Recent Selection Changes in Human Genes under Long-Term Balancing Selection

Cesare de Filippo,*† Felix M. Key,† Silvia Ghirrotto,‡ Andrea Benazzo,‡ Juan R. Meneu,† Antje Weihmann,† NISC Comparative Sequence Program,§ Genés Parra,† Eric D. Green,§ and Aida M. Andrés*†

1Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
2Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy
3National Human Genome Research Institute, National Institutes of Health, Bethesda, MD

*Corresponding author: E-mail: cesare_filippo@eva.mpg.de; aida.andres@eva.mpg.de.
Associate editor: Ryan Hernandez

Abstract
Balancing selection is an important evolutionary force that maintains genetic and phenotypic diversity in populations. Most studies in humans have focused on long-standing balancing selection, which persists over long periods of time and is generally shared across populations. But balanced polymorphisms can also promote fast adaptation, especially when the environment changes. To better understand the role of previously balanced alleles in novel adaptations, we analyzed in detail four loci as case examples of this mechanism. These loci show hallmark signatures of long-term balancing selection in African populations, but not in Eurasian populations. The disparity between populations is due to changes in allele frequencies, with intermediate frequency alleles in Africans (likely due to balancing selection) segregating instead at low- or high-derived allele frequency in Eurasia. We explicitly tested the support for different evolutionary models with an approximate Bayesian computation approach and show that the patterns in PKDREJ, SDR39U1, and ZNF473 are best explained by recent changes in selective pressure in certain populations. Specifically, we infer that alleles previously under long-term balancing selection, or alleles linked to them, were recently targeted by positive selection in Eurasian populations. Balancing selection thus likely served as a source of functional alleles that mediated subsequent adaptations to novel environments.

Key words: natural selection, environmental changes, out-of-Africa.

Introduction
Natural selection drives the adaptation of populations to their environment (Darwin and Wallace 1858). Balancing selection maintains advantageous polymorphisms in populations and, as a consequence, it increases genetic diversity. This is in contrast to the reduction in diversity that results from favoring the single, most advantageous allele via positive or purifying selection. Mechanisms of balancing selection include overdominance (Allison 1956), frequency-dependent selection (Wright 1939), fluctuating selection (Gillespie 1978), and pleiotropy (Gendzekhadze et al. 2009), although when selection is old the genetic signatures of all these types of selection can be similar (Andrés 2011; Key, Teixeira, et al. 2014). The first signature is an excess of polymorphic over divergent sites. Old selection maintains the advantageous polymorphism and linked neutral polymorphisms longer than expected under neutrality (Wiuf et al. 2004; Charlesworth 2006). This results in an unusual accumulation of polymorphisms that is typically reflected in a local excess of diversity over divergence (Hudson et al. 1987), and the intensity of this signature depends mostly on the age of the balanced polymorphism (Charlesworth 2006). The second signature is a shift in allele frequencies. When a frequency equilibrium (an allele frequency that maximizes fitness in the population) exists, balancing selection maintains the selected polymorphism close to the frequency equilibrium; neutral variants also accumulate at a similar frequency due to linkage, shifting the local distribution of allele frequencies (the site frequency spectrum, SFS) toward the frequency equilibrium (Andrés 2011). For instance, if the frequency equilibrium is 0.5, the SFS is expected to show a shift toward intermediate frequency alleles close to 0.5.

Balancing selection can act for long periods of time. Some polymorphisms persist for millions of years and can even be shared among species as trans-species polymorphisms, which exist in humans (Loisel et al. 2006; Ségurel et al. 2012; Leffler et al. 2013; Teixeira et al. 2015) but are rare (Asthana et al. 2005). Most balanced polymorphisms are present in single species, with the catalog of human candidate targets of balancing selection (Andrés et al. 2009; DeGiorgio et al. 2014; Rasmussen et al. 2014) far surpassing the catalog of trans-species polymorphisms. This is because selection is rarely old and constant enough (for more than 6 My) to create trans-species polymorphisms. Within species, targets of balancing selection are classically assumed to be shared across populations, with unusually low $F_{ST}$ values flagging such cases (Schierup et al. 2000; Bamshad and Wooding 2003; Key, Teixeira, et al. 2014). This builds on the reasonable expectation that selection that has maintained a polymorphism for millions of years (and is thus detectable on the patterns of
linked variation) is likely shared across populations. Yet this is not necessarily the case.

For example, there are several loci where signals of long-term balancing selection are detected in human populations of African origin, but not in populations outside of Africa (Andrés et al. 2009; DeGiorgio et al. 2014). This is an unexpected observation because the balanced polymorphisms are old and predate the out-of-Africa migration, and it raises the question of whether these population differences are explained by drift alone (e.g., during and after the out-of-Africa migration, characterized by population bottlenecks and expansions) or by changes in the selective pressure outside of Africa, on previously balanced loci. Changes in selective pressure might be reasonable because, while the ancestors of modern humans have lived in Africa for millions of years (long adapting to the environment), the colonization of the rest of the world happened only in the last 50,000 years (Gravel et al. 2011). These migrating human populations encountered new environments, and they experienced novel, local adaptations (Cavalli-Sforza 1966; Lewontin and Krakauer 1973; Akey et al. 2004; Coop et al. 2009; Pickrell et al. 2009; Fumagalli et al. 2011) or changes in the strength of selection (Key, Peter, et al. 2014).

