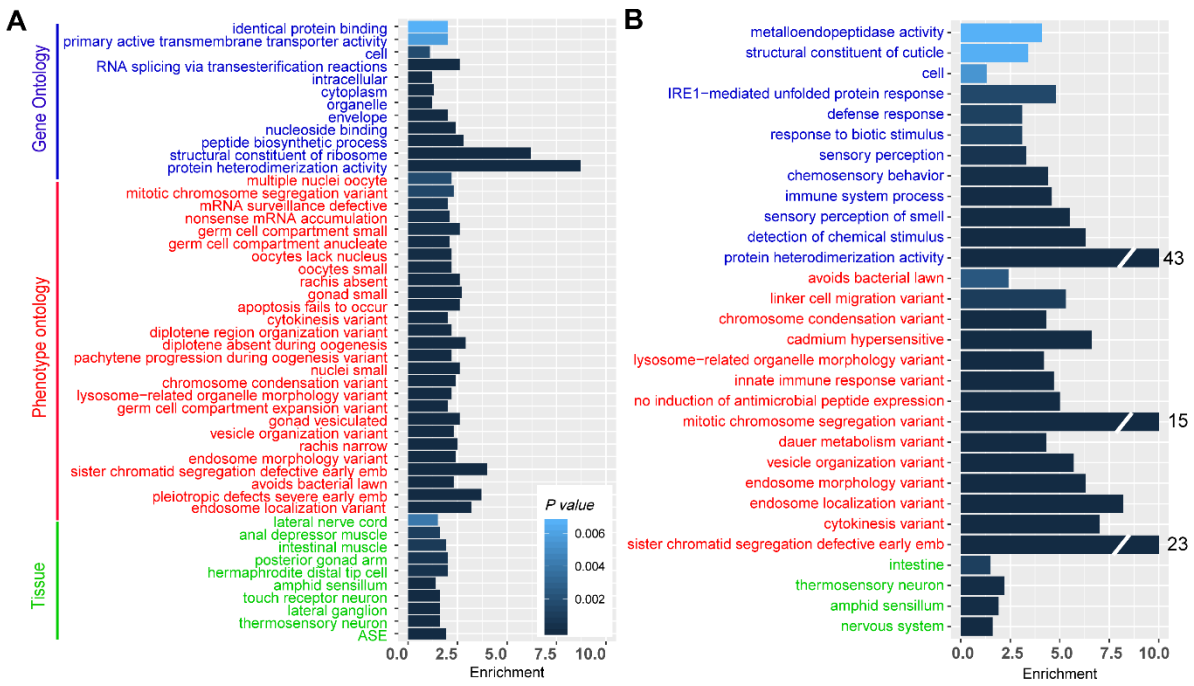
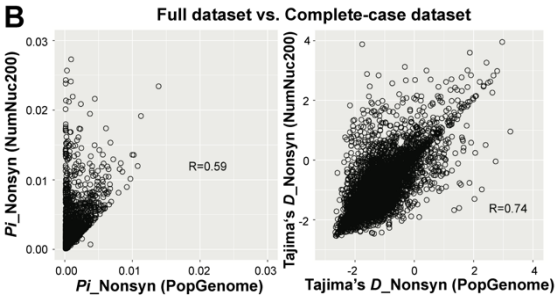


Supplementary Figure S1. Gene ontology enrichment analysis of genes with no nonsynonymous SNVs. (A) Gene enrichment analysis for the 1143 genes with no nonsynonymous SNVs and small indels in the CDS region. (B) Gene enrichment analysis for the 302 genes that have no nonsynonymous SNVs and small indels in the whole gene. Results of gene ontology enrichment (blue), phenotype enrichment (red), and tissue expression enrichment (green) were obtained using the enrichment analysis tools on Wormbase (WS275).

