PNAS
www.pnas.org

Supplementary Information for

Polygenic adaptation and convergent evolution on growth and cardiac genetic pathways in African and Asian rainforest hunter-gatherers

C. M. Bergey, M. Lopez, G. F. Harrison, E. Patin, J. A. Cohen, L. Quintana-Murci, L. B. Barreiro, and G. H. Perry

Corresponding Authors: Christina M. Bergey and George H. Perry.
E-mail: cxb585@psu.edu and ghp3@psu.edu

This PDF file includes:

Supplementary text
Figs. S1 to S12
Tables S1 to S22
Captions for Databases S1 to S15
References for SI reference citations

Other supplementary materials for this manuscript include the following:

Databases S1 to S15

Supporting Information Text

Text S1 - Positive selection signatures on growth-associated genes

We examined whether gene-specific signatures of strong positive selection (using an "outlier-based" designation of genes with PBS index values < 0.01) in the rainforest populations were enriched for known functional associations with growth using *a priori* lists of 4,888 total growth-related genes, consisting of (with some redundancy among individual categories, as expected): i) 3,996 genes that affect growth or size in mice (MP:0005378) from the Mouse/Human Orthology with Phenotype Annotations database (1); ii) 266 genes associated with abnormal skeletal growth syndromes in the Online Mendelian Inheritance in Man (OMIM) database (https://omim.org/), as assembled by (2); iii) 427 genes expressed substantially more highly in the mouse growth plate, the cartilaginous region on the end of long bones where bone elongation occurs, than in soft tissues (lung, kidney, heart; >= 2.0 fold change; (3); and iv) 955 genes annotated with the Gene Ontology "growth" biological process (GO:0040007). Separately, we also considered in our analyses the set of 166 genes located within the 16 genomic regions previously associated with the pygmy phenotype in the Batwa, using an admixture mapping approach (4), as well as GH1- and IGF1-associated genes using data from OPHID database of protein–protein interaction (PPI) networks (5).

We used each of the curated *a priori* growth-related gene lists for testing the hypothesis that such loci are enriched for genes with signatures of strong positive selection (outlier PBS selection index values) or have a shift in the distribution of PBS selection index values consistent with subtle polygenic adaptation in the Batwa and Andamanese rainforest hunter-gatherer but not the Bakiga and Brahmin agriculturalist populations. We identified 202, 188, 291, and 252 outlier strong selection candidate genes (with PBS index values < 0.01) in each of the Batwa, Bakiga, Andamanese, and Brahmin populations, respectively. Genes in the *a priori* growth-related gene lists were not significantly overrepresented among PBS outliers in any populations, except for those associated with mouse growth phenotype in the Brahmin (68 observed, 47.7 expected; Fisher $p = 0.0179$) (Dataset S1i). Though the lack of over-representation of growth-related gene lists among loci with outlier signatures of strong positive natural selection related to growth is perhaps unsurprising considering the polygenic phenotype, our distribution shift-based test also showed no significant shifts in the distribution of PBS indices for any population (Dataset S1k). Genes in genomic regions previously associated with the pygmy phenotype in the Batwa (4) were enriched for genes with outlier PBS selection index values in the Batwa (outlier-based test: 5 observed, 1.39 expected; Fisher $p = 0.017$; Dataset S1i) and the PBS distribution for the phenotype-associated genes was shifted relative to the genome-wide distribution (distribution shift-based test: Kolmogorov-Smirnov test $p = 0.056$; Dataset S1k). We found no evidence that genes associated with GH1 and IGF1 were enriched for outlier or polygenic selection.

Text S2 - Impact of cross-annotated genes between growth factor- and cardiac-related pathways

To assess whether shared genes in GO categories relating to the heart and growth factor binding were responsible for the significant shift in PBS selection index values for genes in these annotations, we compared the distributions of PBS selection indices before and after removing 9 genes common to heart pathways and growth factor binding. The heart GO terms assessed were: 'cardiocyte differentiation' (GO:0035051), with a shift in the Andamanese hunter-gatherers; 'cardiac ventricle development' (GO:0003231), with a shift in the Batwa hunter-gatherers; and 'cardiac muscle tissue development' (GO:0048738) with a convergent shift in the Batwa and Andamanese. Of the 123 heart related genes contained in these pathways, 9 were also annotated to the GO molecular function 'growth factor binding' (GO:0019838): *ACVR1*, *EGFR*, *ENG*, *FGFR2*, *FGFRL1*, *LTBP1*, *SCN5A*, *TGFBR1*, and *TGFBR3*.