Here we aim to explore a model of adaptation where balancing selection turned into positive selection. We purposely focus on a small number of genes in order to perform detailed analyses and computationally intensive inferences that allow us to distinguish between competing models of adaptation. We selected a number of genes previously identified (Andrés et al. 2009) as showing African-specific signatures of balancing selection, as they are prime candidates for having experienced shifts in selective pressure outside of Africa. We analyzed six human populations and confirmed both the African signatures of long-term balancing selection and the absence of these signatures in Eurasians. We investigated the probability that population differences are due to demography or to changes in selective regime, showing that a model where selection changed after the out-of-Africa migration favoring an existing or linked new variant best explains the patterns of genetic variation in three genes. This reveals a shift in selective pressure in previously balanced loci that created genetic differences among human populations.

### Results

**Genes with Signatures of Balancing Selection in Africa**

We initially investigated 14 genes (supplementary table S1, Supplementary Material online) previously shown to have significant signatures of long-term balancing selection in Africa only (Andrés et al. 2009). In addition, we analyzed 49 "control" loci (old, processed pseudogenes) as our proxy for neutrality (see Materials and Methods and supplementary table S1, Supplementary Material online). We produced a combination of Sanger and Illumina-derived sequence data for a total of near 230 kb, and obtained high-quality polymorphism data in the coding and adjacent noncoding regions of these genes from five human populations (each N = 30): Yoruba (YRI) and Luhya (LWK) from Africa, Toscani (TSI) from Europe, and Gujarati (GIH) and Han Chinese (CHB) from Asia. We identified signatures of balancing selection with two neutrality tests: Hudson–Kreitman–Aguadé (HKA) (Hudson et al. 1987) and Mann–Whitney U (MWU) (Nielsen et al. 2009). These tests detect departures from the neutral expectation in the density of polymorphisms and in the SFS, respectively (see Materials and Methods), and significant signatures for both tests are expected only under long-term balancing selection. At the 5% P value cut-off, four genes (CLCNKB [chloride channel, voltage-sensitive kb], PKDREJ [polycystic kidney disease and receptor for egg jelly], SDR39U1 [short-chain dehydrogenases/reductases family 39U member 1], and ZNF473 [zinc finger protein 473]) show both significant excess of polymorphism and significant shifts toward intermediate-frequency alleles in African populations (fig. 1). We note that these signatures are not due to mapping errors or partial duplications (supplementary material section 2, Supplementary Material online). All four genes thus display strong signatures of balancing selection in both African populations and, conservatively, we focused only on these for the remainder of the study.

Outside of Africa, two of the four genes (CLCNKB and PKDREJ) display significant signatures of long-term balancing selection in the European TSI, but no gene shows similar signatures in the Asian populations (GIH and CHB). The differences between African and non-African populations are not surprising because these genes were originally selected for their discordant signatures among human groups (Andrés et al. 2009). But they confirm that these loci are adequate for our purposes.

**Excess of Polymorphism**

As discussed above, all four genes have unexpectedly high levels of polymorphism in African populations (HKA test, in fig. 1). To better understand the distribution of single nucleotide polymorphisms (SNPs) in the genomic region, we extended the analysis to a larger genomic region (400,000 bp centered on each gene) and computed, in sliding windows, the ratio of “polymorphism to divergence” (PtoD, the number of SNPs divided by the number of substitutions to the chimpanzee genome). PtoD is thus a measure of diversity that controls for local heterogeneity in mutation rate and grows with older local coalescent times (Hudson et al. 1987;
Recent Selection Changes in Human Genes · doi:10.1093/molbev/msw023

McDonald 1998). PtoD was computed in sixth populations from the 1000 Genomes (see Materials and Methods), and compared with values generated via neutral coalescent simulations (fig. 2). As expected, all genes show high SNP density in Africans (fig. 2A), although the excess does not reach significance in ZNF473. Non-African populations also show overall high mean PtoD, with the exception of PKDREJ in East Asians and ZNF473 in TSI, CHB, and JPT, which show significantly low PtoD mean values (fig. 2A).

When PtoD is investigated along each genomic region (fig. 2B–E), in African populations all genes contain peaks of PtoD above the 95% confidence interval (CI) of the neutral expectation, with PKDREJ, SDR39U1, and CLCNKB having peaks above the 99% CI. The highest local PtoD peak always falls within or very close to the gene, except in ZNF473 (fig. 2E). In non-African populations, when peaks of high diversity exist they overlap those in Africans (fig. 2B–E). The observed reduction of diversity in PKDREJ for Asians (fig. 2B) is also in agreement with a previous study (Pickrell et al. 2009).

Alleles at Intermediate Frequency in Africa and at Low or High Frequency Out-of-Africa

We next investigated the distribution of derived allele frequencies in our set of four genes. The density of alleles at intermediate frequency is higher in Africans than in non-Africans (supplementary figs. S5 and S6, Supplementary Material online). To compare the SFS across populations, we used the joint SFS, which shows the frequency of every polymorphic allele in two populations (fig. 3A). In the neutral control regions, allele frequencies correlate well among populations: Pearson’s correlation coefficient $r^2 = 0.96$ in the comparison between two African populations and $r^2 > 0.81$ in the comparisons between African and non-African populations (table 1).

