Standing genetic variation as the predominant source for adaptation of a songbird

Yu-Ting Lai[a], Carol K. L. Yeung[b], Kevin E. Omland[c], Er-Li Pang[d], Yu Hao[d], Ben-Yang Liao[e], Hui-Fen Cao[f], Bo-Wen Zhang[d], Chia-Fen Yeh[a], Chih-Ming Hung[g], Hsin-Yi Hung[a], Ming-Yu Yang[h], Wei Liang[i], Yu-Cheng Hsu[j], Cheng-Te Yao[k], Lu Dong[d], Kui Lin[d], and Shou-Hsien Li[a,1]

[a]School of Life Science, National Taiwan Normal University, 11677 Taipei, Taiwan; [b]Novogene Bioinformatics Institute, 100083 Beijing, People's Republic of China; [c]Department of Biological Sciences, University of Maryland, Baltimore County, MD 21250; [d]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beijing Normal University, 100875 Beijing, People's Republic of China ; [e]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, 35053 Miaoli County, Taiwan; [f]Institute of Genomics, School of Biomedical Sciences, Huaqiao University, 361021 Xiamen, People's Republic of China ; [g]Biodiversity Research Center, Academia Sinica, 11529 Taipei, Taiwan; [h]School of Life Sciences, Peking University, 100080 Beijing, People's Republic of China ; [i]Ministry of Education Key Laboratory for Ecology of Tropical Islands, College of Life Sciences, Hainan Normal University, 571158 Haikou, People's Republic of China ; [j]Department of Natural Resources and Environmental Studies, National Dong Hwa University, 97401 Hualien, Taiwan; and [k]Division of Zoology, Endemic Species Research Institute, 55244 Nantou, Taiwan

Edited by Scott V. Edwards, Harvard University, Cambridge, MA, and approved December 17, 2018 (received for review August 7, 2018)

What kind of genetic variation contributes the most to adaptation is a fundamental question in evolutionary biology. By resequencing genomes of 80 individuals, we inferred the origin of genomic variants associated with a complex adaptive syndrome involving multiple quantitative traits, namely, adaptation between high and low altitudes, in the vinous-throated parrotbill (*Sinosuthora webbiana*) in Taiwan. By comparing these variants with those in the Asian mainland population, we revealed standing variation in 24 noncoding genomic regions to be the predominant genetic source of adaptation. Parrotbills at both high and low altitudes exhibited signatures of recent selection, suggesting that not only the front but also the trailing edges of postglacial expanding populations could be subjected to environmental stresses. This study verifies and quantifies the importance of standing variation in adaptation in a cohort of genes, illustrating that the evolutionary potential of a population depends significantly on its preexisting genetic diversity. These findings provide important context for understanding adaptation and conservation of species in the Anthropocene.

standing variation | population genomics | adaptation | postglacial expansion

Adaptation lies at the heart of Darwinian evolution. It requires the presence of advantageous alleles that are either new mutations (1, 2) or preexisting standing variants (3, 4). Population genetic theory conventionally considers novel mutations to be the genetic source of adaptation (1, 5), and their significance has been highlighted in various empirical studies (6, 7). However, a high level of standing variation may allow a faster response to environmental changes than waiting for appropriate mutations to arise. Recent genomic studies also provide evidence for the role of standing variation in adaptation (8, 9). Nevertheless, the relative contribution of new mutations and standing variation to adaption has rarely been evaluated. This is fundamental to the identification of the major driver underpinning this critical evolutionary process and to the forecast of the fate of species in the Anthropocene, as the ability to adapt is central to species' survival in changing environments (10).

Evaluating the relative contribution of standing and new genetic variation to adaptation requires assessing a cohort of genes that are under directional selection. Screening for the genetic basis underlying a complex adaptive syndrome is likely to yield a set of such genes. Here, we screened for genetic variants that underlie adaptation between high and low altitudes that likely involves many quantitative traits (11) of the vinous-throated parrotbill (*Sinosuthora webbiana*), one of the most widely distributed nonmigrant passerines in East Asia (12). In Taiwan, this bird is found from sea level up to 3,100 m (13). Individual parrotbills only disperse within a limited distance (14). Field data suggests that it exhibits no seasonal altitudinal migration and may only range within 100 m in altitude (15).

To evaluate the relative contribution of different sources of adaptive genetic variants, we inferred the relative age of genetic variants across the entire genome. As a continental island, Taiwan was periodically connected with the Asian mainland when land bridges surfaced during the Pleistocene glacial periods (16), allowing Taiwan and mainland populations to resume gene flow. Consequently, genetic variants shared by the two populations are likely to represent standing variation inherited from their most recent common ancestors. In contrast, new mutations which have arisen since the latest episode of gene flow would only be found in either one of the populations.

Since the last glacial period, temperatures in Taiwan have risen by about 9.6 °C (17), likely driving species to shift or expand their ranges from lowland glacial refugia toward higher altitudes (18, 19). While populations that expanded to higher altitudes faced challenges such as lower temperature and oxygen partial pressure, those that remained in the lowlands would also have found themselves in a warmer environment, an aspect that has been little investigated. Knowledge of how species adapted to

Significance

It is a tenet of modern biology that species adapt through natural selection to cope with the ever-changing environment. By comparing genetic variants between the island and mainland populations of a passerine, we inferred the related age of genetic variants across its entire genome and suggest that preexisting standing variants played the predominant role in local adaptation. Our findings not only resolve a long-standing fundamental problem in biology regarding the genetic sources of adaptation, but imply that the evolutionary potential of a population is highly associated with its preexisting genetic variation.