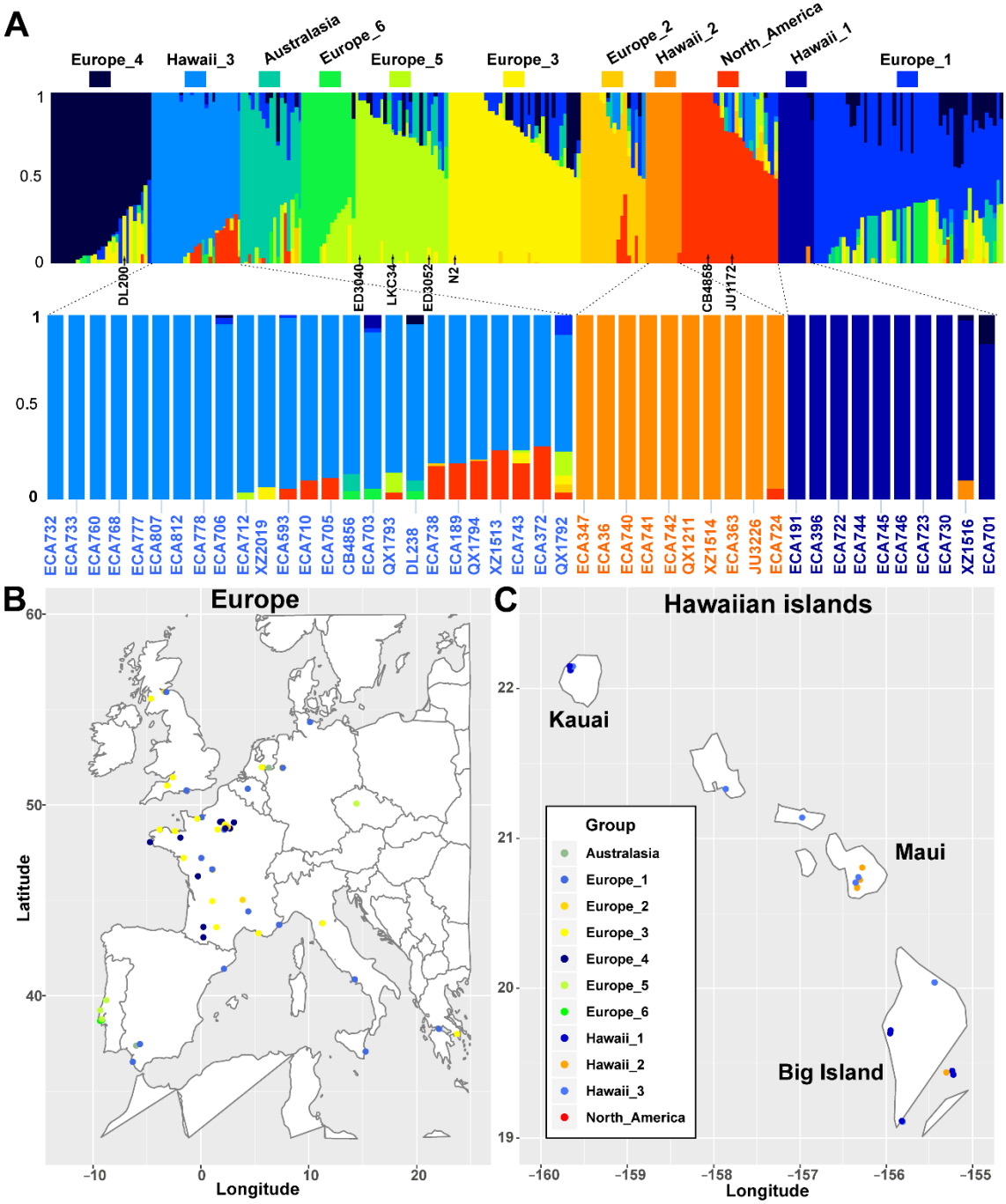
A

	Complete-case dataset		Full dataset	
	DnaSP	PopGenome	VariScan_200	VariScan_450
		Pi		
Syn	0.00053	0.00051	0.00110	0.00110
Nonsyn	0.00040	0.00040	0.00092	0.00086
Intron	0.00067	0.00070	0.00200	0.00170
UTR	0.00110	0.00110	0.00230	0.00210
		Tajima's *D*		
Syn	-0.93	-1.09	-0.92	-0.93
Nonsyn	-1.08	-1.26	-1.08	-1.10
Intron	-1.24	-1.51	-1.19	-1.23
UTR	-0.86	-1.04	-0.82	-0.86