In the four genes, allele frequencies are also similar between the two African populations (Pearson’s correlation $r^2 = 0.94$; table 1) and among all non-African populations (Pearson’s $r^2 > 0.89$; table 1 and supplementary fig. S7B, Supplementary Material online). In fact, the correlation between pairs of non-African populations is 4–10% higher in the four genes than in the controls (a significant difference, all Fisher r-to-z transformation two-tailed $P < 0.001$), and it is also 8% higher for these alleles than for non-genic alleles in the 1000 Genomes data set (supplementary table S4, Supplementary Material online). When we contrast African and non-African populations (fig. 3B), though, the correlation in allele frequencies between any African and any non-African population is significantly lower in the four genes than in the neutral controls (all Fisher r-to-z transformation two-tailed $P < 0.005$). This corresponds to a 6–26% weaker correlation between Africans and non-Africans in the four genes than in the neutral controls (in the 1000 Genomes data set this is also 8% higher for these alleles than for non-genic alleles in the four genes than in the controls (all Fisher r-to-z transformation two-tailed $P < 0.001$), and it is also 8% higher for these alleles than for non-genic alleles in the 1000 Genomes data set (supplementary table S4, Supplementary Material online). When we contrast African and non-African populations (fig. 3B), though, the correlation in allele frequencies between any African and any non-African population is significantly lower in the four genes than in the neutral controls (all Fisher r-to-z transformation two-tailed $P < 0.005$). This corresponds to a 6–26% weaker correlation between Africans and non-Africans in the four genes than in the neutral controls (in the 1000 Genomes data set this is also 8% higher for these alleles than for non-genic alleles in the 1000 Genomes data set (supplementary table S4, Supplementary Material online).

To understand the basis of these population differences, we focus on the alleles at intermediate frequency in Africa, which are most interesting from the standpoint of balancing selection. We define alleles with derived allele frequency 0.20 ≤ DAF ≤ 0.80 in Africa as “intermediate in Africa alleles” (I-alleles). We then ask what proportion of them has very low (DAF ≤ 0.05) or very high (DAF ≥ 0.95) frequency in
non-Africans; we call these “intermediate in Africa different Out-of-Africa alleles” (iAdO-alleles). In control regions, on average only 18.4% of iA-alleles are iAdO-alleles; in the four genes this proportion is 52.8% (table 2). This corresponds to an average 3.3-fold increase in iAdO-alleles in the genes when compared with neutral regions (table 2), a significant enrichment (all exact binomial tests \( P < 0.008 \)). Therefore, the four genes are enriched in alleles at intermediate frequency in Africa but at high or low frequency outside of Africa. Many of these SNPs are among the most differentiated alleles between African and non-African populations in the 1000 Genomes data set, although only SNPs in PKDREJ remain significant after accounting for their intermediate frequency in Africa (fig. 4 and supplementary table S5, fig. S9, and section 3, Supplementary Material online). Importantly, these SNPs also drive the reduced allele frequency correlation among populations (fig. 3 and fig. 5).

### Table 1. Correlation of Allele Frequencies between Populations in Genes and Controls.

<table>
<thead>
<tr>
<th></th>
<th>LWK</th>
<th>YRI</th>
<th>TSI</th>
<th>CHB</th>
<th>GIH</th>
</tr>
</thead>
<tbody>
<tr>
<td>LWK</td>
<td>—</td>
<td>0.96</td>
<td>0.87</td>
<td>0.84</td>
<td>0.86</td>
</tr>
<tr>
<td>YRI</td>
<td>0.94</td>
<td>—</td>
<td>0.85</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>TSI</td>
<td>0.81</td>
<td>0.71</td>
<td>—</td>
<td>0.86</td>
<td>0.95</td>
</tr>
<tr>
<td>CHB</td>
<td>0.73</td>
<td>0.60</td>
<td>0.89</td>
<td>—</td>
<td>0.88</td>
</tr>
<tr>
<td>GIH</td>
<td>0.78</td>
<td>0.68</td>
<td>0.96</td>
<td>0.95</td>
<td>—</td>
</tr>
</tbody>
</table>

**Note:** The values above and below the diagonal show the correlation coefficients (as Pearson’s \( r^2 \)) for SNPs in the control and the four candidate genes, respectively. All values are highly significant \( (P < 1 \times 10^{-6}) \). In each pairwise comparison, we consider only sites that are polymorphic in at least one of the two populations.

### Table 2. Proportion and Number of iA-alleles and iAdO-alleles.

<table>
<thead>
<tr>
<th></th>
<th>Genes</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YRI (155)</td>
<td>LWK (199)</td>
</tr>
<tr>
<td></td>
<td>TSI (0.406 (63)</td>
<td>0.482 (96)</td>
</tr>
<tr>
<td></td>
<td>CHB (0.677 (105)</td>
<td>0.714 (142)</td>
</tr>
<tr>
<td></td>
<td>GIH (0.400 (62)</td>
<td>0.487 (97)</td>
</tr>
<tr>
<td></td>
<td>YRI (0.297 (22)</td>
<td>0.254 (18)</td>
</tr>
<tr>
<td></td>
<td>TSI (0.297 (22)</td>
<td>0.254 (18)</td>
</tr>
<tr>
<td></td>
<td>CHB (0.149 (11)</td>
<td>0.127 (9)</td>
</tr>
<tr>
<td></td>
<td>GIH (0.149 (11)</td>
<td>0.127 (9)</td>
</tr>
</tbody>
</table>

**Note:** DAF is defined as “intermediate” in Africans according to two different criteria. In both cases, the table shows the proportion of sites with frequency defined as low (DAF \( \leq 0.05 \)) or high (DAF \( \geq 0.95 \)) in non-Africans. All values of the “Genes” are significantly higher than those of the Controls (all exact binomial tests \( P < 0.008 \)). The numbers in parenthesis correspond to the number of iA-alleles in Africans and iAdO-alleles in non-Africans.

*The values correspond to the mean of the ratio Genes/Controls across the three populations (TSI, CHB, and GIH). Note that the values are very similar with both criteria used to define intermediate allele frequencies.