After removing the 9 shared genes, the mean PBS selection index for the Andamanese among genes annotated to 'cardiocyte differentiation' decreased slightly from 0.444 to 0.443 and the pre- and post-filtration distributions were not significantly different (Kolmogorov-Smirnov $D = 0.023$, $p = 1$). Similarly, the mean PBS selection index for the Batwa for genes in 'cardiac ventricle development' decreased slightly from 0.654 to 0.652, and the distributions were not significantly different ($D = 0.044$, $p = 1$). Finally, for 'cardiac muscle tissue development', the mean PBS selection index for the Andamanese increased from 0.450 to 0.453, and for the Batwa increased from 0.474 to 0.486. Again the pre- and post-filtering distributions were not significantly different for the Andamanese ($D = 0.015$, $p = 1$) or Batwa ($D = 0.015$, $p = 1$).

Similarly, after removing 9 shared genes, the mean PBS selection index for genes annotated to 'growth factor binding' (GO:0019838) for the Batwa increased slightly from 0.437 to 0.440 and for the Andamanese decreased from 0.455 to 0.437. Again, the pairs of distributions were not significantly different (Batwa: $D = 0.030$, $p = 1$; Andamanese: $D = 0.036$, $p = 1$).

Text S3 - Correcting for potential bias from differing gene size or global minor allele frequency (MAF)

In order to assess the potential biases related to differences in gene length (e.g. number of SNPs) or in SNP global minor allele frequencies (MAF), we repeated the analysis after modifying how the PBS selection index was computed. As in the uncorrected analysis, these corrected PBS selection index values were computed using 1,000 iterations.

First, to control for gene size, we sampled the PBS values for each SNP from only genes with the same number of SNPs during the computation of the selection index. For larger genes, gene sizes were binned to ensure sufficient SNPs from which to sample, using sets $[11, 15]$, $[16, 20]$, and $[21, \infty)$.

Second, to control for differing MAF values for SNPs, we did the permutation-based computation of the PBS selection index while matching SNPs on global MAF (computed using the African or Asian datasets for within-continent analyses.) SNPs were grouped by MAF into bins of size 0.01, and for each SNP in a gene, SNPs were sampled from only the set in the MAF bin.

Neither modification to the PBS selection index computation algorithm majorly affected the PBS selection index values nor the GO-based downstream analyses. Corrected and uncorrected PBS selection index values were highly correlated ($R^2 = 0.993$ to 0.997 and 0.953 to 0.985 for the gene size- and MAF-corrections respectively; SI Appendix, Fig. S4).

The GO biological processes and molecular functions with the strongest evidence of enrichment for strong selection were similar for the convergent (SI Appendix, Tables S5 and S6, Figs. S5 and S6) and population-specific selection analyses (SI Appendix, Tables S9 and S10, Figs. S5 and S6). The only mentioned growth- or heart-associated pathway that was no longer significant after correction was the biological process "negative regulation of growth," which was significantly enriched for genes with evidence of strong selection in the Batwa in the original analysis, but its p-value rose to 0.0448 after correction for gene size. In contrast, "cardiac muscle tissue development" (GO:0048738) which originally had a convergent empirical p-value of 0.025, was significantly enriched for strong positive selection convergently in the Batwa and Andamanese after MAF-based filtration ($p = 0.001$).

Similarly, the top GO categories with evidence of polygenic selection were largely unchanged for the convergent (SI Appendix, Tables S13 and S14, Figs. S9 and S10) and population-specific selection analyses (SI Appendix, Tables S17 and S18, Figs. S9 and S10). Minor changes include "growth factor binding" (GO:0019838) which rose to be no longer significant with the MAF-based correction (original convergent empirical $p < 0.001$; MAF corrected $p = 0.005$).