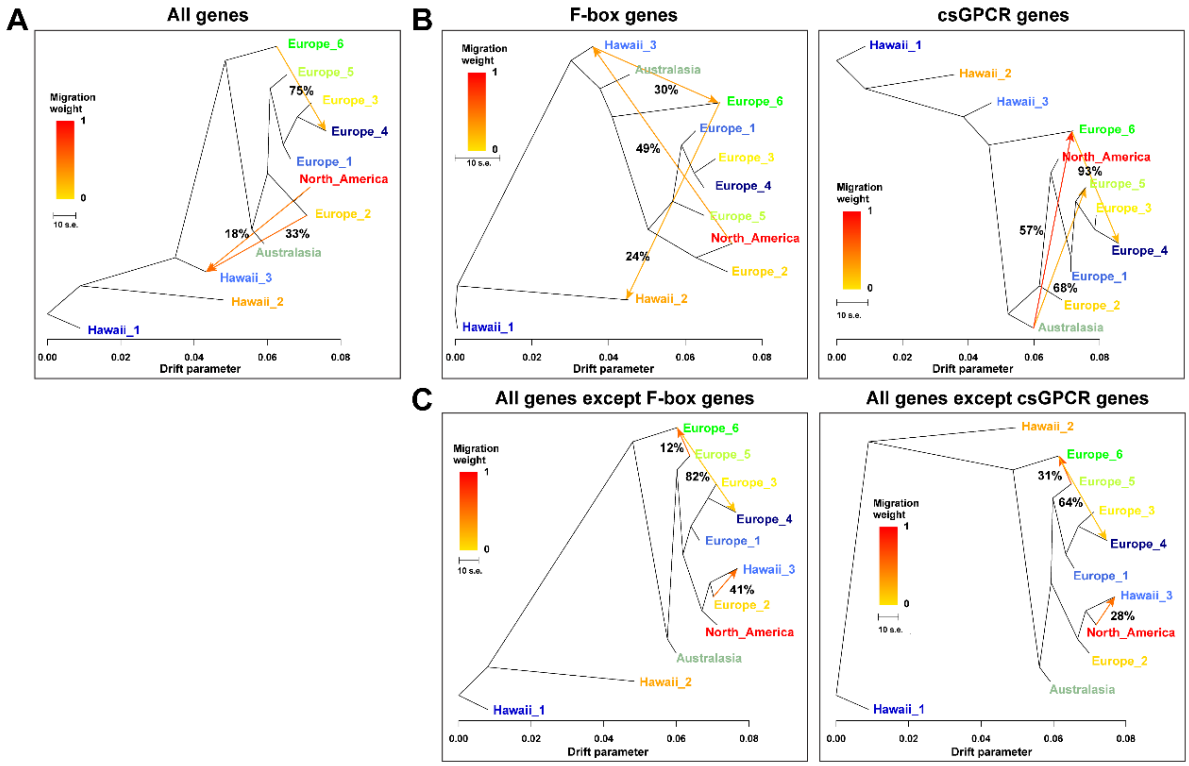
B



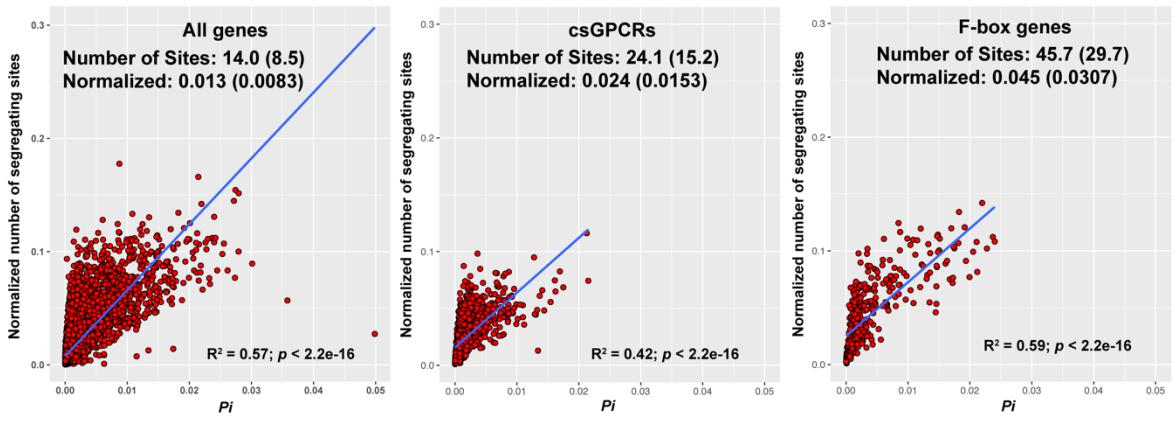
Supplementary Figure S2. Computation of nucleotide polymorphism (*Pi*) and Tajima's *D* for synonymous, nonsynonymous, intron, and UTR SNVs in genes using different software. (A) The mean of *Pi* and Tajima's *D* calculated for synonymous, nonsynonymous, intron and UTR SNVs in all genes. (B) The correlation of *Pi* and Tajima's *D* computed using the full dataset (VariScan; NumNuc = 200) and the complete-case dataset (PopGenome; excluding sites with any missing genotype data).



Supplementary Figure S3. Population structure and geographical distribution of *C. elegans* wild isolates. (A) The ancestral population proportions of wild isolates (color-coded) inferred by Admixture, when the number of population (K value) was set at 11. The bar indicates the ancestral fraction for each strain. Only strains with one ancestral proportion larger than 0.5 were shown (see Table S2 for the data on all strains). Each group (or subpopulation) was named after the geographical location of most strains in that group. The names of the strains from the three Hawaiian subpopulations were shown in the enlarged panels. (B-C) Geographical locations of strains isolated in Europe (B) and Hawaii (C). Each dot represents a strain, but some dots are entirely overlapping given the scale of the map. The latitude and longitude data for each strain was obtained from CeNDR. Color code is consistent with (A).



Supplementary Figure S4. Gene flow of nonsynonymous variants of F-box and csGPCR genes among different subpopulations. Gene flow among the three Hawaiian and eight non-Hawaiian subpopulations inferred by TreeMix using the nonsynonymous SNVs of all genes (A), the F-box genes, csGPCRs (B) and all genes excluding the F-box or csGPCR genes (C). Three migration events were allowed and 1000 bootstrap replicates were run with "Hawaii_1" set as the outgroup. The three most probable migration events from the 1000 bootstraps were shown and the percentage indicated the probability for the migration event. The color of the arrow indicates migration weight.