Author contributions: K.L. and S.-H.L. designed research; Y.-T.L., W.L., Y.-C.H., C.-T.Y., and L.D. performed research; W.L., Y.-C.H., C.-T.Y., and L.D. contributed new reagents/analytic tools; Y.-T.L., E.-L.P., Y.H., B.-Y.L., H.-F.C., B.-W.Z., C.-F.Y., C.-M.H., and M.-Y.Y. analyzed data; and Y.-T.L., C.K.L.Y., K.E.O., B.-Y.L., B.-W.Z., C.-M.H., H.-Y.H., M.-Y.Y., and S.-H.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequence reported in this paper has been deposited in the National Center for Biotechnology Information Sequence Read Archive (accession nos. SRX5087906 to SRX5087990; BioProject ID PRJNA504683).

[1]To whom correspondence should be addressed. Email: t43028@ntnu.edu.tw.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1813597116/-/DCSupplemental.

Published online January 18, 2019.

global warming in the past may help us understand the challenges they face in a continuously warming planet.

De Novo Assembly of a Draft Genome for the Parrotbill. We first sequenced and assembled de novo a reference genome of a male Taiwanese vinous-throated parrotbill. The assembled genome, 1.06 Gb in length (genomic coverage = 110×, *SI Appendix*, Table S1), contains 6,508 scaffolds (2,034 scaffolds > 10 kb) with N50 size of 1.94 Mb. This draft genome includes a total of 72 Mb (6.8%) of repetitive sequences (*SI Appendix*, Table S2). We found 20,149 predicted genes, 9,581 of which could be functionally annotated. In total, we obtained 278 Mb of genes (26% of total genome) consisting of 253 Mb of introns and 25 Mb of coding regions.

Genetic Variation and Geographic Structure of Parrotbills in Taiwan. We resequenced the genomes of 40 parrotbill individuals that were collected from four populations (10 individuals per population; Fig. 1*A*) with an average depth of 5.8× genome coverage (*SI Appendix*, Table S3). They formed two high- (>2,000 m) and low- (<100 m) altitudinal comparison pairs, one from the east side of the Central Mountain Range (CMR) in Taiwan and one from the west side. The altitudinal gradients between the two high- and low-altitude populations are steep (43 km and 55 km apart for the east and west population pairs, respectively). The two low-altitude localities were much warmer (over 10 °C higher annual mean temperature) and drier (50% lower annual precipitation) than the two high-altitude localities (*SI Appendix*, Table S3).

Comparing with the reference genome, 20,017,403 polymorphic sites (S) within the 40 resequenced genomes were found. The levels of intrapopulation genetic diversity are similar (π and mean individual heterozygosity, $\bar{H}$, *SI Appendix*, Tables S3 and S5). However, the number of polymorphic sites (S), a demographically sensitive index, is about 12% greater in the low-altitude (mean S = 11,329,479) than in the high-altitude populations (mean S = 10,132,251). This is consistent with the conjecture that populations at or near glacial refugia retain more stable and larger effective population sizes (N_e) (hence more genetic variants) than populations at the expanding edges during postglacial expansion (20, 21).

The population structure was revealed in an unrooted neighbor-joining population tree (*SI Appendix*, Fig. S3) based on genome-wide F_{ST} (fixation index, a genetic measurement of population differential) values among all four populations (*SI Appendix*, Table S6), which indicates that pairs of populations on the same side of the CMR are clustered together. Thus, the two population pairs can serve as independent samples for illustrating parallel adaptation along altitudinal gradients (22).

Demographic History of the Vinous-Throated Parrotbill in Taiwan. Bottleneck and other events could leave genetic signatures resembling those of selection (23). We used the diploid sequence of the reference genome and a high-coverage genome sequence (24.5×) of an additional male Taiwanese parrotbill to test for recent severe bottleneck events. Results of pairwise sequentially Markovian coalescent (PSMC) analysis (24) indicate that their