### The Model of Selection

In our set of four genes the double signature of balancing selection in Africans (with increased diversity and intermediate-frequency alleles) indicates long-term balancing selection. The incomplete signature of balancing selection outside of Africa, with excess of polymorphism yet absence of intermediate-frequency alleles, suggests a possible change in the frequency of the balanced polymorphism(s) and linked variation in these populations.
Recent Selection Changes in Human Genes · doi:10.1093/molbev/msw023

Genetic drift or a change in selective forces might explain this unexpected observation. In what follows, we aimed to infer the selective history of these genes and the possible causes for the different signatures between African and non-African populations. To do this, we modeled five evolutionary scenarios that include one African, one European, and one East-Asian population (fig. 6A) and performed an approximate Bayesian computation (ABC) analysis (Beaumont et al. 2002).

Because signatures of balancing selection are clear in Africa, we keep the selective history in Africans identical in all models: A balanced polymorphism arose in the ancestor of all human populations and selection acted continuously in Africa by maintaining the balanced polymorphism at approximately 0.5 (see Materials and Methods). For simplicity, balancing selection is simulated with overdominance, which here is also appropriate to simulate other mechanisms of long-term balancing selection that leave patterns of diversity that are compatible with those observed in Africa (e.g., frequency-dependent selection that favors intermediate-frequency alleles or mild fluctuating selection that maintains polymorphisms for long periods of time; see Discussion and supplementary material section 8, Supplementary Material online, on the likelihood of other types of balancing selection).

Because all non-African populations show similarly incomplete signatures of selection and a higher correlation in allele frequencies compared with controls (fig. 2 and table 1), we kept the selective history identical in the two non-African populations (supplementary material section 4.4, Supplementary Material online). In order to model the changes outside of Africa, we considered five scenarios (fig. 6A):

1. Balancing to Balancing (B-B), in which balancing selection continued acting after the out-of-Africa migration in non-Africans.
2. Balancing to Neutrality (B-N), in which balancing selection stopped acting after the out-of-Africa migration, and the gene evolved neutrally in non-Africans.
3. Balancing to Positive (B-P), in which balancing selection stopped acting after the out-of-Africa migration, and the gene evolved under different types of natural selection in non-Africans as follows:

   i. Balancing to Positive on standing variation (B-Psv): One of the two alleles of the balanced polymorphism became directionally (positively) selected in non-Africans right after the out-of-Africa migration. This is a model of positive directional selection acting on a previously balanced allele, which is similar to a soft sweep or selection on standing variation (sv). It also closely models positive selection on an intermediate-frequency allele that is closely linked to the balanced polymorphism.

   ii. Balancing to Positive on de novo mutation (B-Pdn): The balanced polymorphism became neutral, and a de novo (dn) advantageous mutation appeared in the ancestors of non-Africans right after the out-of-Africa migration, and immediately became directionally (positively) selected.

   iii. Balancing to Positive due to change in frequency equilibrium (B-Pcfe): The frequency equilibrium of the balanced polymorphism changed from ~0.50 to 0.07 in non-Africans right after the out-of-Africa migration, and so the allele was subject to selection to change its allele frequency. We chose the value of 7% because it yielded similar summary statistics to the B-Psv model, and we sought to discriminate between these two models. Because the change is recent this model also mimics changes in other types of balancing selection (e.g., recent changes in long-term mild fluctuating selection driving a selected allele to very high frequency). For simplicity, we consider this also a Balancing-to-Positive selection model.

In summary, we have a model B-B where selection did not change outside of Africa, one model B-N where selection stopped acting outside of Africa, and three B-P models where outside of Africa selection favored the increase in allele frequency of a new or existing variant.

We used an ABC framework (Beaumont et al. 2002) to infer the posterior probability of each of these models given the data. ABC is a useful tool that allows probabilistic model testing and parameter estimation when calculating the model’s likelihood function is not feasible (Beaumont et al. 2002; Bertorelle et al. 2010). In short, the ABC procedure is based on running several thousand simulations (160,000 in our analyses) under each model, with the relevant evolutionary parameters being drawn from a prior distribution. The simulated genetic data are then summarized in a number of summary statistics (see Materials and Methods and supplementary material section 4, Supplementary Material online) and simulations producing statistics that resemble best those of real data are selected to estimate the posterior probability of each model (Beaumont 2008).

Our main goal is to distinguish B-B, B-N and B-P models. To account for the presence of three B-P models, we run the ABC model selection with two different approaches. The first ABC model selection approach was carried out in two hierarchical steps. We first compared the three B-P models (B-Pcfe, B-Psv, and B-Pdn) and performed an ABC model...
choice analysis to select the model with the highest posterior probability. Then the chosen B-P model was compared with the other two models (B-B and B-N) via a second ABC model choice analysis (supplemental material section 4.4, Supplementary Material online).

We determined the accuracy of our model choice inferences calculating the true and false positive rates using 1,000 simulations as Pseudo-Observed-Data for each model (supplementary material section 4.4.2, Supplementary Material online, for the full procedure). The results (supplementary tables S7 and S8, Supplementary Material online) indicate that the true positive rate is good for model B-B (81%), moderate for the three B-P models (on average 63%), and weak for B-N (47%). The false positive rate is relatively high (roughly 12%) and very similar for the three models. Therefore, our analysis is somehow biased in favor of the B-B model. When we compare the three B-P models, the true positive rate is low for each of them (all lower than 47%) although the false positive rate is quite low as well (<8%). This is not surprising given that the tested scenarios are similar and difficult to differentiate by a set of summary statistics (supplementary fig. S13B, Supplementary Material online). Nevertheless, as stated above, we focused on the distinction among three main models (B-B, B-N, and B-P).