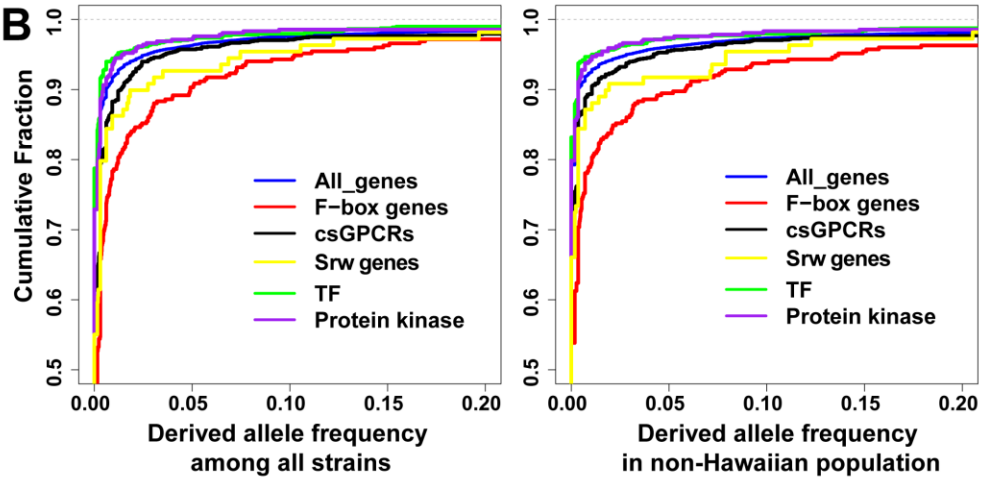


Supplementary Figure S5. Correlation between polymorphism and the number of segregating sites. Significantly positive Pearson correlation between CDS length-normalized number of segregating sites and *Pi*, for all genes, csGPCRs, and F-box genes. The number in parentheses indicate the corresponding values excluding singletons.

A

	# of genes with CNVs	Total # of genes	Enrichment fold	# of CNVs	# of CNVs per gene
All genes	5586	20220	N.A.	8740	0.43
F-box genes	185	336	1.99	669	1.99
csGPCRs	552	1301	1.54	939	0.72
Srw genes	49	112	1.58	86	0.77
TF	198	891	0.80	454	0.51
Protein Kinase	115	402	1.04	239	0.59

