats, with abiotic environments varying from mesic to extremely xeric conditions (Nakazato et al., 2010). These species are proposed to exist as structured populations with many demes (over 100) linked by migration (Arunyawat et al., 2007; Städler et al., 2009; Nakazato et al., 2010). We test the prediction that the mean $E(s)$ of the DFE is identical among habitats and species, but that $V(s)$ differs. We quantify and compare the strength of purifying selection acting on coding regions and on NC (intronic) regions for eight nuclear loci with putatively known housekeeping functions. These housekeeping genes are conserved among the species, and we thus expect strong purifying selection acting on their coding sequences. We also assess whether purifying selection acts on the intronic regions.

The four closely related wild tomato species studied here are characterized by fragmented population structure and various degrees of local adaptation to abiotic conditions (Xia et al., 2010). Thus, our second objective is to investigate how spatial structure of populations affects the efficacy of purifying selection. We quantify the strength of purifying selection using the distribution of deleterious mutations among populations at the eight studied loci, using F_{ST}-based methods (Foll and Gaggiotti, 2008). Finally, we compare the strength of selection inferred at the species level (that is, via the species-wide DFE) with that based on estimates of the fixation index across each structured population of species.

Materials and methods

Plant material and DNA sequencing
Solanum section *Lycopersicon* consists of 13 nominal species found in a relatively small area in western Peru, Chile and Ecuador and includes the domesticated tomato, *Solanum lycopersicum* (formerly *L. esculentum*; Peralta et al., 2008). These species are closely related diploids ($2n = 24$) sharing a high degree of genomic synteny (Ji and Chetelat, 2007). The four studied species, *S. chilense*, *S. peruvianum*, *S. arcanum* and *S. habrochaites* are characterized by spatially structured populations (Arunyawat et al., 2007). For this study, new population samples were collected in central and northern Peru by

T Städler and C Merino in September 2006: Canta, Otuzco, Contumaza and Lajas for *S. habrochaites*, and Otuzco, Rupe, San Juan and Cochabamba for *S. arcanum*. The population samples and geographic locations are summarized in Supplementary Table S1. Voucher specimens have been deposited at the herbarium of the Universidad San Marcos (Lima, Peru). Basic population genetic analyses of nucleotide polymorphism within species and divergence among species will be published elsewhere (CM, AT, WS and TS, unpublished data).

For each sampled population, usually five or six diploid individuals (that is, 10 or 12 alleles) were sequenced at eight unlinked nuclear loci that were previously studied in similarly sized samples of *S. chilense* and *S. peruvianum* (CT093, CT208, CT251, CT066, CT166, CT179, CT198 and CT268; Arunyawat et al., 2007). These loci are single-copy complementary DNA markers originally mapped by Tanksley et al. (1992) in genomic regions with different recombination rates (Stephan and Langley, 1998). The gene products putatively perform key housekeeping functions, and thus purifying selection is suggested to drive their evolution (Supplementary Table S2; Roselius et al., 2005). Genomic DNA was extracted from silica-dried tomato leaves using the DNeasy Plant Mini Kit (Qiagen GmbH, Hilden, Germany). PCR primers were the same as developed for our previous studies, and PCR conditions followed those of our previous studies of the same loci in *S. peruvianum* and *S. chilense* (Arunyawat et al., 2007); PCR primer information can be accessed at http://evol.bio.lmu.de/downloads/index.html.

PCR amplification was performed with High Fidelity Phusion Polymerase (Finnzymes, Espoo, Finland), and all PCR products were examined with 1% agarose gel electrophoresis. Generally, direct sequencing was performed on PCR products to identify homozygotes and obtain their corresponding sequences. For heterozygotes, a dual approach of both cloning before sequencing and direct sequencing was used to obtain the sequences of both alleles. As before, we developed a series of allele-specific sequencing primers whose 3′-end was anchored on identified SNPs or indels (for details of this approach, see Städler et al., 2005). Haplotype phase was thus completely resolved for all sequences. Sequencing reactions were run on an ABI 3730 DNA Analyzer (Applied Biosystems and HITACHI, Foster City, CA, USA). Two alleles were sequenced for each individual, and a total of 39–52 sequences were obtained for each locus × species combination. Contigs of each locus were first built and edited using the Sequencher program (Gene Codes, Ann Arbor, MI, USA) and adjusted manually in MacClade 4 (version 4.06 for OS X, Sinauer Associates, Sunderland, MA, USA). The new sequences for *S. habrochaites* and *S. arcanum* have been deposited in GenBank under accession numbers GU950656–GU951412. In addition, this study also analyzes our previously published sequences sampled from each of four populations in both *S. peruvianum* and *S. chilense*, as well as outgroup sequences from tomato relatives (Baudry et al., 2001; Roselius et al., 2005; Arunyawat et al., 2007). We also included one previously sequenced sample of *S. habrochaites*, which was obtained from the Tomato Genetics Resource Center at UC Davis (http://tgrc.ucdavis.edu; accession LA1775, 'Ancash', see Supplementary Table S1 and Städler et al., 2005).

Basic analyses of the sequence data

For each species, we analyzed polymorphic sites per locus and for all concatenated loci together, using the three categories of sites: S, NS and NC or intronic polymorphic sites. We allowed for multiple hits and polarized nucleotide states as ancestral or derived using either *S. ochranthum* or *S. lycopersicoides* as outgroups (depending on availability; Roselius et al., 2005; Arunyawat et al., 2007). We quantified the strength of purifying selection at the species level by analyzing the pooled samples within species, that is, all four or five populations together, because these should best represent the species-wide diversity (Städler et al., 2009).

Tajima's D (D_T; Tajima, 1989) summarizes the site-frequency spectrum. Statistically significant deviations from zero suggest that a locus has not evolved under neutrality, or that past demographic events have affected the site-frequency spectrum. We also conducted tests based on a comparison of the divergence between two species, taking into account the ratios of S and NS substitutions to the ratios of S and NS polymorphisms (K_a/K_s and π_a/π_s ratios), the McDonald and Kreitman (1991) test and the proportion of adaptive substitutions α (Bierne and Eyre-Walker, 2004). Statistical analyses were performed using DnaSP v. 5.0 (Librado and Rozas, 2009) and the program SITES (Hey Lab, Department of Genetics, Rutgers University).