**Fig. 5.** Two-dimensional SFS for each of the candidate genes. See figure 3 for more details and supplementary figure S8, Supplementary Material online for other pairwise comparisons.
and give little emphasis to the distinction among the three B-P models (fig. 6C).

Figure 6B shows the results of this model selection approach for each gene and population. For PKDREJ, SDR39U1, and ZNF473, the B-P model has the highest support consistently in all populations; the B-B and B-N models have no and minor support, respectively. For CLCNKB, an ambiguous picture emerges, with modest posterior probabilities favoring the B-B and B-N models. This ABC analysis thus provides little support for stable balancing selection or neutrality outside of Africa in PKDREJ, SDR39U1, and ZNF473, suggesting instead a change in selective pressure as the most likely scenario.

In order to identify potential bias in our estimates due to the hierarchical procedure of selection within the three B-P models, we performed a second model selection with a different approach, which consists in comparing simultaneously the five models (B-B, B-N, B-Pcfe, B-Psv, and B-Pdn). Given that the three B-P models (B-Pcfe, B-Psv, and B-Pdn) produced similar results, we assigned a prior probability of 1/3 for the B-B and B-N models, and of 1/9 for each B-P model. A single ABC model choice analysis was then run to obtain the posterior probability of each of these five models. The results of this approach are extremely similar to those of the hierarchical approach (supplementary fig. S14, Supplementary Material online) and also support the B-P model for PKDREJ, SDR39U1, and ZNF473 and show inconclusive results for CLCNKB.

The joint SFS was not explicitly considered in the ABC analysis, but it clearly displays (fig. 5 and supplementary fig. S8, Supplementary Material online) the differences between CLCNKB (with many alleles in the diagonal of the joint SFS of African vs. non-African populations) and PKDREJ, SDR39U1, and ZNF473, all with a virtually empty diagonal in the joint SFS.

Haplotypes and Populations
The relationships among haplotypes in these three genes showed, as expected, higher diversity in Africa than outside
of Africa. Nevertheless, this is not due to a general loss of low-frequency polymorphisms as in other genomic regions: In all three genes we observe the complete or nearly complete loss of one haplotype lineage outside of Africa (fig. 7). This agrees well with a model where a set of haplotypes increased rapidly in frequency in populations outside of Africa, putatively due to linkage to a variant that is advantageous in non-African populations.

The Functional Effect of iAdO-Alleles

No SNP in these four genes has been associated with diseases or phenotypes in genome-wide association studies (Welter et al. 2014). But the iAdO-alleles are prime candidates to be the targets of changing selective pressures, so we investigated their putative functional consequences. As a group, the 102 iAdO-alleles have significantly higher C-scores for deleterious-ness (Kircher et al. 2014) than expected given C-scores in the rest of the genome (supplementary fig. S15B, Supplementary Material online). This is not the case when only non-synonymous iAdO-alleles are considered (supplementary fig. S15A, Supplementary Material online), suggesting that the potential functional effect is due to regulation. In fact, the set of iAdO-alleles show a significant enrichment (P < 0.001) in high scores for regulatory features (supplementary table S9 and fig. S16, Supplementary Material online) as described in RegulomeDB (Boyle et al. 2012). For example, 31% of iAdO-alleles are predicted with high confidence to affect DNA–protein binding and are associated with changes in gene expression (i.e., mapped to an eQTL; Boyle et al. 2012). Less than 3% of SNPs fall in these functional annotated categories when we randomly sample three genes in the genome (supplementary fig. S16, Supplementary Material online), so this is an unusual enrichment in functional alleles that suggests a possible effect in gene regulation of the observed allele frequency differences. For more details on these analyses, see supplementary material section 7, Supplementary Material online.

The Balanced Alleles through the Out-of-Africa Bottleneck

It is theoretically possible that a change in selective pressure is a direct consequence of demography if, for example, the balanced polymorphism is lost during a bottleneck and selection can no longer act. This is more likely in non-African populations, which experienced a severe out-of-Africa bottleneck (Gravel et al. 2011). The possibility is included in our simulations, but we wanted to formally ask how often we expect it to happen.

The fixation of a balanced allele only due to the increased drift produced by a demographic event is an unlikely scenario, at least for the parameters we considered (i.e., over-dominance with selection coefficient ranging from 0.01% to 10%). In none of the simulations under stable balancing selection (model B-B) was the balanced polymorphism lost. In the scenario with neutrality in non-African populations (model B-N), the probability of fixation of the balanced polymorphism after the out-of-Africa bottleneck is also low according to our simulations: 11% in Europeans, 15% in Asians, and 7% in both Europeans and Asians. This should be considered as a conservative upper bound as we did not allow in the simulations migration between Africans and non-Africans (supplementary material section 4.1, Supplementary Material online), eliminating chance of reintroduction in non-Africans of the lost allele.
Discussion