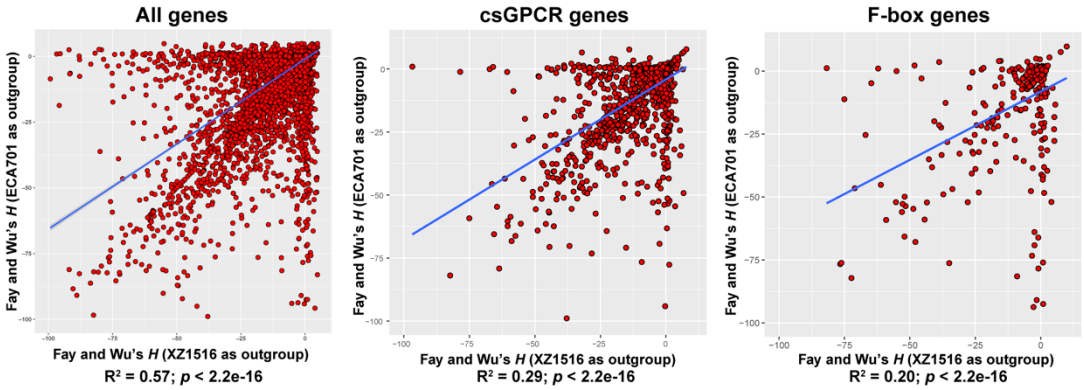
B



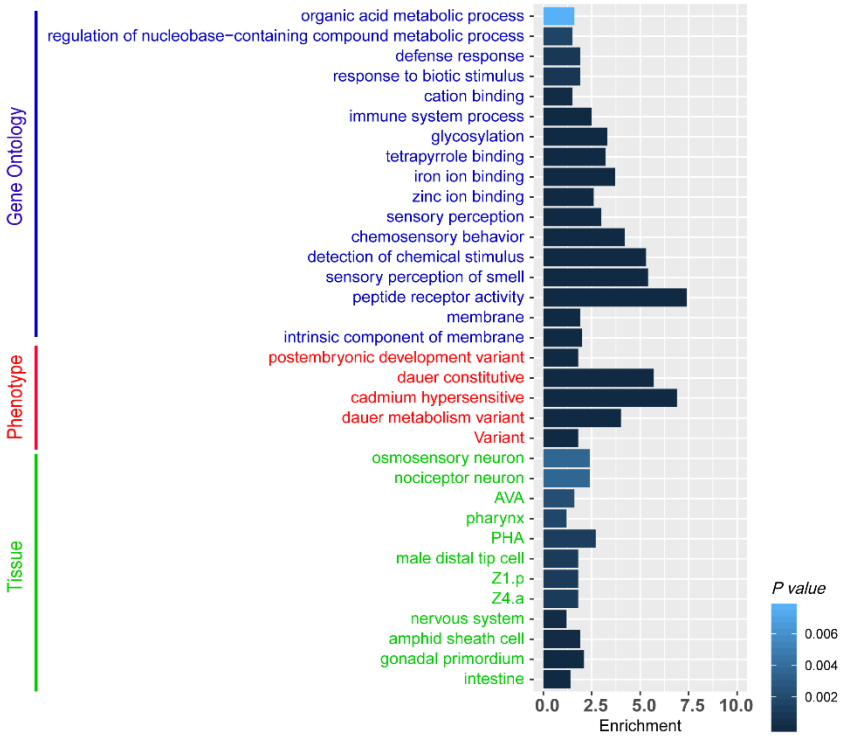
C

	Derived allele frequency	
	In all strains	In non-Hawaiian population
All genes	0.0130	0.0154
F-box genes	0.0243	0.0326
csGPCRs	0.0166	0.0186
Srw genes	0.0180	0.0180
TF	0.0096	0.0110
Protein Kinase	0.0120	0.0123

Supplementary Figure S6. Polymorphism and derived allele frequency for copy number variants (CNVs) in the coding region of different families of genes. (A) The number of genes with CNVs in coding region of all genes, F-box genes, csGPCRs, *Srw* genes, TF, and Protein kinase genes, as well as the number of CNVs occurring in these gene families. (B) The cumulative distribution of derived CNVs in different gene families in all strains or the non-Hawaiian population using XZ1516 as the outgroup. (C) The mean value of derived CNV allele frequency for different gene families.

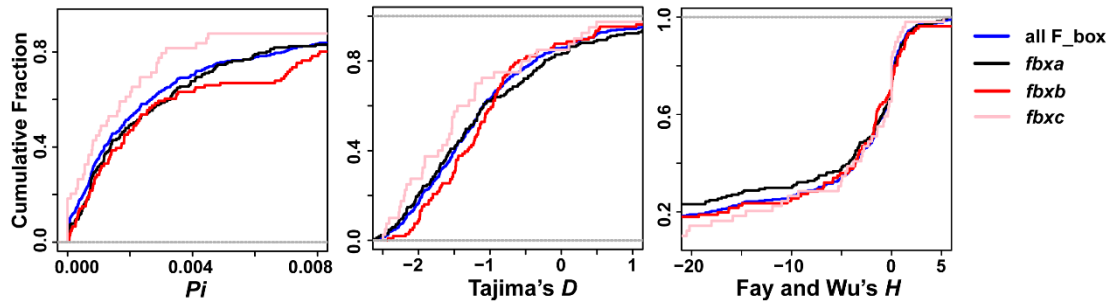


Supplementary Figure S7. Correlation between Fay and Wu's *H* values with XZ1516 and ECA701 as outgroups. Significantly positive Pearson correlation between the *H* values calculated using XZ1516 and ECA701 as outgroups for all genes, csGPCRs, and F-box genes.



Supplementary Figure S8. Gene enrichment analysis for genes whose Fay and Wu's *H* lower than -20. Results of gene ontology enrichment (blue), phenotype enrichment (red), and tissue expression enrichment (green) were obtained using the enrichment analysis tools on Wormbase (WS275).

A



B

Gene	Tajima's *D*	*Pi*	Fay and Wu's *H*
skr-1	-0.84	0.00002	0.01
skr-2	-0.51	0.00055	-3.18
skr-3	-1.13	0.00100	-4.40
skr-4	-1.36	0.00600	-7.84
skr-5	1.14	0.02300	-17.06
skr-6	0.20	0.00410	0.99
skr-7	-2.03	0.00041	-1.75
skr-8	-1.80	0.00020	-1.87
skr-9	-1.14	0.00100	0.25
skr-10	-2.00	0.00014	-9.86

Gene	Tajima's *D*	*Pi*	Fay and Wu's *H*
skr-12	-1.82	0.00030	0.15
skr-13	-1.02	0.00058	0.09
skr-14	-1.40	0.00110	-1.50
skr-15	-2.05	0.00130	-37.58
skr-16	-1.04	0.00008	0.04
skr-17	0.41	0.00140	-1.44
skr-18	-0.55	0.00110	-2.72
skr-19	-0.79	0.00004	0.02
skr-20	-1.11	0.00005	-1.96
skr-21	-0.40	0.00084	-0.61

Gene	Tajima's *D*	*Pi*	Fay and Wu's *H*
cul-1	-1.16	0.00029	-3.49
cul-2	-1.33	0.00018	0.41
cul-3	-1.46	0.00002	-3.89
cul-4	-1.34	0.00007	-1.83
cul-5	-1.69	0.00003	0.07
cul-6	-0.83	0.00046	-6.78