Partitioning polymorphic sites and the site-frequency spectra

We computed the total number of S and NS sites (including segregating and non-segregating) using the approximate method *yn00* of Yang and Nielsen (2000), as implemented in phylogenetics analysis and maximum likelihood. This method takes into account transition/transversion rate bias and base/codon frequency bias. The size of the coding regions, number of S sites and NS sites are found in Supplementary Table S2, using *S. habrochaites* as an example. All models of sequence evolution used here assume that all sites have the same mutation rate and that no multiple hits occur. We thus corrected for multiple hits in our datasets by calculating the number of substitutions and polymorphisms (S, NS and NC) using the DnaSP conservative criteria (Nei and Gojobori, 1986).

We first calculated the simplified version of the site-frequency spectrum comprising three categories developed by Fay et al. (2001). The minor allele at each polymorphic SNP is called 'rare' if its frequency is <5%, 'intermediate' if the frequency is between 5% and 20%, and 'common' if its frequency is >20%. These categories of SNP frequencies are calculated for the pooled samples (39–60 sequences, depending on the locus and species) separately for S, NS and NC sites. The proportions of polymorphism S*, NS* and NC* are computed for each frequency class (rare, intermediate and common), where asterisk (*) denotes the ratio of the number of SNPs over the total number of sites (Fay et al., 2001). Calculations were made using R scripts (R Development Core Team, 2005).

Under simplified assumptions, NS and NC sites fall into three classes: neutral, slightly deleterious and strongly deleterious (Fay et al., 2001). Neutral NS or NC sites are responsible for all common SNPs in the

site-frequency spectrum and a portion of rare and intermediate SNP classes. Slightly deleterious mutations account for the excess of rare-frequency polymorphism, as well as a small fraction of the intermediate-frequency SNPs. Strongly deleterious mutations are assumed to rarely rise to detectable frequency (Fay et al., 2001). The proportion of NS and NC SNPs in each class is compared with that observed for the S ones. S sites are assumed to be neutral and their site-frequency spectrum thus only determined by past demographic events and population structure. Under purifying selection, an excess of rare-frequency polymorphisms at NS and NC sites is thus expected compared with the proportion of rare-frequency polymorphisms at S sites. On the other hand, similar frequencies of common SNPs are expected for the various classes, reflecting the effectively neutral evolution of common SNPs. The quantities 1–NS*/S* and 1–NC*/S* thus indicate the proportion of sites under selection for NS sites and NC regions, respectively (Fay et al., 2001).

Simulation analyses to estimate the DFE parameters

We aim to estimate the demography, the DFE parameters and α simultaneously for each of the four wild tomato species, using the maximum likelihood method of Eyre-Walker and Keightley (2009). This method is available on PD Keightley's web server (http://homepages.ed.ac.uk/eang33/) and can be summarized as follows. The demographic model is a simple one-step population size change from N_1 (ancestral population size) to $N_2 = N_e$, the present effective population size, assumed to be at equilibrium between mutation, selection and drift. The population expansion ($N_1 < N_2$) or contraction ($N_1 > N_2$) occurred t-generations ago. Each deleterious mutation has a different fitness coefficient s, which is assumed to be drawn from a γ distribution with shape parameter b and mean parameter $N_2E(s) = N_eE(s)$ (Keightley and Eyre-Walker, 2007). It is also assumed that there is a class of neutral sites at which mutant alleles have no effect on fitness. For diploid organisms, the relative fitness of the wild-type genotype, heterozygous mutant and homozygous mutant genotype is 1, $1–s/2$ and $1–s$, respectively (Eyre-Walker and Keightley, 2009). In a second step, the rate of positively selected mutations α is estimated for coding regions. The demography and the DFE parameters are thus used to predict the expected number of substitutions due to deleterious mutations, and the difference between this expected number and the observed number of substitutions yields an estimate of α (Eyre-Walker and Keightley, 2009). In addition, we use the DoFE software available at A. Eyre-Walker's website (University of Sussex, Brighton, UK) (Bierne and Eyre-Walker, 2004), a method based on the MK test (McDonald and Kreitman, 1991) to estimate α. Unlike Eyre-Walker and Keightley (2009), the DoFE method does not take the demography of each species explicitly into account, and NS mutations segregating within a species are assumed to be neutral.

The accuracy of DFE parameter estimates has been tested so far only for datasets with large numbers of SNPs (>1000; Eyre-Walker and Keightley, 2009). We assessed the accuracy of these parameter estimates for datasets mimicking our available wild tomato sequence data. The software SFS_code available at

R. Hernandez website (http://sfscode.sourceforge.net/SFS_CODE) (Hernandez, 2008) was used to simulate a panmictic population with a sample of 20 diploid individuals (that is, 40 chromosomes) sequenced at eight loci of 1250 bp each, using two demographic scenarios: constant population size and a fivefold expansion, the latter occurring $t = 0.8 \times 4 \ N_2$ generations ago. The population mutation parameter of the ancestral population θ_1 varies from 0.005 to 0.02, and 50 simulations with a number of S segregating sites between 100 and 200 were retained for analysis. We simulated 50 independent datasets for a given set of DFE parameter values for the NS mutations with increasing deleterious coefficients ($-N_eE(s) = 5, 50, 500$ and 5000), as well as 50 datasets with only neutral mutations. The shape of the deleterious DFE distribution was fixed at $b = 0.1$. This value is conservative for assessments of power, as for strongly leptokurtic shapes the estimates of the mean and variance of the γ distribution are most inaccurate (Keightley and Eyre-Walker, 2010). Similar simulations were run with a more platikurtic γ distribution ($b = 50$).

Strictly speaking, datasets with only neutral mutations do not have a DFE because 100% of the mutations are in the range $0 < -N_es < 1$. However, in our simulations, datasets with relatively low numbers of neutral SNPs may show departures from the neutral site-frequency spectrum, and thus a DFE is estimated by the method of Eyre-Walker and Keightley (2009) with mutations including $-N_es > 1$.

We recorded the means of estimates and the root mean square error of the ratio of current and ancestral effective population size (N_2/N_1), the time of expansion ($t/4N_2$) and the shape of the DFE γ distribution (b). Root mean square error is the square root of the average squared difference (over $n_{sim} = 50$ datasets) between the estimated value and the simulated value, divided by the simulated value. We also recorded the proportions of mutations in different $-N_es$ ranges ($0 < -N_es < 1$; $1 < -N_es < 10$; $10 < -N_es < 100$; $100 < -N_es$). Multiple mean comparisons were performed for the proportions of mutations in different $-N_es$ ranges as a Tukey's honestly significantly different test (confirmed by a Bonferroni test), as implemented in the R software (R Development Core Team, 2005).