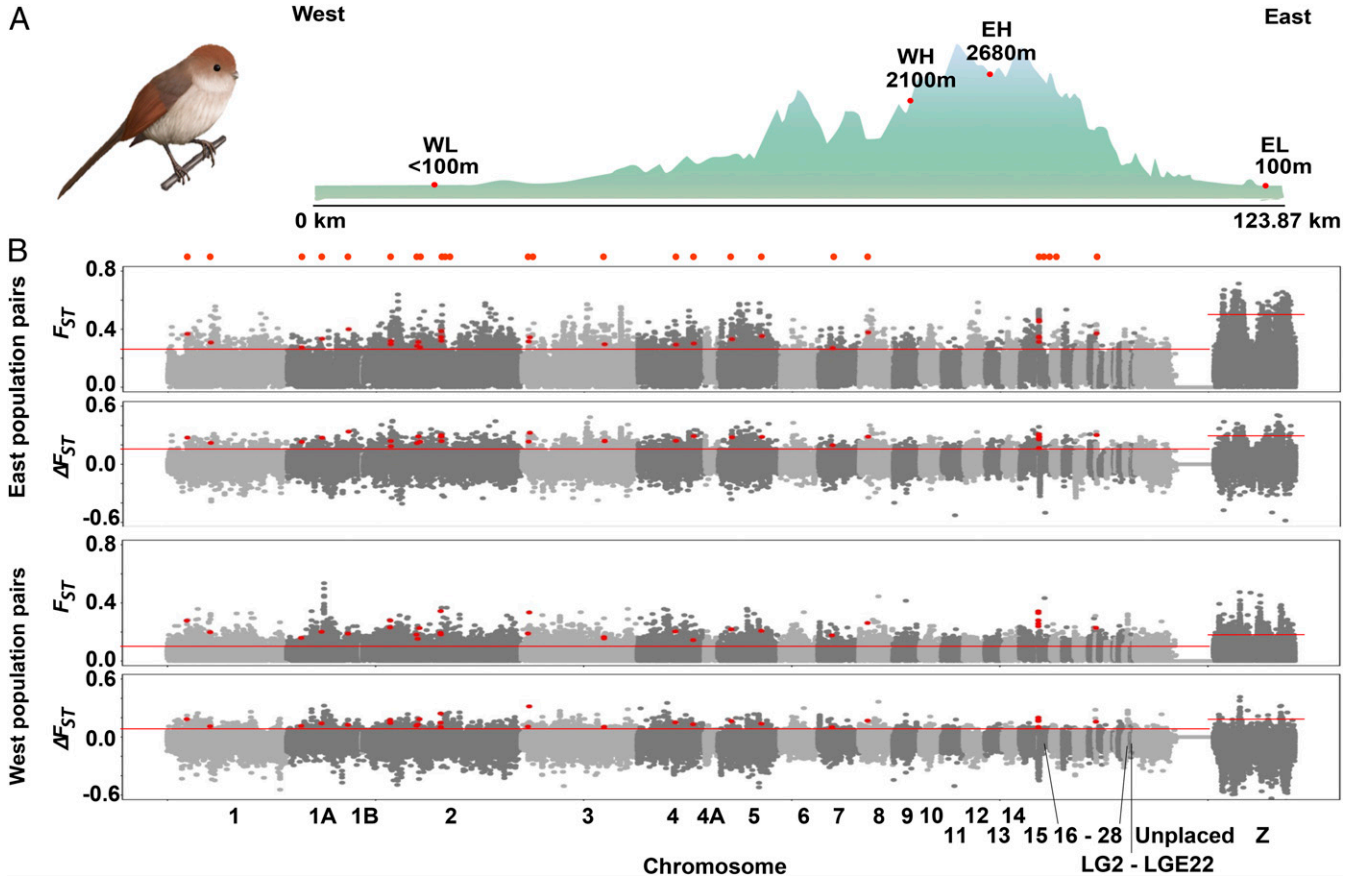

Fig. 1. (*A*) The vinous-throated parrotbill and four sites (red dots) on an east–west section of central Taiwan at which vinous-throated parrotbills were sampled and (*B*) the distribution of the F_{ST} and ΔF_{ST} in the east and west high-/low-altitude local population pairs of each 10-kb nonoverlapping genomic window that was aligned with the published genome of the zebra finch. Red dots on top of and within each panel represent candidate regions on the genome (n = 24); red horizontal lines indicate the top 1% of F_{ST} and ΔF_{ST}. EH, high-altitude population east of CMR; EL, low-altitude population east of CMR; WH, high-altitude population west of CMR; WL, low-altitude population west of CMR.

N_e varied through time, although without evidence of a severe population bottleneck (*SI Appendix*, Fig. S4). N_e was ∼30,000–100,000 from the middle of the last glacial period (∼50–70 thousand years ago) to the present.

Genomic Regions Associated with Divergent Selection Between Different Altitudes. For the east population pair, we found 2,812 outlier regions ($0.715 \geq F_{ST} \geq 0.271$; the top 1% of F_{ST} values; mean genome-wide $F_{ST} = 0.108$, 95% CI = 0.108–0.109); for the west population pair, we found 3,051 outlier pairs ($0.536 \geq F_{ST} \geq 0.140$; mean genome-wide $F_{ST} = 0.052$, 95% CI = 0.051–0.052) (Fig. 1*B*). Of these regions, 152 are shared between the east and west populations. It is important to consider that both ecologically divergent selection and purifying selection would reduce genetic diversity on functionally conserved regions and lead to high F_{ST} in the linked genomic region. Whereas divergent selection should only affect populations in different environments, by contrast, purifying selection would affect all populations (25). For the Taiwanese parrotbills, F_{ST} of 10-kb windows was highly correlated between all population pairs (*SI Appendix*, Fig. S5), implying that the high-F_{ST} regions across genome is largely the result of purifying selection. Purifying selection on linked genomic regions is further supported by the negative correlation between F_{ST} and π values (*SI Appendix*, Fig. S6), positive correlation between F_{ST} values and the level of linkage disequilibrium, r^2 (*SI Appendix*, Fig. S7), and negative correlation between π and r^2 values (*SI Appendix*, Fig. S8). Therefore, we employed ΔF_{ST} analysis to reduce the effect of purifying selection on genetic differentiation of genomic regions.

Only 24 10-kb nonoverlapping autosomal regions in the top 1% for both F_{ST} and ΔF_{ST} values ($F_{ST} \geq 0.464$ and 0.344 in the east and west population pairs, respectively; and $\Delta F_{ST} \geq 0.164$ and 0.101 in the east and west population pairs, respectively) were considered candidate regions for adaptation to altitudes along both sides of the CMR (Fig. 1*B*). These regions have high levels of genetic differentiation between different altitudes and are less likely to be affected by purifying selection. Only six of the candidate regions are concatenated into two longer divergent genomic islands that were 20 and 40 kb in length.