Text S4 - Modification of significance testing in empirical test for convergent evolution

We also modified and repeated the analysis that computes the significance of the convergence GO tests using a permutation-based approach. Whereas we originally permuted gene-PBS relationships to generate the random null distributions of PBS selection index values for two populations considered jointly, we instead permuted the gene-GO relationships to preserve LD patterns. The one-to-many relationships between genes and GO terms were shuffled, maintaining the groupings of GO terms that were assigned together to an original gene. We repeated the GO-based analyses for enrichment of strong selection or polygenic selection 1,000 times with these randomized gene-GO annotations, and compared our actual observed values to this randomly-generated null distribution. As before, we then defined the p-value of our empirical test for convergent evolution as the probability that this statistic was more extreme (lower) for the observed values than for the randomly generated values. The resultant p-value summarizes the test of the null hypothesis that both results could have been jointly generated under random chance. The results of the modified test were only slightly different than the original for both convergence in strong outlier selection (SI Appendix, Table S21) and in a shifted PBS selection index (SI Appendix, Table S22).

Text S5 - Assessment of bias due to uncertainty in allele frequency

Given the differing samples sizes available for each population, we performed a simple analysis to determine how allele frequency uncertainty could have affected our results. First, we subsampled the larger Batwa and Bakiga populations down to random sets of no more than 30 individuals (the size of the British and Kenyan outgroup populations), recomputed F_{ST} and PBS, and performed the gene ontology-based tests with this modified input. The analyses based on the subsampled populations did appear to have slightly lower power to detect signatures of selection based on relatively higher p-values of the most significant results compared to the analysis of our full sample. However, our main results typically remained among the top enriched GO terms for each analysis. Specifically:

- As in the full analysis, the set of GO terms with evidence of significant convergent enrichment for strong positive selection included GO biological processes related to limb morphogenesis including 'embryonic limb morphogenesis' (empirical test for convergence $p < 0.001$; SI Appendix, Fig. S7, Table S7).

- The two highlighted growth-related GO terms remained among the set of terms with significant population-specific enrichment for strong positive selection (SI Appendix, Fig. S7, Table S11): 'negative regulation of growth' was significantly enriched for the Batwa ($p = 0.005$) and 'negative regulation of cell differentiation' for the Andamanese ($p = 0.0045$).

- As in the full analysis, the GO terms with significant evidence of convergent shifts in PBS select index values included 'growth factor binding' (empirical test for convergence $p = 0.003$), but 'organ growth' and 'cardiac muscle tissue development' were no longer significant (SI Appendix, Fig. S11, Table S15).

- Finally, the GO terms with significant population-specific shifts in distribution of PBS selection index again included the heart-related terms 'cardiac ventricle development' for the Batwa ($p = 0.001$) and 'cardiocyte differentiation' for the Andamanese ($p = 0.009$), though related term 'cardiac muscle tissue development' was not significant (SI Appendix, Fig. S11, Table S19).

We also performed a second subsampling analysis because our populations with the smallest sample sizes, the Andamanese and Brahmin, were represented by 10 individuals per population. We repeated all of our analyses after randomly subsampling each of the Batwa and Bakiga populations down to 10 from 50 individuals. We again observed growth- and heart-related genetic pathways to be the most enriched for signatures of selection, consistent with the original analysis:

- GO biological processes related to limb morphogenesis remained among the set of GO terms with evidence of convergent enrichment for strong positive selection, including 'embryonic limb morphogenesis' (empirical test for convergence $p < 0.001$; SI Appendix, Fig. S8, Table S8).

- For population-specific enrichment for strong positive selection, ‘negative regulation of growth’ was not significantly enriched for the Batwa (SI Appendix, Fig. S8, Table S12), unlike in the full analysis. The growth-related GO term, ‘negative regulation of cell differentiation,’ remained significantly enriched for the Andamanese ($p = 0.004$).

- The GO terms with significant evidence of convergent shifts in PBS select index values included ‘growth factor binding’ ($p = 0.003$), ‘organ growth’ ($p = 0.002$), and ‘cardiac muscle tissue development’ ($p = 0.003$), as in the full analysis. A related term, ‘heart morphogenesis,’ was also significantly enriched ($p = 0.002$; SI Appendix, Fig. S12, Table S16).

- Finally, as before, the terms with significant population-specific shifts in distribution of PBS selection index included the heart-related terms ‘cardiac ventricle development’ for the Batwa ($p = 0.003$) and ‘cardiocyte differentiation’ for the Andamanese ($p = 0.010$), though related term ‘cardiac muscle tissue development’ was not significant (SI Appendix, Fig. S12, Table S20).