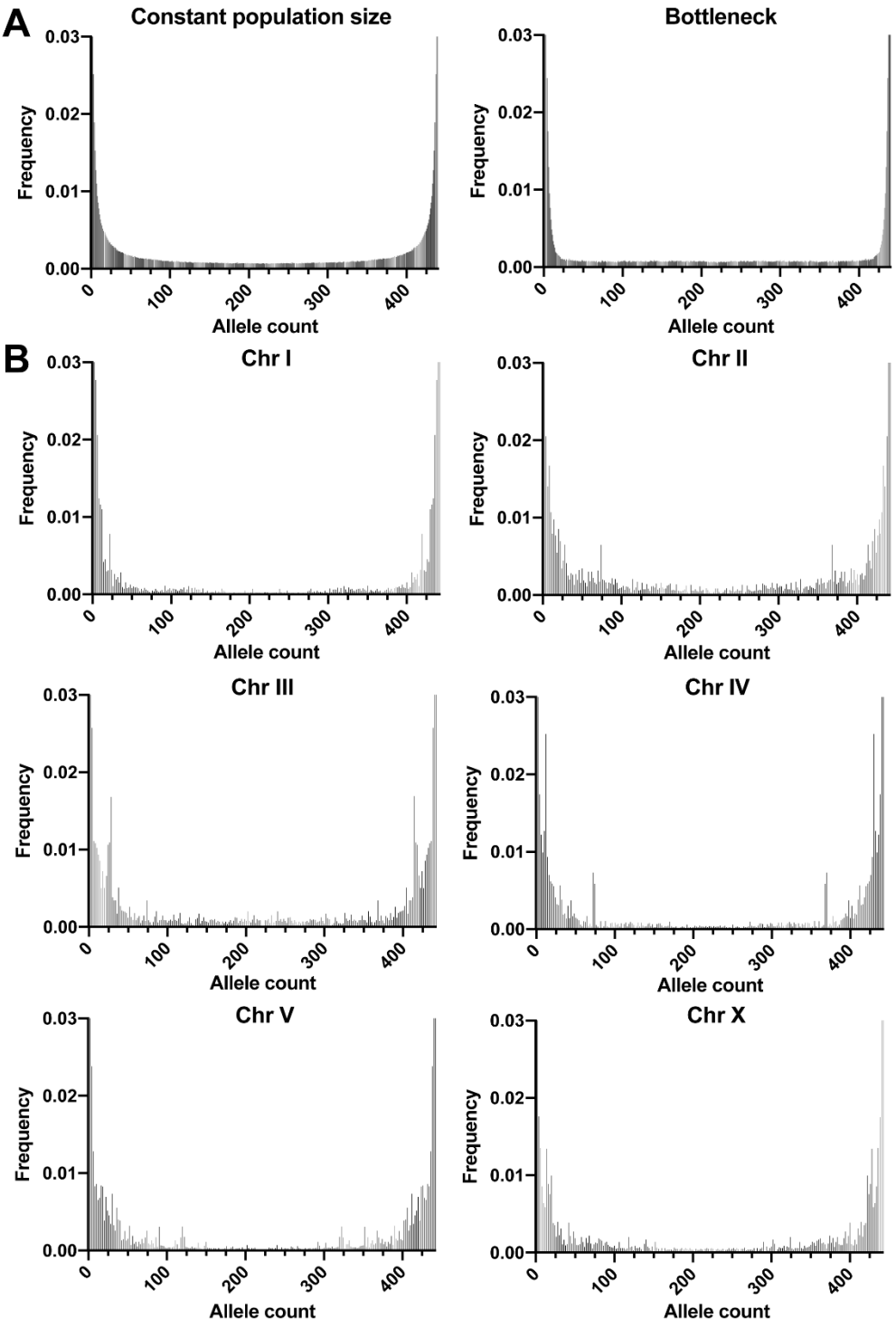
Supplementary Figure S9. Population genetics analysis for SCF complex factors and genes in TF subgroups. (A) The cumulative distribution of *Pi*, Tajima's *D*, and Fay and Wu's *H* values for *fbxa*, *fbxb*, and *fbxc* type F-box genes. (B) The *Pi*, Tajima's *D*, and Fay and Wu's *H* values of Skp1-homologus genes and Cullin genes. *skr-11* is a pseudogene and not shown. (C) The cumulative distribution of *Pi*, Tajima's *D*, and Fay and Wu's *H* value of genes in TF subgroups. See Table S1 for the classification of TF genes into subgroups. (D) The mean and median value of *Pi*, Tajima's *D*, and Fay and Wu's *H* values for all TFs and the five TF subgroups. Double asterisks indicate $p < 0.05$ in Wilcoxon's rank-sum test.