Seventeen candidate genes were identified (*SI Appendix*, Table S8). Among them, four are involved in oxygen cascade: *VAV3* and *COL15A1* are related to angiogenesis; *IGF2* to respiratory system phenotype; and *TPPP* to hemoglobin content. Three candidate genes are related to aspects of metabolism that could be involved in thermoregulation: *SUPT7L* (26) and *HBP1* are related to lipid metabolism; and *OLA1* to ATP hydrolysis. High-altitude mammals have physiological adaptations in their circulatory, hematologic, respiratory, and thermoregulatory systems as well as in their metabolism in comparison with their sea-level counterparts (27, 28). Therefore, the genes we identified may be associated with adaptation between high and low altitudes. Among them, *VAV3* (29), *TPPP* (30), *IGF2* (31), and *COL1A1* (32) have been reported to be associated with altitudinal adaption or hypoxia syndrome in other organisms. We also found that *HBP*1 (33) and *OLA*1 (34) are associated with temperature acclimation and heat shock phenotypes, respectively, in other vertebrates. The phenogenotypic associations of these genes in other systems not only validate our inferences, but also support the view that multiple quantitative traits, including oxygen cascade and thermal regulation, were involved in the adaptation to different altitudes.

Although adaptation could have been achieved through regulatory evolution of individual genes (e.g., those mentioned above), functionally related genes could also have changed expression simultaneously in the adaptive process. To examine this possibility, we performed an enrichment analysis on gene ontology (GO) terms. Among the 23 GO terms in which the 17 candidate genes were enriched [false discovering rate (FDR) < 0.05; $P < 0.0019$], many were found to be associated with glycogen metabolism (7 GO terms) and epigenetic control of gene expression (11 GO terms), especially histone acetylation (8 GO terms) for the latter group (*SI Appendix*, Table S9). This observation is consistent with previous studies showing reduced activity of glycogen metabolism accompanied by increased metabolism of fatty acids in organisms moving to a higher altitude (35, 36). This result also implies that epigenetic control, especially histone modification involving acetylation, played a critical role in the adaptation of vinous-throated parrotbills to different altitudes and their ability to cope with rapid environmental changes (37). Consistently with that possibility, the involvement of chromatin modifications in high-altitude adaptation has been observed in several mammals (38, 39).

Predominant Role of Noncoding Genomic Regions in Adaptation to Different Altitudes. Within the 24 candidate regions, 35 (average $F_{ST} = 0.767$; range: 0.631–1.000) and 32 (average $F_{ST} = 0.628$; range: 0.461–0.832) SNPs with the highest F_{ST} values in the east and west altitudinal pairs, respectively, were considered as candidate SNPs (*SI Appendix*, Tables S10–S13). The F_{ST} values of these candidate SNPs are significantly higher than those for all SNPs in both altitudinal pairs (Welch two-sample t test: $t = 19.795$ in the east pair and 21.592 in the west pair, $P < 1 \times 10^{-15}$). Furthermore, we found that most of the candidate SNPs are fixed or nearly fixed (frequency of major allele ≥ 0.9) in one of the four populations (34 and 27 of the candidate SNPs for the eastern and western populations, respectively; *SI Appendix*, Tables S10–S13), suggesting that they may have had major effects on the associated phenotypes. The higher polymorphism levels of the other candidate SNPs suggest that they are probably under soft sweep or polygenic selection and thus unable to drive these advantageous genetic variants to fixation (40). Noticeably, among these candidate SNPs, only 10 (29.4%) and 4 (14.8%) of them were fixed or nearly fixed in both high and low altitudes in the east and west populations, respectively. This implies that the intensity of divergent selection could be asymmetric at different altitudes for these candidate SNPs.

Intriguingly, none of these candidate SNPs is located within coding regions: eight are located in intronic regions of the *HBP1* and *AAK1* genes, and the other 59 are located in intergenic regions. The disproportionate number of candidate SNPs located in intergenic regions (Fisher's exact test: $P = 0.008$) supports the predominant role of regulatory regions in adaptation (41, 42). Alleles in regulatory regions are often codominant in their effect, have mild pleiotropic consequences, and modify only the expression of genes rather than their functions (43).

Standing Genetic Variation as the Predominant Source of Adaptation. The prerequisite for inferring standing variation using the comparative approach is recent coancestry between sampled taxa. We found that the mainland and Taiwan parrotbill populations diverged early in the last glacial period with estimated N_es of 482,065 for Taiwan and 1,527,174 for the mainland (*SI Appendix*, Table S14). Their recent divergence is also supported by the low genome-wide net genetic distance (D_a) between them, 0.0007 ($D_{xy} = 0.0051$ between Taiwan and mainland populations; πs of Taiwan and mainland populations are 0.0040 and 0.0048, respectively). This implies a divergence time ($t = D_a/2\mu t$; t, divergence time; μ, mutations per site per year) of about 152,000 y, meaning that genetic variants shared between the extant Taiwan and mainland parrotbill populations are unlikely to have arisen independently. However, the F_{ST} values between the mainland population and the low-altitude populations in Taiwan are lower (mean = 0.122) than those between the mainland and the high-altitude ones (mean = 0.133). This supports the view that the high-altitude populations are the leading edge of a postglacial population that possesses lower genetic diversity and a smaller N_e than the low-altitude refuge population.

Assuming genetic panmixia across the vinous-throated parrotbill's mainland range (44), we further resequenced the