Humans are a young and quite homogeneous species, with substantial genetic and phenotypic similarity among populations (Rosenberg et al. 2002). In fact, although humans inhabit a wide variety of environments, they colonized areas outside of the African continent only in the last 50,000 years (Gravel et al. 2011). These migrating populations adapted to their new habitats biologically and/or culturally (Richerson and Boyd 2008; Coop et al. 2009), and these local adaptations undoubtedly explain some of the phenotypic differences that exist among human groups today. Because the rate of new mutations is low in humans (Scally and Durbin 2012; Fu et al. 2014), mostly due to our low effective population size (Lynch 2010, 2011), it is likely that these novel adaptations are largely mediated by selection on previously existing variation (Pritchard et al. 2010; Messer and Petrov 2013). The classical definition of positive selection from standing variation considers that alleles segregate neutrally (or nearly neutrally) before becoming advantageous upon environmental change (Innan and Kim 2004; Przeworski et al. 2005; Pennings and Hermisson 2006; Messer and Petrov 2013). However, experimental evolution on yeast has shown that these alleles usually have significant fitness effects (often deleterious) before changes in the environment turn them advantageous (Hietpas et al. 2013). It is indeed logical that the environment will rarely determine whether an allele has functional and phenotypic consequences that affect fitness. Most likely, environmental shifts will instead modify the magnitude and perhaps the direction of the fitness effect of a given mutation.

Therefore variants that have been under balancing selection with a significant (and likely complex) effect on fitness are prime candidates to be affected by selection from standing variation. In addition, loci that contain balanced polymorphisms accumulate a high number of additional variants, some of which are not neutral and may later become advantageous. Given the demographic history of humans, alleles under long-term balancing selection in Africa (or functional, linked alleles) could have contributed to recent human local adaptation.

We investigate this possibility by exploring in detail four genes in humans, and their patterns suggest that this might be the case. The four genes show hallmark signatures of long-term balancing selection in Africa that combines an excess of polymorphism and an excess of intermediate-frequency alleles in both African groups. These patterns are expected under balancing selection with frequency equilibrium around 0.5, such as overdominance (with similar fitness of both homozygotes) or frequency-dependent selection (with favored frequency close to 0.5). Moderately fluctuating selection (with a selected allele varying mildly in frequency around 0.5) lacks a frequency equilibrium but it could produce similar patterns. Strongly fluctuating selection and negative frequency-dependent selection would likely leave different genetic signatures (lacking the excess of polymorphism, the excess of intermediate-frequency alleles, or both; supplementary material section 8, Supplementary Material online). We thus focused on the mechanisms that best predict the patterns observed in Africa. Importantly, these three mechanisms would result in modest differences between populations if selection remained unchanged.

Nevertheless, these four genes show extremely different patterns in Europe and/or Asia, with an absence of the hallmark double signature of balancing selection. It is interesting that the levels of diversity are overall high in non-African populations, and the main difference between African and non-African populations is in the distribution of allele frequencies, which in non-Africans lacks the excess of intermediate-frequency alleles observed in Africans. Correlation in allele frequency between Africans and non-Africans is weaker than in neutral regions, and a substantial amount of alleles present at intermediate frequency in Africans segregate at low or high frequency in populations outside of Africa. These signatures, combined with the incomplete signatures of balancing selection outside of Africa, are compatible with a recent change in selective pressure, which would have changed the haplotype landscape and shifted allele frequencies but not wiped out (quite yet) segregating alleles. The strong correlation in allele frequencies among the non-African populations (stronger than in neutral controls) suggests that the similarities between Europeans and Asians are due to their shared demographic and selective histories.

We formally tested this hypothesis by considering different evolutionary scenarios and conclude that a model with changes in selection outside of Africa (where selection favored an existing or new mutation) best explains the data for three genes: PKDREJ, SDR39U1, and ZNF473. For these genes a model of continuous balancing selection (B-B) or change to neutrality (B-N) has little support. For a fourth gene CLCNKB, results are less conclusive. This could be the result of a reduction in the strength of selection without a change in the selective regime, a possibility that we did not consider. We note that the power of the ABC analysis to distinguish among the three main models is moderate because these scenarios do not produce strikingly different signatures on top of the pre-existing signatures of long-standing balancing selection. This is, per se, a challenging exercise. Still, our main limitation is a bias toward the B-B model (continuous, unchanged balancing selection outside of Africa), which shows extremely weak support in the three genes that have a robust result.

A change from balancing to positive selection seems most likely given our observations, although the change in selective regime could in principle be more complex. For example, the locus might experience drastic changes in the frequency equilibrium (as in our B-Pcfe model) or changes in previously mild fluctuating selection such that the selected allele reaches very low or high frequency in Eurasian populations (while keeping similar, intermediate-frequency alleles in the two African groups). In all cases, one allele increases fast in frequency due to changes in selective pressure (supplementary material section 8, Supplementary Material online).
Table 3. Summary of Population Genetics Results for Eurasian Populations and Biological Features of the Four Genes.

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistics</th>
<th>CLCNKB</th>
<th>PKDREJ</th>
<th>SDR39U1</th>
<th>ZNF473</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSI</td>
<td>PtoD</td>
<td>bal</td>
<td>bal</td>
<td>bal</td>
<td>_*</td>
</tr>
<tr>
<td>SFS</td>
<td>bal</td>
<td>bal</td>
<td>pos/neg</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ABC</td>
<td>B-B</td>
<td>B-P</td>
<td>B-P</td>
<td>B-P</td>
<td></td>
</tr>
<tr>
<td>GHI</td>
<td>PtoD</td>
<td>bal</td>
<td>bal</td>
<td>bal</td>
<td>–</td>
</tr>
<tr>
<td>SFS</td>
<td>–</td>
<td>pos/neg</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ABC</td>
<td>B-B</td>
<td>B-P</td>
<td>B-P</td>
<td>B-P</td>
<td></td>
</tr>
<tr>
<td>CHB</td>
<td>PtoD</td>
<td>bal</td>
<td>_*</td>
<td>bal</td>
<td>_*</td>
</tr>
<tr>
<td>SFS</td>
<td>–</td>
<td>–</td>
<td>pos/neg</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ABC</td>
<td>B-N</td>
<td>B-P</td>
<td>B-P</td>
<td>B-P</td>
<td></td>
</tr>
</tbody>
</table>

Note.—SFS, PtoD: summary of the evidence provided by the neutrality tests (MWU for SFS and HKA for PtoD from fig. 1). "bal" is for balancing selection, "pos/neg" for positive or negative selection (in the specific case of SFS, it refers to excess of low-frequency variants but not high-frequency derived). "_" stands for neutrality, and "_" indicates evidence of positive selection in the 1000 Genomes data (fig. 2) but not for the neutrality tests carried on our data (fig. 1). ABC: Type of model supported in the ABC model choice (see main text and fig. 6 for more details).