	Pi		Tajima's *D*		Fay and Wu's *H*	
	Syn	Nonsyn	Syn	Nonsyn	Syn	Nonsyn
All genes	0.00099	0.00091	-0.91	-1.08	-6.68	-4.17
All genes except F-box	0.00096	0.00086	-0.91	-1.08	-6.64	-3.75
All genes except csGPCRs	0.00087	0.00086	-0.9	-1.07	-5.74	-3.47
F-box genes	0.00260	0.00420	-0.97	-1.07	-9.28	-11.87
csGPCRs	0.00260	0.00179	-1.09	-1.2	-20.05	-9.76
Srw genes	0.00360	0.00340	-1.2	-1.27	-33.55	-22.46
TF	0.00110	0.00078	-0.92	-1.08	-7.88	-4.27
Protein kinase	0.00051	0.00035	-1.01	-1.21	-5.89	-2.81

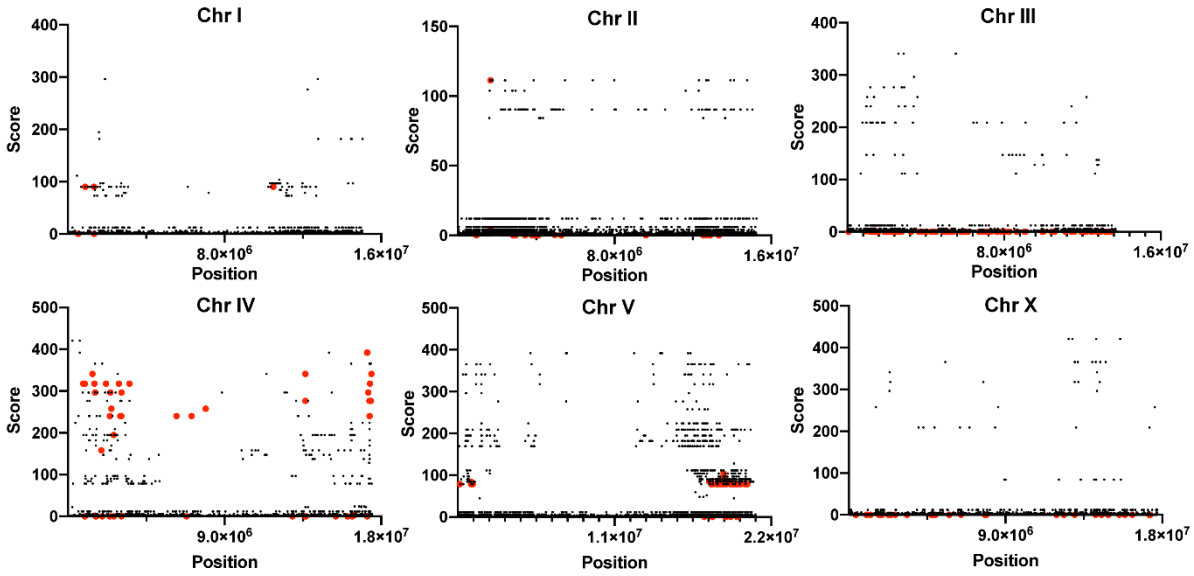
Supplementary Figure S10. Comparison of the polymorphisms and neutrality test statistics for synonymous and nonsynonymous SNVs. Mean values for *Pi*, Tajima's *D*, and Fay and Wu's *H* for different groups of genes calculated using synonymous or nonsynonymous SNVs. To compare the same set of genes for average *Pi*, we included the genes which have no synonymous or nonsynonymous SNVs ($Pi = 0$). So, the mean of *Pi* is slightly smaller than that in Figure 1E, which excluded the genes without nonsynonymous SNVs.

		Non-Hawaiian population				Hawaiian population			
	Sampling methods	# of sites	*Pi*	Tajima's *D*	Fay and Wu's *H*	# of sites	*Pi*	Tajima's *D*	Fay and Wu's *H*
All genes	Scattered	5.4	0.00045	-0.74	-3.89	6.9	0.00092	-0.33	-2.13
	Pooled	5.2	0.00041	-0.77	-4.00	6.7	0.00088	-0.38	-2.28
F-box genes	Scattered	19.5	0.00245	-1.02	-11.03	24.2	0.00457	-0.44	-5.30
	Pooled	18.6	0.00220	-1.07	-11.56	23.0	0.00428	-0.50	-5.97
csGPCRs	Scattered	10.7	0.00101	-1.03	-9.43	13.2	0.00202	-0.53	-6.56
	Pooled	10.4	0.00092	-1.09	-9.64	12.8	0.00196	-0.56	-6.64
Srw genes	Scattered	22.6	0.00208	-1.29	-22.32	26.9	0.00384	-0.73	-16.31
	Pooled	22.0	0.00185	-1.39	-22.94	26.1	0.00372	-0.74	-16.72
TF	Scattered	5.6	0.00034	-0.76	-4.28	7.3	0.00077	-0.32	-2.58
	Pooled	5.5	0.00032	-0.78	-4.40	7.1	0.00074	-0.36	-2.66
Protein kinases	Scattered	4.3	0.00015	-0.82	-2.81	5.4	0.00038	-0.30	-1.07
	Pooled	4.1	0.00014	-0.84	-2.89	5.2	0.00035	-0.39	-1.29