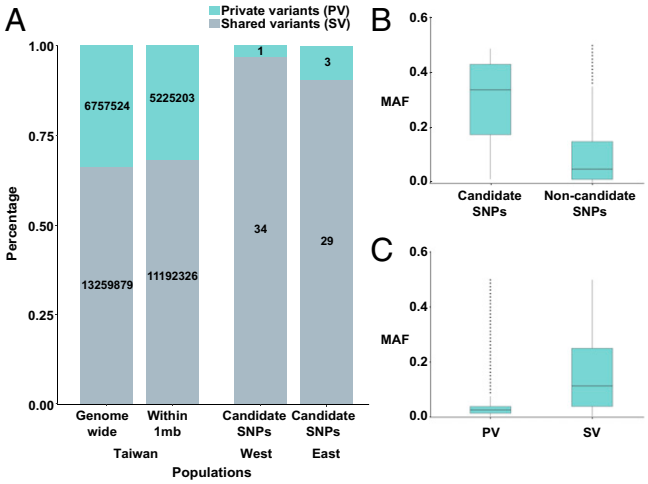

Fig. 2. (*A*) The proportion of polymorphic SNPs shared with the mainland population is significantly higher for candidate SNPs inferred in both high- and low-altitude population pairs than that of the entire genome and SNPs within 1 Mb downstream and upstream regions of all genes in the genome (within 1 mb) of the Taiwan population (East, east population pair; West, west population pair; Fisher's exact test, candidate SNPs: $P = 1.6 \times 10^{-5}$ and $P = 0.002$ for the east and west high-/low-altitude population pairs, respectively; all of the SNPs within 1 Mb regions: $P < 0.00001$ for both east and west high-/low-altitude population pairs). (*B*) The minimum allele frequency (MAF) of candidate SNPs in the mainland population is significantly higher than that of noncandidate SNPs (Welch two-sample *t* test: $t = -10.524$, df = 63, $P = 1.644 \times 10^{-15}$). (*C*) The MAF of shared variants is significantly higher than private variants (Welch two-sample *t* test, $t = -2{,}909.8$, df = 19,750,000, $P < 2.2e\text{-}16$). PV, private variants; SV, shared variants.

genomes of 40 parrotbill individuals (*SI Appendix*, Table S4) to represent the species' genetic diversity in the mainland. We found ~1.3 times as many polymorphic sites in the mainland (S = 27,879,100) as in the Taiwan population (S = 20,017,403). About 66.2% of polymorphic sites in Taiwan (S = 13,259,879 sites) were also polymorphic in the mainland. These shared polymorphic sites likely represent common standing SNPs (minimum allele frequency, MAF, of SNPs ≥ 1.25%) in both the mainland and Taiwan populations, while the private genetic variants of each population are more likely to be rare standing SNPs (MAF < 1.25%) that our sampling scheme failed to detect in the other population, together with some new mutations that occurred since the populations split. About 94% of the candidate SNPs in the Taiwan populations are also polymorphic (or shared) in the mainland populations (Fig. 2*A*), significantly more than the 66.2% of all Taiwan SNPs (Fisher's exact test: $P = 1.6 \times 10^{-5}$ and $P = 0.002$ for the east and west altitudinal pairs, respectively; Fig. 2*A*). The genomic regions around functional elements including genes should experience different evolutionary trajectories from those around nonfunctional regions because of the distinct densities of selection targets within them (45). Therefore, we compared the proportion of shared SNPs in a range 1 Mb upstream and downstream of all genes in the genome with the proportion of the candidate SNPs themselves that were shared. The proportion of candidate SNPs that was shared was still significantly higher (Fisher's exact test: $P < 0.001$ and $P = 0.004$; Fig. 2*A*), confirming that shared standing genetic variation was the predominant genetic source of adaptation.

In the mainland parrotbill population, the MAF of candidate SNPs (average MAF = 0.300, $\sigma^2 = 0.021$) was significantly higher than that of other SNPs (average MAF = 0.107, $\sigma^2 = 0.016$; Welch two-sample *t* test: $t = -10.524$, df = 63, $P = 1.644 \times 10^{-15}$; Fig. 2*B*). Assuming that the mainland population with its large long-term N_e (*SI Appendix*, Table S14) is more likely to retain the ancestral allelic frequencies, our results support the view that standing variants with high initial frequencies facilitate a swift response to selection posed by changing environments (46, 47). Because the N_e of parrotbills is relatively large in Taiwan, genetic drift should only have a limited effect on the dynamics of allelic frequency in the population. Because the average MAF of private SNPs in Taiwan (average MAF = 0.038, $\sigma^2 = 0.002$) is significantly lower than that of shared ones (average MAF = 0.158, $\sigma^2 = 0.019$; Welch two-sample *t* test, $t = -2909.8$, df = 19,750,000, P value < 2.2e-16; Fig. 2*C*), these private SNPs are likely to have preexisted in the ancestral population in low frequency. The low proportion of private SNPs in candidate SNPs implies that rare genetic variants in the ancestral population are less likely to enable response to rapid environmental changes as predicted by the theory (4).

The significant role of shared standing genetic variation in adaptation revealed in this study (Fig. 2) has important implications for our understanding of adaptation and biological conservation. Our results suggest that to cope with environmental changes quickly, a population may benefit significantly from the short waiting time needed to exploit standing genetic variation, especially for variants with high initial frequencies. Our results are consistent with the accumulating evidence that species with high levels of genetic diversity can persist in changing environments (48) and that they are more likely to colonize novel environments (49). Therefore, a species' evolutionary potential can be constrained by a lack of genetic variation (50). This underscores the need to preserve the genetic diversity of species in the face of accelerating environmental changes. Conversely, species with low genetic variation could fail to cope with accelerating anthropogenic environmental changes, which warrants more attention to their conservation status.