DFE and demography estimates for each tomato species

We concatenated the eight loci using all polymorphism data (shared and private among species) for each species, and estimated the ratio of current and ancestral effective population size (N_2/N_1), the time of expansion ($t/4N_2$) and the shape of the DFE γ distribution (b) for the NS and NC sites. Means and standard errors of parameter estimates were obtained using 50 bootstraps for each site type and each species. This method assumes that all sites are unlinked, which is in broad agreement with previous studies revealing that S. peruvianum and S. chilense exhibit high values of the population recombination parameter (Stephan and Langley, 1998; Arunyawat et al., 2007). The number of SNPs used for estimations is shown in Table 1 and Supplementary Table S7.

Purifying selection and population structure

Assuming linkage equilibrium and that polymorphism at S sites is mainly driven by past changes in population size and population structure, purifying selection is expected to increase the within-species fixation index

Table 1 Multilocus values of Tajima's D_T per species for the pooled samples (arithmetic means across eight loci)

Species	Site category	Number of SNPs	Average D_T
S. peruvianum	Synonymous	278	−0.825
	Silent	616	−0.985
	All sites	742	−1.061
S. chilense	Synonymous	189	−0.227
	Silent	402	−0.301
	All sites	518	−0.527
S. habrochaites	Synonymous	84	−0.322
	Silent	202	−0.526
	All sites	260	−0.757
S. arcanum	Synonymous	171	−0.809
	Silent	377	−0.704
	All sites	457	−0.731

Abbreviation: SNP, single-nucleotide polymorphism.
SNPs are grouped by categories: synonymous, silent sites (that is, non-coding and synonymous) and all sites. The number of SNPs for each class is indicated for each species.

F_{ST} at the NS and NC sites, compared with S sites (Charlesworth et al., 1997; Whitlock, 2003). This is because negative selection changes the depth and shape of the underlying coalescent, reducing the effective population size with effects analogous to the increased rates of genetic drift in smaller populations (for example, reduced nucleotide diversity, higher F_{ST}). However, very strong purifying selection (markedly reduced nucleotide diversity) and linkage disequilibrium can decrease the within-species fixation index F_{ST} at the NS and NC sites, compared with S sites (Pamilo et al., 1999).

We estimated F_{ST} values within species (that is, using all population samples available per species) across all polymorphic SNPs, using the BayeScan program by Foll and Gaggiotti (2008). For each type of site (S, NS and NC), we compared the distribution of F_{ST} values for all polymorphic SNPs within each species. Because of the non-normal distribution of F_{ST} values, non-parametric statistical tests were used to compare the F_{ST} distributions for the different types of sites. The effect of site type (that is, NC, S or NS) on the distribution of F_{ST} was evaluated using a one-way Kruskal–Wallis test. If the effect of site type was significant at the 5% level, we used pairwise Wilcoxon tests to determine which type of site exhibits higher F_{ST} values (R Development Core Team, 2005).

Results

Evidence for purifying selection
The pooled population samples show an excess of low-frequency variants at S sites for S. peruvianum and S. arcanum, indicating marked species-wide expansions (strongly negative D_T; Table 1). A comparison of S and silent sites (S + NC) indicates that NC regions exhibit a slight excess of low-frequency polymorphisms at the species level, except in S. arcanum. Considering all polymorphic sites (S + NC + NS) together yields lower (more negative) D_T values than those for silent sites, indicating that purifying selection acts on coding regions by keeping NS deleterious mutations at low frequency (except in S. arcanum; Table 1; Supplementary Table S3). Methods based on divergence show low power to detect signatures of selection at these loci. The K_a/K_s ratios are lower than one for all species and all loci (Supplementary

Table S5), and loci CT166 and CT208 contain no or very few NS SNPs. However, the McDonald–Kreitman test does not show significant departures from neutrality, except for two marginally significant loci: CT066 in S. habrochaites and CT198 in S. arcanum (Supplementary Table S4). These tests of neutrality indicate that the strength of purifying selection varies between loci (Supplementary Tables S3–S5). Note, however, that the signature is identical across species. For example, loci CT208 and CT166 show a similar lack of NS polymorphism, i.e. evidence for strong purifying selection, in all four species.

Values of NS*/S* and NC*/S* indicate the presence of purifying selection acting on these genes. In coding regions, 85–89% of the sites are under purifying selection in all species (Supplementary Table S6). In introns, 24–30% of the sites are under purifying selection in S. arcanum, S. chilense and S. habrochaites, but this proportion reaches 62% for S. peruvianum (Supplementary Table S6).

We find no evidence of positive selection acting on these eight genes. Both the methods of Bierne and Eyre-Walker (2004) and Eyre-Walker and Keightley (2009) infer the proportion of adaptively driven substitutions (α) to be negative or not different from zero in all species. Indeed, α values estimated by the method of Eyre-Walker and Keightley (2009) are included in the confidence intervals obtained by the DoFE method of Bierne and Eyre-Walker (2004) and are centered around zero (Supplementary Table S8).

Distribution of fitness effects
Estimates from simulated datasets: On the basis of only 100–200 SNPs, demographic and DFE parameters are mostly overestimated by Eyre-Walker and Keightley's (2009) method. In fact, population expansion is inferred in cases with no expansion and the expansion factor is overestimated in simulated expansion scenarios (Table 2). The root mean square error of estimates of the expansion factor increases with increasing deleterious mutation effects (increasing $-N_2E(s)$; Table 2) and the simulated time of the expansion ($t/4N_2 = 0.8$) was correctly estimated only once for $-N_eE(s) = 5$ (Table 2). Therefore, the method of Eyre-Walker and Keightley (2009) cannot be used to derive meaningful demographic estimates for our wild tomato species using the current datasets.

We find that the mean and the shape of the distribution of fitness effects cannot be retrieved with such a small number of SNPs when assuming a shape $b = 0.1$ of the γ distribution. The estimated means ($-N_eE(s)$) are not informative, as they range from very small to extremely large values for all simulated datasets (Supplementary Table S9). Therefore, the shape b of the DFE is always overestimated, except when high $-N_eE(s)$ values are simulated (500 and 5000 in Table 2). The log likelihood ratios for the fit to the data are higher for simulated datasets without population expansion and with low selection coefficients ($-N_eE(s) = 5$ or 50; Supplementary Table S9). For our low number of SNPs, the method of Eyre-Walker and Keightley (2009) has no statistical power to jointly estimate the shape and the scale of the DFE, because they both influence the mean and the variance of the DFE (Keightley and Eyre-Walker, 2010; PD Keightley, personal communication).