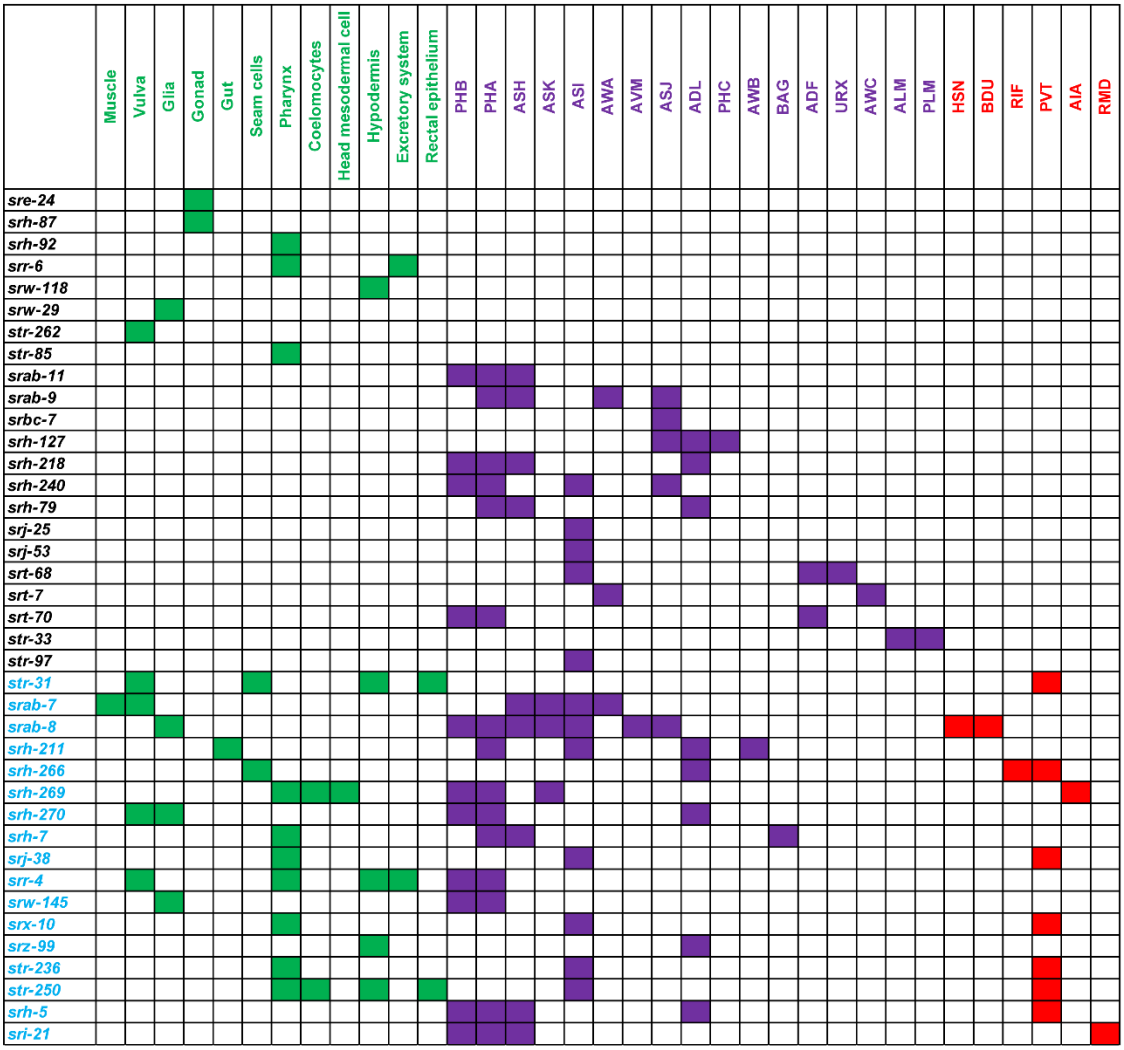
Supplementary Figure S11. Polymorphisms and neutrality test statistics calculated using the nonsynonymous SNV data of strains randomly chosen under two different sampling methods. Independent sampling were repeated 100 times to obtain 100 sets of strains and their SNVs. The mean of the number of segregating sites, *Pi*, *D*, and *H* for a given list of genes were calculated for each set of data. The average of these mean values are shown.



Supplementary Figure S12. The site frequency spectrum (SFS) pattern. (A) Site frequency spectrum of simulated SNV data under constant population size or bottleneck model. (B) The SFS pattern of the SNVs in the non-Hawaiian population from Chromosome I to X.



Supplementary Figure S13. Selective sweep site prediction based on site frequency spectrum. The distribution of α score predicted by SweeD for Chromosome I to X. X axis indicates chromosomal location. Red points indicate the alleles with significance level of 1%.



Supplementary Figure S14. The expression atlas of some csGPCRs under strong positive selection. The known expression pattern of csGPCR genes with Fay and Wu's *H* value smaller than -20 according to Vidal, et al. (2018). Tissue types were divided into sensory (purple blocks), interneurons and motor neurons (red blocks) and non-neuronal tissues (green). Genes in black are expressed in one of the three types of tissues and genes in blue are expressed in more than one types of tissue.