Both High and Low Altitudes Could Have Been Harsh for the Parrotbill Since the End of the Last Glacial Period. During the course of postglacial range expansion, parrotbills at the leading expansion edge are thought to have encountered different environmental challenges from those at the trailing edge, and high altitudes are conventionally considered to have harsher environments than low altitudes (11). We would thus expect high-altitude parrotbills to be under more stringent recent selection. To test the "harsh high altitudes" hypothesis, we examined the polymorphic patterns

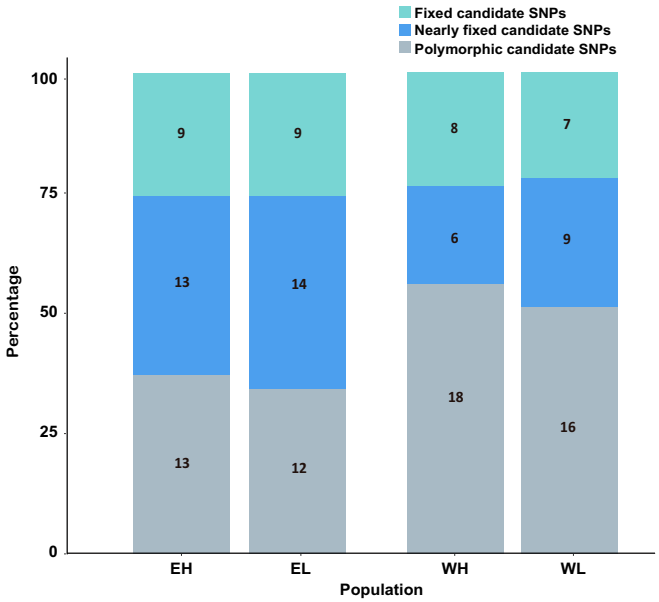

Fig. 3. The proportion of candidate SNPs that are fixed or nearly fixed (90% ≤ the major allele frequency < 100%) in the low- and high-altitude populations.

of the candidate SNPs at different altitudes and found similar numbers of fixed or nearly fixed candidate SNPs on both sides of the island (Fisher's exact test $P = 0.809$ and 0.803 for the east and west pairs, Fig. 3). Assuming that fixation probability depends on the strength of directional selection in a given environment, this suggests that the strength of selection pressures at different altitudes was comparable. Evidence of strong selection pressure in the low-altitude populations also arose from the results of the Tajima's D test that we used to detect recent selective sweep [<0.1 N_e generations (51, 52)], in 10-kb sliding regions across the genome. Recent selective sweeps that occurred since the last glacial period as indicated by the lowest 1% of Tajima's D value in each local population were detected for 6 (Tajima's D ranges from: −1.3569 to −1.6825) out of 24 candidate regions, with similar frequencies in population pairs on both sides of the CMR (none and four for the east high- and low-altitude populations, respectively; one and two for the west high- and low-altitude populations, respectively; Fisher's exact test: $P = 0.112$ and $P = 1.000$ for the east and west population pairs, respectively).

Our results suggest that during postglacial expansion, the population remaining at low altitudes (the trailing edge of the expansion) could also have experienced strong selection pressure. These low-altitude selection pressures may have been caused by increased temperatures, whereas those at the high-altitude leading edge of the expansion arose from the lower oxygen partial pressure. This finding calls for more research on adaptation at low altitudes or latitudes where populations at the trailing edge of postglacial expansion are located and are considered as long-term reservoirs of species' genetic diversity and cradles of speciation (53).

Materials and Methods

Reference Genome and Genome Annotation. To assemble a reference genome, we constructed two paired-end and mate-paired libraries for Illumina short-read sequencing from the DNA of a male Taiwan vinous-throated parrotbill (*Sinosuthora webbiana bulomachus*). The draft genome was assembled with *ALLPATHS-LG* (54) version-44099.

Repeated DNA sequences were masked with *RepeatMasker* (55) version open-4.0.5. We followed the procedure in Ellegren et al. (56) to remove potential contaminating DNA in the draft genome. Finally, 6,508 scaffolds were left for use in the draft parrotbill genome assembly. Then we used *Satsuma* (57) to align the parrotbill draft genome assembly to the well-annotated and assembled zebra finch (*Taeniopygia guttata*) genome. Furthermore, we used *Augustus* (58) for gene prediction and the ALDB database (59) for identifying lncRNAs.

Population Resequencing and Variants Calling. Forty parrotbills (34 males and 6 females) were collected from two high-altitude and two low-altitude locations in Taiwan (*SI Appendix*, Table S3). Six females were sampled from the high-altitude local population in Hualien County; due to the low population density at high altitudes (60), other samples were all males. Climatic information (annual temperature and annual mean precipitation) for each locality was extracted from global climate data (Worldclim v1.4, ref. 61).

To infer whether SNP sites in Taiwan's population are shared or private variants, we also sampled male parrotbills from four mainland populations with 10 individuals from each local population (*SI Appendix*, Table S4).

A whole-genome resequencing library was constructed for each individual. The average coverage of the parrotbill population from Taiwan and the mainland was 5.8× and 12.1×, respectively (*SI Appendix*, Tables S3 and S4).

We used the algorithm *BWA-MEM* to map the raw reads on the reference genome. Then we used *Samtools* (62) for variant calling and *Vcf-tools* (63) to generate a consensus sequence for each individual, and then we randomly assigned one allele from heterozygous genotypes to two putative haplotypes.

Divergence Demography of the Vinous-Throated Parrotbill. The demographic history of the vinous-throated parrotbill in Taiwan was estimated by a hidden Markov model (HMM) implanted in the pairwise sequentially Markovian coalescent (*PSMC*) method. Effective population sizes were inferred from autosomes of two individuals with average genome coverage of 24.5×.

We used *G-PhoCS* v1.3 (64) to estimate long-term N_e and the divergent time (τ) between the Taiwan and mainland parrotbill populations, based on genome sequences of four individuals (two from mainland and two from Taiwan) with high coverages (>20×).

Genome Scan.

Calculating summary statistics of genetic variation. The *PopGenome package* (65) in R with its sliding window method was used to calculate the following statistics in the Taiwan populations: interpopulation differentiation, F_{ST} and D_{xy}, intrapopulation diversity, π, and *Tajima's D* with 10-kb nonoverlapping sliding windows. To evaluate linkage statistics, R^2, within each 10-kb window, we used *beagle* version 4.1 (66) to phase all of the 40 individuals from Taiwan.