The Candidate Genes and Their SNPs

In this section, we summarize the results for each gene (table 3) and provide further information about the genes and their SNPs.

CLCNKB shows strong evidence of long-term balancing selection in virtually all analyses, although the two Asian populations (CHB and GHI) do not show an excess of intermediate-frequency alleles (fig. 1). The ABC does not clearly favor one model across populations. Therefore, it is unclear whether selection remained stable or if it weakened outside of Africa, and we have no convincing evidence of a change in selective pressure in particular human groups.

The remaining three genes show instead clear evidence of a change in selective pressure outside of Africa.

PKDREJ encodes a protein known to play a role in fertilization by generating a Ca\(^{2+}\) transporting channel that is directly involved in initiating the acrosome reaction of the sperm (Butscheid et al. 2006). Its highest expression is in testis (Kissopoulou et al. 2013) and mice knockout spermatozoa are sperm (Butscheid et al. 2006). Its highest expression is in testis than in other tissues (Kissopoulou et al. 2013). Expression of the gene is ubiquitous (Kissopoulou et al. 2013). The gene shows clear signatures of long-term balancing selection in Africans and very different signatures in all Eurasians, where some populations even show classical signatures of positive selection such as an excess of low-frequency alleles. In agreement with these patterns, the ABC results strongly favor a change in selective pressure involving B-P models in all non-African populations (fig. 6B and C).

ZNF473 encodes for a protein involved in histone 3'-end pre-mRNA processing. ZNF473 associates with U7 small nuclear ribonucleoprotein, which mutated in Xenopus blocks histone pre-mRNA processing and disrupt oogenesis (Dominski et al. 2002). Like PKDREJ, ZNF473 is more expressed in testis than in other tissues (Kissopoulou et al. 2013). Despite showing signatures of long-term balancing selection in both African populations, non-African populations lack any signature of balancing selection, including excess of diversity (fig. 1). In addition, the ABC analysis supports the model B-P (fig. 6B and C). Together these results suggest that ZNF473 has experienced drastic changes in selection outside of Africa, probably involving positive selection, although the pre-existing signatures of long-term balancing selection hide classical signatures of a selective sweep. We note that the highest PtoD peaks fall up- and downstream of ZNF473 (fig. 2E) so we cannot discard that regulatory elements or neighboring genes are the targets of natural selection (supplementary material section 6, Supplementary Material online).

Conclusion

In conclusion, our study suggests that balancing selection can create reservoirs of genetic variants that mediate later adaptation. We focused on a number of genes to define this mechanism, but additional cases likely exist in the human genome. Ultimately, these represent events of positive selection on standing variation or soft sweeps, selective events that are notably difficult to identify with classical population genetics methods (Innan and Kim 2004; Przeworski et al. 2005; Pennings and Hermisson 2006) unless selection is recent and very strong (Albrechtsen et al. 2010; Peter et al. 2012; Messer and Petrov 2013; Ferrer-Admetlla et al. 2014). We expect that investigating shifts in selection of previously balanced alleles will help refine the catalog of loci that have contributed to recent adaptation of humans to their local environments.

Materials and Methods

Samples and Populations

We analyzed a total of 150 HapMap samples from five populations (30 individuals per population): YRI from Nigeria, LWK from Kenya, TSI from Italy, GHI from India, and CHB from Beijing. The DNA was purchased from Coriell Cell Repositories. In addition, for analyses where we need empirical genome-wide distributions or longer genomic regions, we also analyzed six populations from the 1000 Genomes phase 1 data (1000 Genomes Project Consortium et al. 2012): Two Africans (YRI and LWK), two Europeans (CEU and TSI), and two East Asians (CHB and JPT).
Targeted Regions
We investigated 14 genes (Supplementary table S1, Supplementary Material online). Four genes were reported by Andrés et al. (2009) as having signatures of long-term balancing selection in African Americans (significant departures from neutral expectations in two neutrality tests) and clearly lacking these signatures in European Americans, with \( P > 0.2 \) in at least one neutrality test. Because that analysis was performed in a potentially admixed group, we included ten additional genes where the signatures of balancing selection did not reach significance in African Americans and were absent in European Americans (Andrés et al. 2009).

We used Sanger and Illumina sequencing technologies to sequence the coding region and adjacent non-coding region of all target genes (Supplementary table S1, Supplementary Material online). We also used 49 control regions described previously as a proxy for neutrality (Andrés et al. 2010). These regions are unlinked, ancient processed pseudogenes, which are distant from genes, do not overlap functional elements, and have GC content similar to coding genes and thus serve as adequate proxy for neutrality.