Fig. S1: Population Branch Statistic (PBS) schematic.

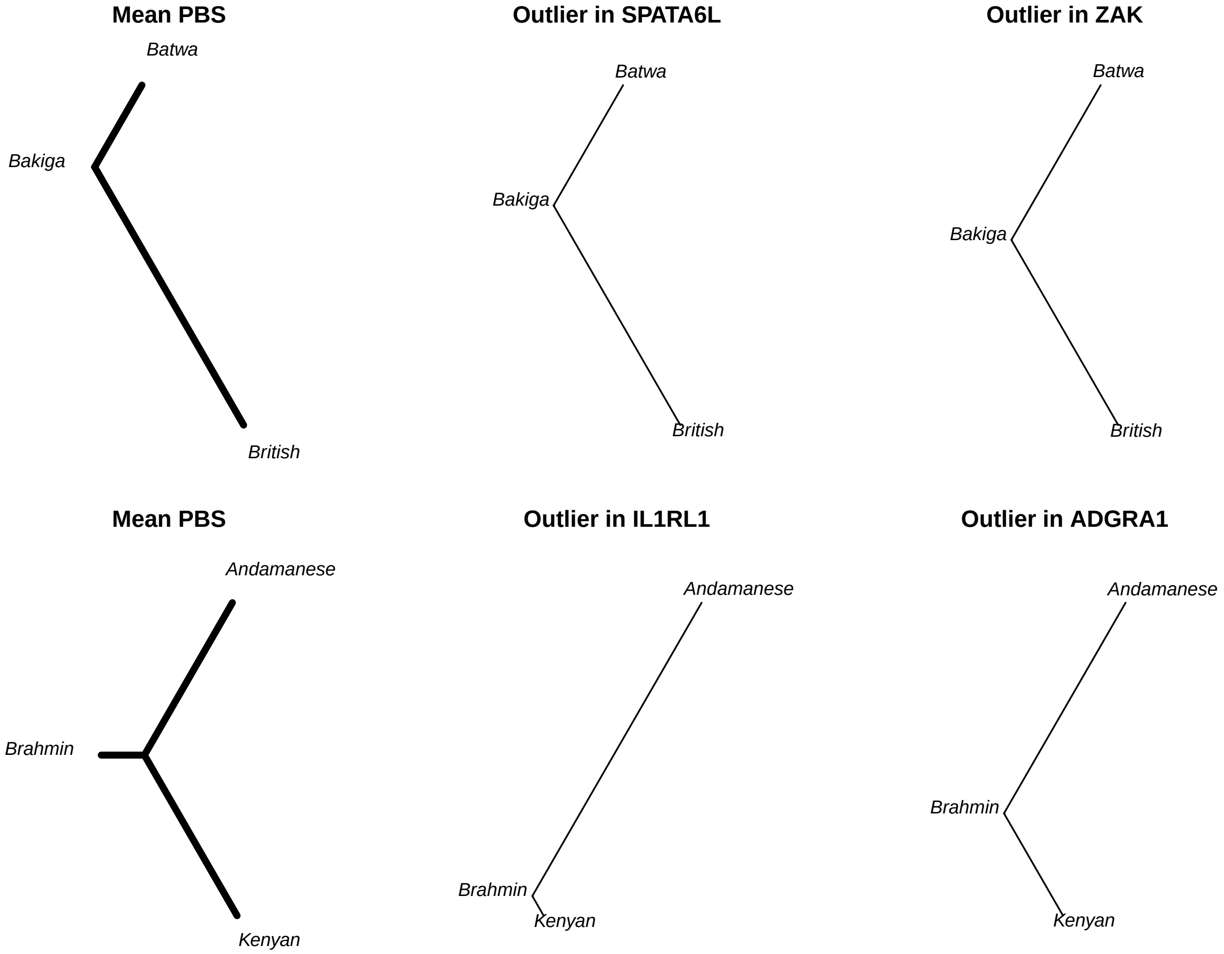


Fig. S1. Mean values of the Population Branch Statistic (PBS; left) for the African dataset (Batwa, Bakiga, and outgroup British populations; upper row) and Asian dataset (Andamanese, Brahmin, and outgroup Kenyan populations; lower row). Middle and right columns contain PBS values for two outlier SNPs in each population.