Inferring candidate regions. An unrooted neighboring-join tree based on the pairwise comparison of F_{ST} values between the four populations in Taiwan was constructed using *MEGA7* (*SI Appendix*, Fig. S3) (67). Based on this result, four local populations were assigned to two high-/low-altitudinal pairs from the east and west side of the Central Mountain Range (CMR) in Taiwan.

To identify outlier regions that are likely associated with divergent selection between different altitudes we calculated the z-transformed F_{ST} value (zF_{ST}) of the two high-/low-altitudinal local population pairs of each 10-kb window. The 10-kb windows with zF_{ST} over 2.33 (the top 1%) were arbitrarily defined as outlier regions. Because there is a difference of about 580 m in altitude between the two high-altitude populations, which might result in some altitudinal effects, ΔF_{ST} was obtained by subtracting $F_{ST(low/low)}$ from $F_{ST(high/low)}$, and then z-transforming ($z\Delta F_{ST}$). Regions with higher ΔF_{ST} values are more likely to be under divergent than linked selection. Only regions with the top 1% values of both F_{ST} and ΔF_{ST} are referred to as candidate regions.

Defining candidate SNPs and inferring whether they are standing variation or private variants. Genetic variants with the highest F_{ST} value within each candidate region were referred to as candidate SNPs. If these candidate SNPs were also found to be polymorphic in the mainland population, they were regarded as standing variants; otherwise, they were considered to be private variants.

Genes related to adaptation between high and low altitudes. Expression levels of a gene can be modulated by enhancers that are up to 1 Mb upstream or downstream of its start codon (68). We therefore defined genes within 1 Mb of the candidate region as candidate genes under divergent selection between different altitudes.

Gene ontology enrichment analysis. Protein sequences and GO annotations of zebra finch genes annotated by *Ensembl* were retrieved from BioMart. Orthologs of the vinous-throated parrotbill and zebra finch were identified using the *InParanoid* (v4.1, ref. 69) algorithm using default parameters). As a result, 8,635 vinous-throated parrotbill genes with at least one ortholog in the zebra finch genome were identified. The *Ensembl*-annotated GO terms of the zebra finch genes and the corresponding upstream GO terms retrieved using the R package "*GO.db*" (70) were assigned to the vinous-throated parrotbill genes. Fisher's exact test (R function: *fisher.test*; null hypothesis no enrichment) was used to test the statistical significance of each GO term; the false discovering rate (FDR) was determined by R function: *p.adjust*.

ACKNOWLEDGMENTS. We thank J. Q. Li, C. T. Wei, and L. W. Wang for providing parrotbill samples from China; Y. C. Zhang, T. C. Chu, and K. H. Chao for bioinformatics assistance; A. Watson for improving the manuscript; Y. Liu, F. Dong, W. N. Bai, and J. H. Hu for commenting on the early version of manuscript; and C. Y. Chen for illustration of parrotbill. This project was funded by a grant from Taiwan Ministry of Science and Technology (NSC 99 2321-B-003-004-MY3).

1. Smith JM, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genet Res* 23: 23–35.
2. Barton N (1998) Evolutionary biology. The geometry of adaptation. *Nature* 395:751–752.
3. Lande R (1975) The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet Res* 26:221–235.
4. Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends Ecol Evol* 23:38–44.
5. Barton NH (2000) Genetic hitchhiking. *Philos Trans R Soc Lond B Biol Sci* 355:1553–1562.
6. Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE (2009) On the origin and spread of an adaptive allele in deer mice. *Science* 325:1095–1098.
7. Van't Hof AE, et al. (2016) The industrial melanism mutation in British peppered moths is a transposable element. *Nature* 534:102–105.
8. Jones FC, et al.; Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.
9. Reid NM, et al. (2016) The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* 354:1305–1308.
10. Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proc Biol Sci* 275:649–659.
11. Körner C (2007) The use of 'altitude' in ecological research. *Trends Ecol Evol* 22:569–574.