Table 2 Results of power analyses for estimates of demographic and DFE parameters

$-N_eE(s)$	Expansion	N_2/N_1	RMSE (N_2/N_1)[a]	$t/4N_2$	RMSE $(t/4N_2)$[a]	b	RMSE (b)
0 (Neutral)	None	4.37 (0.5)	—	4.37 (1.57)	—	8.39 (3.84)	28.42
	Fivefold	4.97 (0.1)	3.57	1.76 (0.5)	3.71	17.93 (5.1)	3.71
5	None	6.63 (0.46)	—	0.8 (0.28)	—	20.68 (5.6)	44.72
	Fivefold	7.08 (0.42)	3.6	0.77 (0.10)	0.73	17 (5.17)	0.73
50	None	6.9 (0.47)	—	1.29 (0.32)	—	17.4 (5.13)	40.18
	Fivefold	5.95 (0.41)	3.07	0.53 (0.08)	0.67	6.06 (3.35)	0.67
500	None	7.03 (0.41)	—	0.5 (0.16)	—	7.89 (3.72)	27.43
	Fivefold	7.6 (0.51)	4.47	1.3 (0.08)	0.73	0.13 (0.02)	0.73
5000	None	6.57 (0.46)	—	0.8 (0.26)	—	0.09 (0.01)	0.07
	Fivefold	8.15 (0.45)	4.48	1.38 (0.1)	0.9	0.095 (0.008)	0.9

Abbreviations: DFE, distribution of fitness effect; RMSE, root mean square error.
Estimated parameters are the ratio of current and ancestral effective population size (N_2/N_1), the time of expansion ($t/4N_2$), and the shape of the DFE γ distribution (b). Means of estimates (standard error) and RMSE were calculated using 50 independent simulated datasets with true parameter values: $-N_eE(s) = 0$, 5, 50, 500 and 5000, $N_2/N_1 = 5$, $t/4N_2 = 0.8$ and $b = 0.1$.
[a]The RMSE of the expansion factor or time of expansion was not calculated in simulations with constant population size.

This means that on the basis of the joint estimates of $-N_eE(s)$ and b, one cannot distinguish neutral mutations from purifying selection if the γ distribution is leptokurtic and skewed toward very strongly deleterious mutations ($b = 0.1$). The shape of the DFE, b, can be estimated for strong purifying selection ($-N_eE(s) > 500$). We observed that estimations of $-N_es$ do not show such extremely large values when the shape of the γ distribution is 50 (data not shown). The mean log likelihood of the estimates is higher than -1300, much higher than that for $b = 0.1$, but joint estimates of the shape and the scale are still unreliable (data not shown).

To circumvent the lack of power to jointly estimate the mean and variance of the DFE, we test the accuracy of summarizing the DFE for comparing distributions with different $-N_eE(s)$ and variances. The DFE can be summarized as four values representing the proportions of mutations falling into given ranges of $-N_e\bar{E}(s)$, as used by Gossmann *et al.* (2010) and Keightley and Eyre-Walker (2010). We show that this summary of the DFE can discriminate between the different purifying selection regimes (Figures 1a and b). The means and standard errors for the estimated proportions of neutral mutations (in the range $0 < -N_es < 1$) and of very strongly deleterious mutations ($-N_es > 100$) are statistically significant between neutral simulations and those with purifying selection (Figure 1). We can thus statistically distinguish between the DFE shape of purifying selection regimes differing by two orders of magnitude: between $-N_es = 5$ and $-N_es = 500$ or between $-N_es = 50$ and $-N_es = 5000$ (Figure 1; Tukey honestly significantly different multiple test at $P < 0.01$). Note also that there is a larger difference in the proportion of neutral mutations ($0 < -N_es < 1$) between the regimes of purifying selection under population expansion (Figure 1b versus 1a). Furthermore, combining results in Figures 1a and b, we demonstrate a significant difference between the DFE shape of $-N_es = 5$ with a fivefold expansion and $-N_es = 500$ without expansion, or between $-N_es = 50$ with expansion and $-N_es = 5000$ without expansion (pairwise t-test, $P < 0.01$). Therefore, despite non-accurate estimation of demographic scenarios, we can distinguish weak purifying selection under population expansion from strong purifying selection in a stationary population.

Estimates for wild tomato species: For NS and NC sites, the DFE has a negative mean for all species ($-N_eE(s)$),
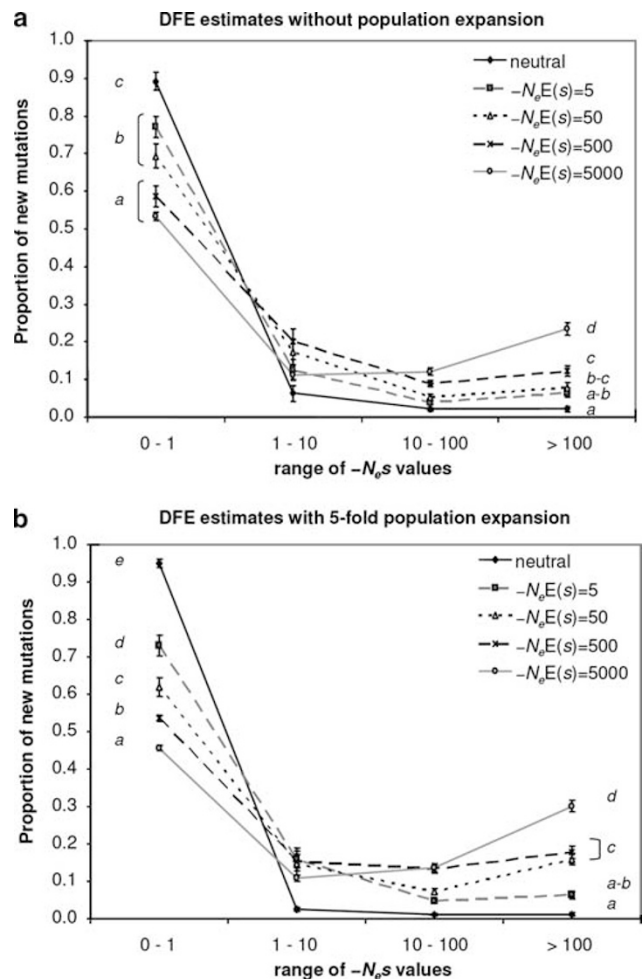