Together with Illumina and Sanger sequences, we analyzed a total of 230,452 bp (Supplementary table S1, Supplementary Material online), and after stringent quality filters (Supplementary material section 1.3, Supplementary Material online) we retrieved a total of 1,708 and 1,109 high-quality biallelic SNPs for Illumina and Sanger technologies, respectively. Supplementary table S3, Supplementary Material online, reports for each gene and population the number of segregating sites and fixed differences relative to the chimpanzee genome (PanTro3).

Population Genetics Analyses
We compared the patterns of the region of interest (each gene) with neutral regions (the 49 controls) with two neutrality tests, and assessed significance with neutral simulations. We thus determined how unusual the patterns of our genes are by comparing them both with neutral regions of the genome and with expectations under neutrality (Andrés et al. 2009).

The first neutrality test is a modified version of the MWU test that detects departures of the SFS in each gene when compared with the neutral regions (Nielsen et al. 2009). In particular, we compared the folded SFS of each gene with the folded SFS of all control regions with two MWU tests, one to detect an excess of low-frequency alleles and one to detect an excess of intermediate-frequency alleles. The second neutrality test is the HKA test to identify excess of polymorphic over divergent sites in a region of interest (Hudson et al. 1987). Specifically, we compare the ratio of polymorphisms over neutrality tests; the only exception is that we did not allow migration after the out-of-Africa migration (for a full description of the models, see Results and fig. 6A).

ABC Analyses
An ABC framework (Beaumont et al. 2002) was used to infer the most likely evolutionary model. We used 160,000 simulations for each of the five evolutionary scenarios to model changes in selective pressure after the out-of-Africa migration. In all models an overdominant balanced polymorphism arises Tbs generations ago, and is maintained until present-day in African populations. This balanced polymorphism has a selection coefficient (Sbs, drawn from a uniform prior distribution) and a dominance coefficient (h) fixed to 25.5 in order to achieve a frequency equilibrium of 0.51. This frequency equilibrium cannot be exactly 0.50 because the fitness model in "SLiM" (Messer 2013) (one of the two programs used to produce the simulations, see below and supplementary material section 4.1, Supplementary Material online) does not allow the two homozygous to have exactly the same fitness (Gillespie 1978). We refer the reader to the Discussion section and supplementary material section 8, Supplementary Material online, for a discussion on other mechanisms of balancing selection.

The five models differ in the selective regime of the non-African populations, where selection changes right after the out-of-Africa migration (for a full description of the models, see Results and fig. 6A). The following parameters were drawn from uniform prior distributions: Mutation rate, \( \mu = U(1 \times 10^{-8} , 4 \times 10^{-8} \) per site per generation); recombination rate, \( \rho = U(0, 4 \times 10^{-8} \) per site per generation); time since balancing selection, \( Tbs = U(40,000, 240,000 \) generations); selection coefficient of the balanced polymorphism, \( Sbs = U(0.0001, 0.1) \); selection of the de novo advantageous mutation in models B-Pdn, Spd = log \( U(0.0001, 0.01) \). Other parameters were identical to those in the neutral simulations used in the neutrality tests; the only exception is that we did not allow migration (see later in the paragraph). The divergence time between human and chimpanzee was set to 6.5 My, that is, 260,000 generations considering a generation time of 25 years.

Given the complexity of the models and the limitations of current simulation software, we combined strategies of coalescent and forward simulations. Specifically, we used the coalescent simulator msns (Ewing and Hermisson 2010) to generate the genetic data until the time of the out-of-Africa migration. We then used the forward simulator SLiM (Messer 2013),
which can model more complex scenarios, to simulate the evolution of all populations after the out-of-Africa event (supplementary fig. S10, Supplementary Material online). However, due to limitations in SLiM we did not include migration between populations because it can produce the coexistence of different types of natural selection in a population (see supplementary material section 4.3, Supplementary Material online, for more details). Supplementary figure S10, Supplementary Material online, illustrates the simulated models and their demographic parameters.

We considered the 27 summary statistics described in supplementary table S6, Supplementary Material online; we calculated them using “msstats” package from “Libsequence” (Thornton 2003) and in-house scripts in R-language (R Core Team 2013) and selected a subset of 16 informative summary statistics that show only moderate correlation (Pearson’s $r^2 < 0.8$) and together give the greatest power of discrimination among the models (supplementary material section 4.2, Supplementary Material online). We corrected $F_{ST}$ (Weir and Cockerham 1984) to take into account the absence of migration in our simulations (supplementary material section 4.3, Supplementary Material online). The model selection analysis (supplementary material section 4.4, Supplementary Material online) was performed independently for Europeans and Asians using the logistic regression approach (Beaumont 2008) and retaining 50,000 simulations out of 480,000.

Supplementary Material
Supplementary tables S1–S9, figures S1–S16, and sections 1–8 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments
We are thankful to Warren W. Kretzschmar for help with implementation of his code to run the neutrality test; we also thank Carlos Bustamante, Andy Clark, Rasmus Nielsen, and Scott Williamson for insights during the design of the project, and Bárbara Bitarello, Michael Dannemann, Romain Laurent, Fabrizio Mafessoni, Diogo Meyer, Kay Prüfer, Fernando Racimo, Gabriel Renaud, Joshua Schmidt, Mark Stoneking, and João Teixeira for helpful discussions and statistical advices. We are thankful to the MPI-EVA’s sequencing group and Martin Kircher for illumina data production, and to Warren W. Kretzschmar and Valerie Maduro for visually checking some Sanger SNPs. We are also thankful to two anonymous reviewers who helped to improve the manuscript. This work was supported in part by the Intramural Research Program of the National Human Genome Research Institute of the U.S. National Institutes of Health.

References