12. Robson C (2007) Family Paradoxornithidae (Parrotbills), *Handbook of the Birds of the World*, vol. 12, eds del Hoyo J, et al. (Lynx Edicions, Barcelona, Spain), pp 292–320.
13. Hachisuka M, Udagawa T (1951) Contribution to the ornithology of formosa: Part 2. *Quart J Taiwan Mus* 4:1–180.
14. Lee J-W, Lee Y-K, Hatchwell BJ (2010) Natal dispersal and philopatry in a group-living but noncooperative passerine bird, the vinous-throated parrotbill. *Anim Behav* 79: 1017–1023.
15. Hsu YC, et al. (2014) Trophic niche width increases with bill-size variation in a generalist passerine: A test of niche variation hypothesis. *J Anim Ecol* 83:450–459.
16. Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations. *J Biogeogr* 27:1153–1167.
17. Lu WC, et al. (2009) Paleoenvironmental study of the Yueh Tan core II in Sun Moon Lake, central Taiwan. *West Pac Earth Sci* 9:119–136.
18. Woodward FI (1987) *Climate and Plant Distribution* (Cambridge Univ Press, New York).
19. Hoffmann AA, Parsons PA (1997) *Extreme Environmental Change and Evolution* (Cambridge Univ Press, New York).
20. Hewitt G (2000) The genetic legacy of the quaternary ice ages. *Nature* 405:907–913.
21. Petit R, et al. (2003) Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* 300:1563–1565.
22. Elmer KR, Meyer A (2011) Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends Ecol Evol* 26:298–306.
23. Thornton KR, Jensen JD, Becquet C, Andolfatto P (2007) Progress and prospects in mapping recent selection in the genome. *Heredity (Edinb)* 98:340–348.
24. Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. *Nature* 475:493–496.
25. Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol* 23:3133–3157.
26. Smith CL, et al. (2018) Mouse Genome Database (MGD)-2018: Knowledgebase for the laboratory mouse. *Nucleic Acids Res* 46:D836–D842.
27. Lenfant C (1973) High altitude adaptation in mammals. *Am Zool* 13:447–456.
28. Moore LG, Niermeyer S, Zamudio S (1998) Human adaptation to high altitude: Regional and life-cycle perspectives. *Am J Phys Anthropol* 107:25–64.
29. Scheinfeldt LB, et al. (2012) Genetic adaptation to high altitude in the Ethiopian highlands. *Genome Biol* 13:R1.
30. Ma Q, et al. (2017) Hypoxia promotes chemotherapy resistance by down-regulating SKA1 gene expression in human osteosarcoma. *Cancer Biol Ther* 18:177–185.
31. Schuler B, et al. (2005) HIF-1 and the adaptation of man to high altitude. *Schweiz Z Med Traumatol* 53:82–87.
32. Falanga V, Zhou L, Yufit T (2002) Low oxygen tension stimulates collagen synthesis and COL1A1 transcription through the action of TGF-β1. *J Cell Physiol* 191:42–50.
33. Scott GR, Johnston IA (2012) Temperature during embryonic development has persistent effects on thermal acclimation capacity in zebrafish. *Proc Natl Acad Sci USA* 109:14247–14252.
34. Mao RF, et al. (2013) OLA1 protects cells in heat shock by stabilizing HSP70. *Cell Death Dis* 4:e491.
35. Grover RF, Weil JV, Reeves JT (1986) Cardiovascular adaptation to exercise at high altitude. *Exerc Sport Sci Rev* 14:269–302.
36. Jackson CGR, Sharkey BJ (1988) Altitude, training and human performance. *Sports Med* 6:279–284.
37. Franks SJ, Hoffmann AA (2012) Genetics of climate change adaptation. *Annu Rev Genet* 46:185–208.
38. Alkorta-Aranburu G, et al. (2012) The genetic architecture of adaptations to high altitude in Ethiopia. *PLoS Genet* 8:e1003110.
39. Dong K, et al. (2014) Genomic scan reveals loci under altitude adaptation in Tibetan and Dahe pigs. *PLoS One* 9:e110520.
40. Pritchard JK, Pickrell JK, Coop G (2010) The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol* 20:R208–R215.
41. King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.
42. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* 134:25–36.
43. Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* 8:206–216.
44. Qu Y, et al. (2012) Incomplete lineage sorting or secondary admixture: Disentangling historical divergence from recent gene flow in the vinous-throated parrotbill (*Paradoxornis webbianus*). *Mol Ecol* 21:6117–6133.
45. Ellegren H, Galtier N (2016) Determinants of genetic diversity. *Nat Rev Genet* 17: 422–433.
46. Lande R, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* 50:434–437.
47. Hermisson J, Pennings PS (2005) Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics* 169:2335–2352.
48. Jump AS, Marchant R, Peñuelas J (2009) Environmental change and the option value of genetic diversity. *Trends Plant Sci* 14:51–58.
49. Prentis PJ, Wilson JR, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends Plant Sci* 13:288–294.
50. Lande R, Barrowclough GF (1987) *Effective Population Size, Genetic Variation, and Their Use in Population Management*, ed Soule ME (Cambridge Univ Press, New York), pp 87–123.
51. Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics* 160:1179–1189.
52. Pennings PS, Hermisson J (2006) Soft sweeps III: The signature of positive selection from recurrent mutation. *PLoS Genet* 2:e186.
53. Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecol Lett* 8:461–467.
54. Gnerre S, et al. (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci USA* 108:1513–1518.
55. Smit AFA, Hubley R, Green P (1996) RepeatMasker Open-4.0. Available at www.repeatmasker.org. Accessed June 11, 2015.
56. Ellegren H, et al. (2012) The genomic landscape of species divergence in Ficedula flycatchers. *Nature* 491:756–760.
57. Grabherr MG, et al. (2010) Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics* 26:1145–1151.
58. Stanke M, Waack S (2003) Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19(suppl_2):ii215–ii225.
59. Li A, et al. (2015) ALDB: A domestic-animal long noncoding RNA database. *PLoS One* 10:e0124003.
60. Hsu YC, Shaner PJ, Chang CI, Ke L, Kao SJ (2014) Trophic niche width increases with bill-size variation in a generalist passerine: A test of niche variation hypothesis. *J Anim Ecol* 83:450–459.
61. Hijmans RJ, et al. (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978.
62. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993.
63. Danecek P, et al.; 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.
64. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nat Genet* 43: 1031–1034.
65. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ (2014) PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol* 31: 1929–1936.
66. Browning BL, Browning SR (2016) Genotype imputation with millions of reference samples. *Am J Hum Genet* 98:116–126.
67. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874.
68. Pennacchio LA, Bickmore W, Dean A, Nobrega MA, Bejerano G (2013) Enhancers: Five essential questions. *Nat Rev Genet* 14:288–295.
69. Remm M, Storm CEV, Sonnhammer ELL (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol* 314:1041–1052.
70. Carlson M (2018) GO.db: A Set of Annotation Maps Describing the Entire Gene Ontology. R Package Version 3.7.0. Available at bioconductor.org/packages/release/data/annotation/html/GO.db.html. Accessed September 25, 2018.