Figure 1 Estimates of the proportions of mutations in different $-N_es$ ranges for simulated datasets with various purifying selection coefficients. (**a**) Under constant population size; (**b**) under a fivefold population expansion, starting in the past at time 0.8 (in units of $4N_2$ generations). Neutral mutations (black solid line with diamonds), $-N_eE(s) = 5$ (long dashed gray line with rectangles), $-N_eE(s) = 50$ (short dashed black line with triangles), $-N_eE(s) = 500$ (long dashed black line with crosses) and $-N_eE(s) = 5000$ (gray solid line with circles). Means ± s.e. are shown for 50 independent simulated datasets for each selection coefficient (means sharing a letter are not significantly different at the 0.01 level according to a Tukey's honestly significantly different test). The demographic parameters are $N_2/N_1 = 5$ and $t/4N_2 = 0.8$, and the shape of the γ distribution (DFE shape) $b = 0.1$.
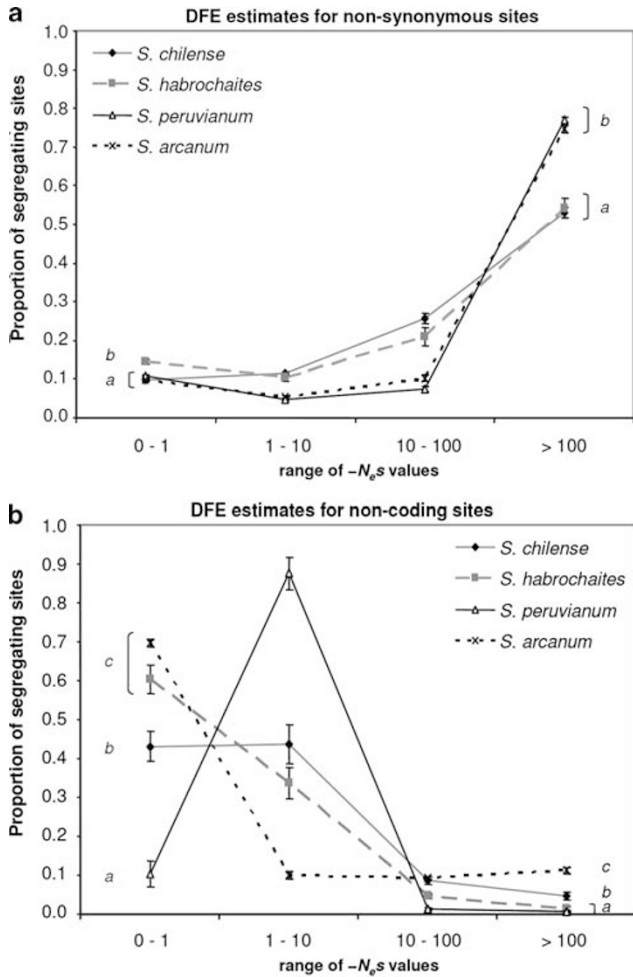
a

b

Figure 2 Proportions of mutations in different $-N_e s$ ranges estimated for the four wild tomato species. (**a**) DFE parameters for non-synonymous polymorphisms; (**b**) DFE parameters for non-coding polymorphisms. *S. chilense* (gray solid line with diamonds), *S. habrochaites* (long dashed gray line with rectangles), *S. peruvianum* (black solid line with triangles) and *S. arcanum* (short dashed line with crosses). Means ± s.e. were calculated for each species using 50 bootstraps (means sharing a letter are not significantly different at the 0.01 level according to a Tukey's honestly significantly different test).

confirming that negative purifying selection is the main force driving the molecular evolution of these loci. Estimates of the shape of the γ distribution show that most species have a leptokurtic DFE distribution, with the notable exception of *S. peruvianum* for NC sites (Table 3, estimated $b \approx 82$). All species exhibit a majority of very strongly deleterious mutations at NS sites, with *S. peruvianum* and *S. arcanum* experiencing stronger purifying selection than the two other species (Figure 2a). The estimates of b and $-N_e E(s)$ do not reveal significant interspecific differences for means and variances of the NS DFE (Table 3; Supplementary Table S10).

The shape of the DFE estimated for the NC (intronic) regions indicates weaker purifying selection than for NS sites. This is demonstrated by the higher estimated proportion of neutral mutations ($0 < -N_e E(s) < 1$ in Figure 2b) and lower proportion of strongly deleterious mutations ($-N_e E(s) > 100$ in Figure 2b), compared with

NS sites (and the much lower estimated $-N_e E(s)$; Supplementary Table S10). In contrast to NS sites, we did not observe excessively high estimates of $-N_e E(s)$, indicating that the DFE estimation is more accurate for intronic sites. Purifying selection on NC regions, however, seems to vary between species. *S. arcanum* shows the highest variance in the DFE with many neutral and weakly deleterious mutations (70%), as well as the highest proportion of very strongly deleterious mutations (15%; Figure 2b). *S. peruvianum*, on the other hand, exhibits a very narrow DFE around $-N_e E(s) = 5.1$ with $b = 82$ (Table 3; Supplementary Table S10), indicating weak purifying selection acting in intronic regions. *S. chilense* and *S. habrochaites* exhibit shapes and means of the DFE that are intermediate between these extremes (Figure 2b; Table 3). The DFE of *S. peruvianum* is significantly different from that of the other three species in all ranges of $-N_e E(s)$ (Figure 2b). Note that, we report the ratios of population sizes (N_2/N_1) and times of expansion (t/N_2) for each species based on S sites and DFE calculations at the NS and NC sites (Table 3). However, following the results from our power analyses (Table 2), we may not interpret these values as robust estimates of past demographic history for these species.

Purifying selection and population structure

Purifying selection increases the proportion of private low-frequency polymorphisms within demes and results in higher intraspecific F_{ST} values at NC and NS sites, compared with those for S sites (Charlesworth *et al.*, 1997). In *S. peruvianum*, F_{ST} values for S sites are lower than for NS sites but not different from those for NC sites (Figure 3a). In *S. chilense* and *S. habrochaites*, all pairwise comparisons show higher F_{ST} values for NC and NS sites compared with S sites, with NC sites being intermediate between lower F_{ST} values at S sites and higher values at NS sites (Figures 3b and c, and Supplementary Information, Section 2). In these three species, the F_{ST}-based results thus agree with the DFE estimates presented above. Purifying selection seems to be very strong in coding regions, and thus F_{ST} for NS sites is higher than for NC (and S) sites. The similar distribution of F_{ST} values for NC and S SNPs in *S. peruvianum* indicates (at best) weak purifying selection on NC sites (Figure 3a). In contrast, *S. arcanum* shows no significant difference between NS and S sites and only a marginally significant difference between S and NC sites (Figure 3d). Furthermore, and contrary to the general expectation outlined above, NS and NC sites show lower mean F_{ST} values in *S. arcanum* than do S sites.

Discussion

The goal of this study was to measure the strength and distribution of purifying selection at eight housekeeping genes in four closely related species of wild tomatoes. Analyzing the species-wide DFE and patterns of F_{ST} within species, we provide evidence for (i) very strong purifying selection acting on NS sites in all species, (ii) weak to strong purifying selection acting on NC regions (introns), (iii) variability in the strength of purifying selection among species in introns and (iv) very low (or non-existent) levels of positive selection on these genes. In the following we discuss how each analysis supports these general conclusions.

Table 3 Estimates of the ratio of current and ancestral effective population size (N_2/N_1), the time of expansion ($t/4N_2$) and the shape of the DFE γ distribution (b) for pooled samples of the four wild tomato species

Species	Non-synonymous sites			Non-coding sites		
	N_2/N_1	$t/4N_2$	b	N_2/N_1	$t/4N_2$	b
S. peruvianum	4.2 (0.27)	0.17 (0.03)	0.16 (0.01)	6.03 (0.44)	0.37 (0.07)	82 (5.3)
S. chilense	8.12 (0.27)	0.04 (0.04)	0.35 (0.02)	7.7 (0.36)	2.48 (0.52)	29.6 (5.99)
S. habrochaites	1.68 (0.28)	8.3 (1.69)	0.26 (0.03)	4.44 (0.42)	0.88 (0.46)	24.7 (5.85)
S. arcanum	5.77 (0.45)	0.36 (0.05)	0.21 (0.02)	3.85 (0.39)	0.38 (0.1)	0.07 (0.021)

Abbreviation: DFE, distribution of fitness effect.
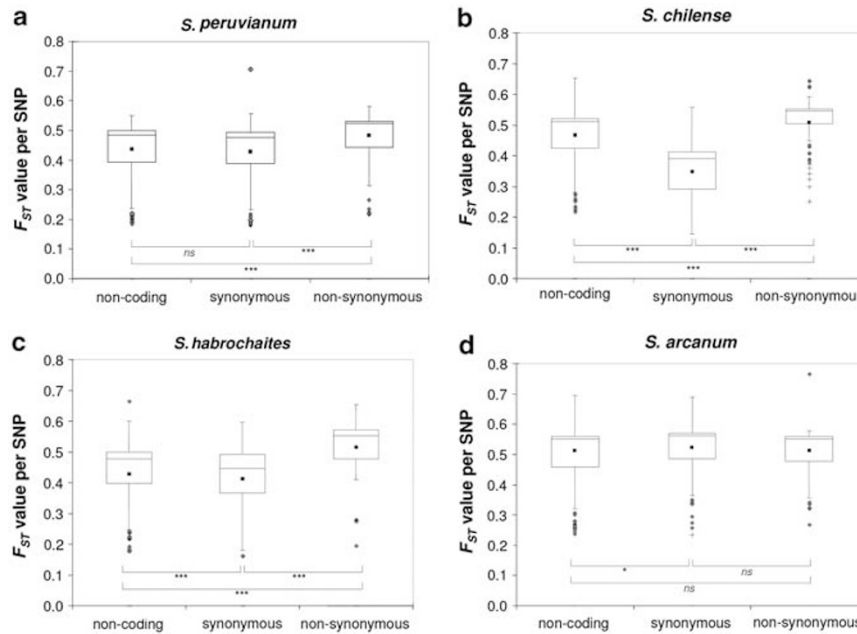The means (standard error) of parameter estimates were calculated using 50 bootstraps.

Figure 3 Boxplots of F_{ST} distributions for non-coding, synonymous and non-synonymous polymorphisms. (**a**) *S. peruvianum*, (**b**) *S. chilense*, (**c**) *S. habrochaites*, (**d**) *S. arcanum*. Whiskers extend from 0.25 to 0.75 quartiles. F_{ST} values that are >1.5 times the interquartile range from the nearest quartile are displayed as diamonds, and for >3 times the interquartile range, they are displayed as crosses. The means of distributions are indicated by full black rectangles. Results of pairwise Wilcoxon tests between the three classes of polymorphisms are indicated below the boxplots as follows: *$P<0.1$; ***$P<0.001$; ns, non-significant (P-values are Bonferroni-corrected by the number of pairwise comparisons).

Estimating the strength of purifying selection

We assessed the effect of purifying selection at the species level by using the pooled sample of all populations for each species. On the basis of site-frequency spectrum analysis (DFE estimates and D_T), purifying selection characterized by a negative mean of the DFE generates an excess of low-frequency polymorphisms (Table 1; Fay *et al.*, 2001; Whitlock, 2003; Keightley and Eyre-Walker, 2007). This explains why we observe a decrease in D_T values when comparing the different types of sites (S, NC and NS; Table 1). Stronger purifying selection on NS sites compared with introns is corroborated by more negative $N_eE(s)$ estimated for the NS DFE in *S. peruvianum*, *S. chilense* and *S. habrochaites* (Figure 2a versus Figure 2b). Such approaches based on the site-frequency spectrum are more powerful to infer parameters of purifying selection, compared with using ratios of polymorphism to divergence between species as in the McDonald–Kreitman test (McDonald and Kreitman, 1991). However, such methods rely on discriminating the signature of selection from that of demographic scenarios

(Fay *et al.*, 2001; Keightley and Eyre-Walker, 2007; Eyre-Walker and Keightley, 2009). In our case, past demographic expansions have created an excess of low-frequency polymorphism at the species level (Städler *et al.*, 2008, 2009), which has to be taken into account when measuring the DFE (Table 3). There is general agreement on the strength of purifying selection between our DFE estimates and the methodologically independent fixation index analyses for three of the four species (F_{ST} for *S. chilense*, *S. peruvianum* and *S. habrochaites*), although inconsistency is found for *S. peruvianum* between the DFE and F_{ST} results and values of Tajima's D_T in NC regions (Table 1).

Note that both the DFE and F_{ST} analyses assume, perhaps unrealistically, complete linkage equilibrium between sites. A theoretical study showed that linkage disequilibrium changes F_{ST} values at S sites compared with NS sites if the recombination rate is five to ten times lower than the mutation rate (Pamilo *et al.*, 1999). Nevertheless, to the extent that purifying selection on NS sites affects the frequency of S polymorphisms due to

linkage disequilibrium (background selection, Charlesworth et al., 1997), our analyses underestimate the strength of purifying selection. This is because S sites would then exhibit a site-frequency spectrum biased toward low-frequency variants (lower Tajima's D_T) and higher F_{ST}, compared with a scenario with complete linkage equilibrium among sites.

For S. arcanum, however, strong purifying selection is inferred at NS and NC sites from the DFE analysis, but F_{ST} estimates for these site classes are lower than for S sites. This result indicates that only neutral mutations that increase in frequency due to drift are shared among populations, whereas there is a lack of private low-frequency polymorphisms. This pattern is also seen in Tajima's D_T, which is similar for S, silent and all sites (Table 1 and Supplementary Table S3). In other words, the S. arcanum site-frequency spectrum does not indicate higher proportions of low-frequency variants at NC and NS sites in excess of that seen at S sites. However, the DFE estimates take into account not only the difference in skew of the site-frequency spectrum between S and NC or NS sites, but also the ratio of segregating sites to potential total sites (NS* or NC*). In fact, when considering the NS*/S* (and NC*/S*) ratios, purifying selection is reflected by a paucity of segregating sites at NS (or NC) sites compared with S sites. Our interpretation, thus, is that purifying selection in S. arcanum is very strong, preventing deleterious mutations from segregating in the populations. As the effective population size of this species is smaller than that of S. peruvianum and S. chilense, one deduces that $-E(s)$ should thus be large in S. arcanum (compared with S. chilense and S. peruvianum). Following results by Pamilo et al. (1999), we suggest that a small effective size, very strong purifying selection and potential linkage disequilibrium in this species might account for the lower F_{ST} at NC and NS sites compared with S sites.

Selection in coding regions
At the eight genes studied here, we find high levels of functional constraint as seen by the elevated deleterious effects of NS mutations. Approximately 90% of the NS sites seem to be under selection (Supplementary Table S6). These loci are housekeeping genes, chosen because they are conserved among species (Roselius et al., 2005), and thus it is unsurprising that they exhibit high levels of purifying selection in coding regions for maintaining their functions. Strong purifying selection was also found in coding regions of other plant species (Gossmann et al., 2010; Slotte et al., 2010).

In coding regions, many new mutations are either nearly neutral or very strongly deleterious. Nearly neutral mutations are expected to segregate in populations with frequencies mainly determined by genetic drift, following the nearly neutral theory (Ohta, 1976). Very strongly deleterious mutations are expected to be eliminated, and thus should not reach high frequency. Recent empirical findings provide evidence for such a shape of the DFE; Fudala and Korona (2009) suggested that it is an intrinsic property of mutations in coding regions to be either nearly neutral, due to the redundancy of protein function in biochemical pathways, or nearly lethal if they hit key functions/amino acids in a pathway/protein. Very strongly deleterious mutations

may occur at active sites of the protein, and nearly neutral ones at functionally less important sites.

There is no evidence for positive selection acting on these genes in the four wild tomato species (α not different from zero; Supplementary Table S8). Two explanations can be proposed. First, the genes were initially chosen because they are conserved between species and mainly encode housekeeping enzymes. As such they do not represent a random sample of the genome-wide DFE for these species. Second, these species exhibit spatial structuring of populations (for example, Arunyawat et al., 2007; Städler et al., 2008). Under population structure, mutations responsible for local adaptation are not considered in species-wide measures of fixed adaptive substitutions. Strong spatial structuring is thus suggested to explain the lack of positive selection found by Gossmann et al. (2010) in various plant species, in contrast to results in Capsella grandiflora, a species with low population subdivision that appears to exhibit high rates of adaptive substitutions (α = 40%; Slotte et al., 2010).

Selection in NC regions
We find that purifying selection also appears to impact NC regions (introns) and its strength to be variable among species. Purifying selection is inferred to be strong in S. arcanum and slightly weaker in S. chilense and S. habrochaites. In S. peruvianum, the estimated DFE indicates weak purifying selection with a small variance of the DFE, corroborated by non-significant F_{ST} differences between S and NC sites. Natural selection on NC regions was found in several species (for example, human and mouse (Jareborg et al., 1999); Arabidopsis thaliana (Thomas et al., 2007) and poplars (Olson et al., 2010)). As intronic regions are known to be of importance in the regulation of expression, splicing and protein synthesis, it is not surprising to find purifying selection on introns, although apparently much weaker than on NS sites (Jareborg et al., 1999).

We verified that our results do not depend on one particular locus being subject to strong selection in the NC regions of S. arcanum. We thus used PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), a database to detect described regulatory elements in DNA sequences, and calculated the number of regulatory motifs disrupted by SNPs or indels. S. arcanum appears to show fewer mutations and indels disrupting motifs than S. chilense, supporting the interpretation that purifying selection is stronger on introns of S. arcanum (Supplementary Information, Section 3). However, none of the loci show significantly high numbers of disrupted motifs across this suite of species.

Ecological differences and selection between species
Comparing the four tomato species with regard to the strength of selection also indicates the adaptive potential of a species and selective constraints imposed by the environment. These four species are found in contrasting environments (for example, mesic/dry, high/low altitude and warm/cold) that generate gradients of abiotic stresses (Nakazato et al., 2010). We note that the inferred DFE for NS sites is similar for the S. peruvianum/S. arcanum and the S. chilense/S. habrochaites pairs (Figure 2a). This significant difference in DFE shape

and/or variance between the pairs of species is not explained, however, by differences in effective population size. Indeed, each pair contains one species with a large N_e (*S. peruvianum* and *S. chilense* with $N_e > 10^6$; Roselius *et al.*, 2005; Städler *et al.*, 2008) and one species with an ~50% lower N_e estimate (*S. arcanum* and *S. habrochaites*, respectively; CM, AT, WS, TS, unpublished data). As argued above, we are also confident that the significant differences in inferred strength of purifying selection between species (Figure 2) are robust to the bias in demographic estimates. For example, the stronger purifying selection in *S. arcanum* compared with other species is not due to a bias in (under-)estimating its expansion factor.

Recent studies suggest that *S. peruvianum* and *S. habrochaites* exhibit more generalist ecological attributes compared with the two other species. Both species have large geographic ranges and wide niche breadth, encompassing diverse habitats ranging from mesic to dry (Nakazato *et al.*, 2010). A screen of candidate genes for abiotic stress response in *S. peruvianum* has found no evidence for molecular signatures of local adaptation (Xia *et al.*, 2010). *S. peruvianum* has the second largest geographic range of the four studied species (Nakazato *et al.*, 2010), the largest effective population size, and exhibits signatures of a marked demographic expansion (Städler *et al.*, 2008). Under such characteristics, geographically varying phenotypic plasticity rather than local adaptation can be promoted (Sultan and Spencer, 2002). We thus propose that lower constraints on the expression of housekeeping genes in *S. peruvianum* correlate with phenotypic plasticity, explaining the apparently relaxed constraints on introns in this species (as for *S. habrochaites*). However, strong functional constraints are evident for the coding regions.

On the other hand, *S. chilense* and *S. arcanum* are found in geographically more restricted areas and characterized by narrower ecological niches (Nakazato *et al.*, 2010). *S. chilense* has been described as a specialist species with particular adaptations to extremely dry environments (Nakazato *et al.*, 2010; Xia *et al.*, 2010). Following our previous argument, we suggest that limited phenotypic plasticity, and thus higher constraints on gene expression of housekeeping genes, may characterize species with narrow ecological distribution (Poot and Lambers, 2008). Compared with *S. peruvianum*, the strong purifying selection inferred for *S. arcanum* indicates that intronic features may have been more conserved in this species. This may suggest that the environment, in which *S. arcanum* is found, exerts some important selective pressures on gene regulation/expression and thus on intronic regions. This could also be due to reduced phenotypic plasticity for gene expression at these genes for a species with quite narrow ecological requirements.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

References

Arunyawat U, Stephan W, Städler T (2007). Using multilocus sequence data to assess population structure, natural selection, and linkage disequilibrium in wild tomatoes. *Mol Biol Evol* **24**: 2310–2322.

Baudry E, Kerdelhué C, Innan H, Stephan W (2001). Species and recombination effects on DNA variability in the tomato genus. *Genetics* **158**: 1725–1735.

Bierne N, Eyre-Walker A (2004). The genomic rate of adaptive amino acid substitution in *Drosophila*. *Mol Biol Evol* **21**: 1350–1360.

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

Charlesworth B, Nordborg M, Charlesworth D (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet Res Camb* **70**: 155–174.

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

Eyre-Walker A, Keightley PD (2009). Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Mol Biol Evol* **26**: 2097–2108.

Fay JC, Wyckoff GJ, Wu CI (2001). Positive and negative selection on the human genome. *Genetics* **158**: 1227–1234.

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

Fudala A, Korona R (2009). Low frequency of mutations with strongly deleterious but nonlethal fitness effects. *Evolution* **63**: 2164–2171.

Gossmann TI, Song BH, Windsor AJ, Mitchell-Olds T, Dixon CJ, Kapralov MV *et al.* (2010). Genome wide analyses reveal little evidence for adaptive evolution in many plant species. *Mol Biol Evol* **27**: 1822–1832.

Hernandez RD (2008). A flexible forward simulator for populations subject to selection and demography. *Bioinformatics* **24**: 2786–2787.

Jareborg N, Birney E, Durbin R (1999). Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. *Genome Res* **9**: 815–824.

Ji Y, Chetelat RT (2007). GISH analysis of meiotic chromosome pairing in *Solanum lycopersicoides* introgression lines of cultivated tomato. *Genome* **50**: 825–833.

Jost L (2008). G_{ST} and its relatives do not measure differentiation. *Mol Ecol* **17**: 4015–4026.

Keightley PD, Eyre-Walker A (2007). Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. *Genetics* **177**: 2251–2261.

Keightley PD, Eyre-Walker A (2010). What can we learn about the distribution of fitness effects of new mutations from DNA sequence data? *Phil Trans Roy Soc Lond* **365**: 1187–1193.

Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Martin G, Lenormand T (2006a). The fitness effect of mutations across environments: a survey in light of fitness landscape models. *Evolution* **60**: 2413–2427.

Martin G, Lenormand T (2006b). A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* **60**: 893–907.

McDonald JH, Kreitman M (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.

Nakazato T, Warren DL, Moyle LC (2010). Ecological and geographic modes of species divergence in wild tomatoes. *Am J Bot* **97**: 680–693.

Nei M, Gojobori T (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* **3**: 418–426.

Ohta T (1976). Role of very slightly deleterious mutations in molecular evolution and polymorphism. *Theor Pop Biol* **10**: 254–275.

Olson MS, Robertson AL, Takebayashi N, Silim S, Schroeder WR, Tiffin P (2010). Nucleotide diversity and linkage disequilibrium in balsam poplar (*Populus balsamifera*). *New Phytol* **186**: 526–536.

Pamilo P, Palsson S, Savolainen O (1999). Deleterious mutations can reduce differentiation in small, subdivided populations. *Hereditas* **130**: 257–264.

Peralta IE, Spooner DM, Knapp S (2008). The taxonomy of tomatoes: a revision of wild tomatoes (*Solanum* section *Lycopersicon*) and their outgroup relatives in sections *Juglandifolium* and *Lycopersicoides*. *Syst Bot Monogr* **84**: 1–186.

Poot P, Lambers H (2008). Shallow-soil endemics: adaptive advantages and constraints of a specialized root-system morphology. *New Phytol* **178**: 371–381.

R Development Core Team (2005). *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing (ed.): Vienna, Austria.

Roselius K, Stephan W, Städler T (2005). The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. *Genetics* **171**: 753–763.

Slotte T, Foxe JP, Hazzouri KM, Wright SI (2010). Genome-wide evidence for efficient positive and purifying selection in *Capsella grandiflora*, a plant species with a large effective population size. *Mol Biol Evol* **27**: 1813–1821.

Städler T, Arunyawat U, Stephan W (2008). Population genetics of speciation in two closely related wild tomatoes (*Solanum* section *Lycopersicon*). *Genetics* **178**: 339–350.

Städler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009). The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics* **182**: 205–216.

Städler T, Roselius K, Stephan W (2005). Genealogical footprints of speciation processes in wild tomatoes: demography and evidence for historical gene flow. *Evolution* **59**: 1268–1279.

Stephan W, Langley CH (1998). DNA polymorphism in Lycopersicon and crossing-over per physical length. *Genetics* **150**: 1585–1593.

Sultan SE, Spencer HG (2002). Metapopulation structure favors plasticity over local adaptation. *Am Nat* **160**: 271–283.

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

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P *et al.* (1992). High-density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**: 1141–1160.

Thomas BC, Rapaka L, Lyons E, Pedersen B, Freeling M (2007). *Arabidopsis* intragenomic conserved noncoding sequence. *Proc Natl Acad Sci USA* **104**: 3348–3353.

Whitlock MC (2003). Fixation probability and time in subdivided populations. *Genetics* **164**: 767–779.

Wright SI, Andolfatto P (2008). The impact of natural selection on the genome: emerging patterns in *Drosophila* and *Arabidopsis*. *Annu Rev Ecol Evol Syst* **39**: 193–213.

Xia H, Camus-Kulandaivelu L, Stephan W, Tellier A, Zhang Z (2010). Nucleotide diversity patterns of local adaptation at drought-related candidate genes in wild tomatoes. *Mol Ecol* **19**: 4144–4154.

Yang ZH, Nielsen R (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* **17**: 32–43.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)