Fig. S2: Population Branch Statistic (PBS) by SNP.

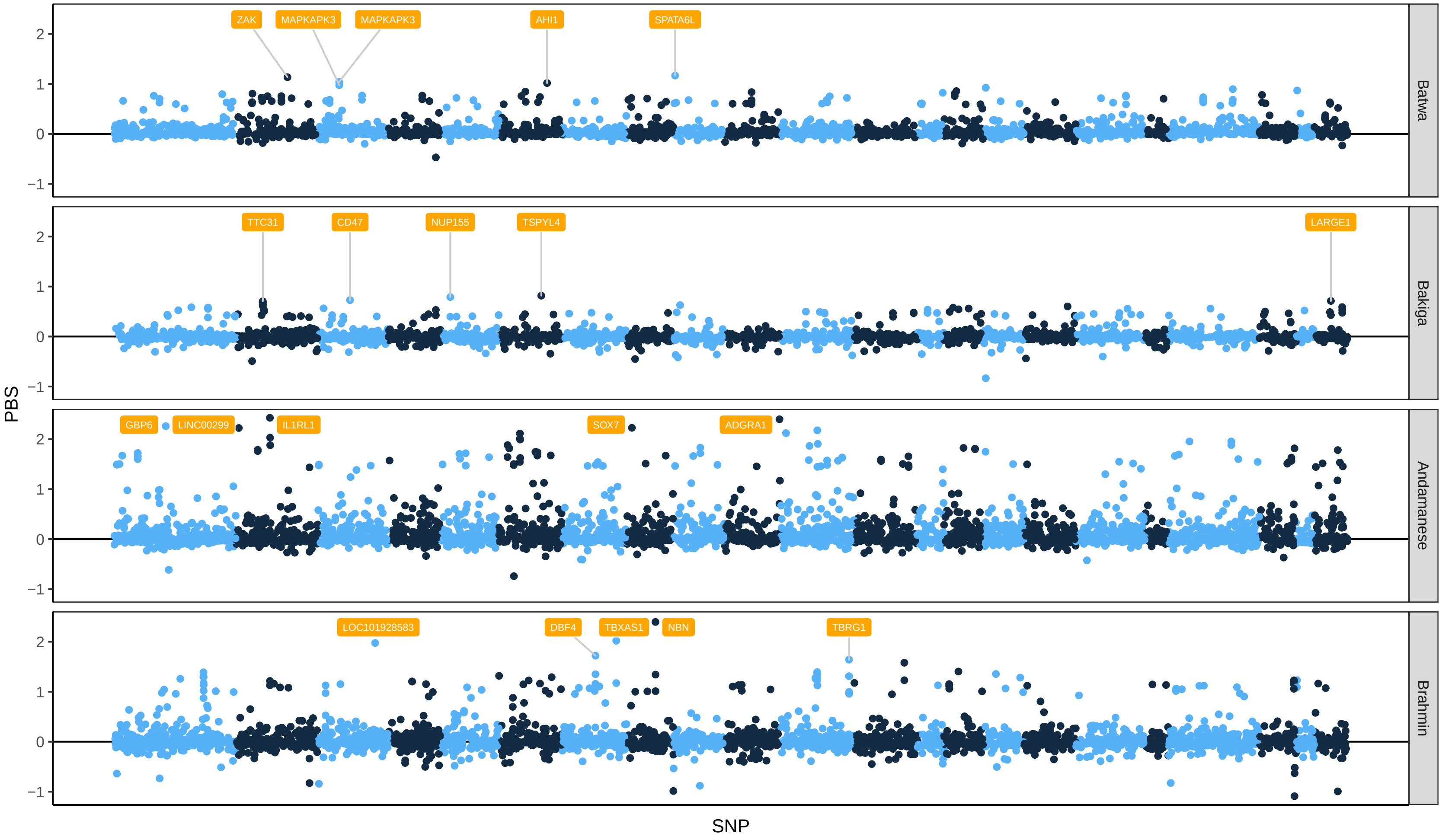


Fig. S2. Population Branch Statistic (PBS) values plotted across the genome for the four focal populations. The genes containing the SNPs with the 5 highest PBS values in each population are labeled.

Fig. S3: Population Branch Statistic (PBS) by gene SNP count.

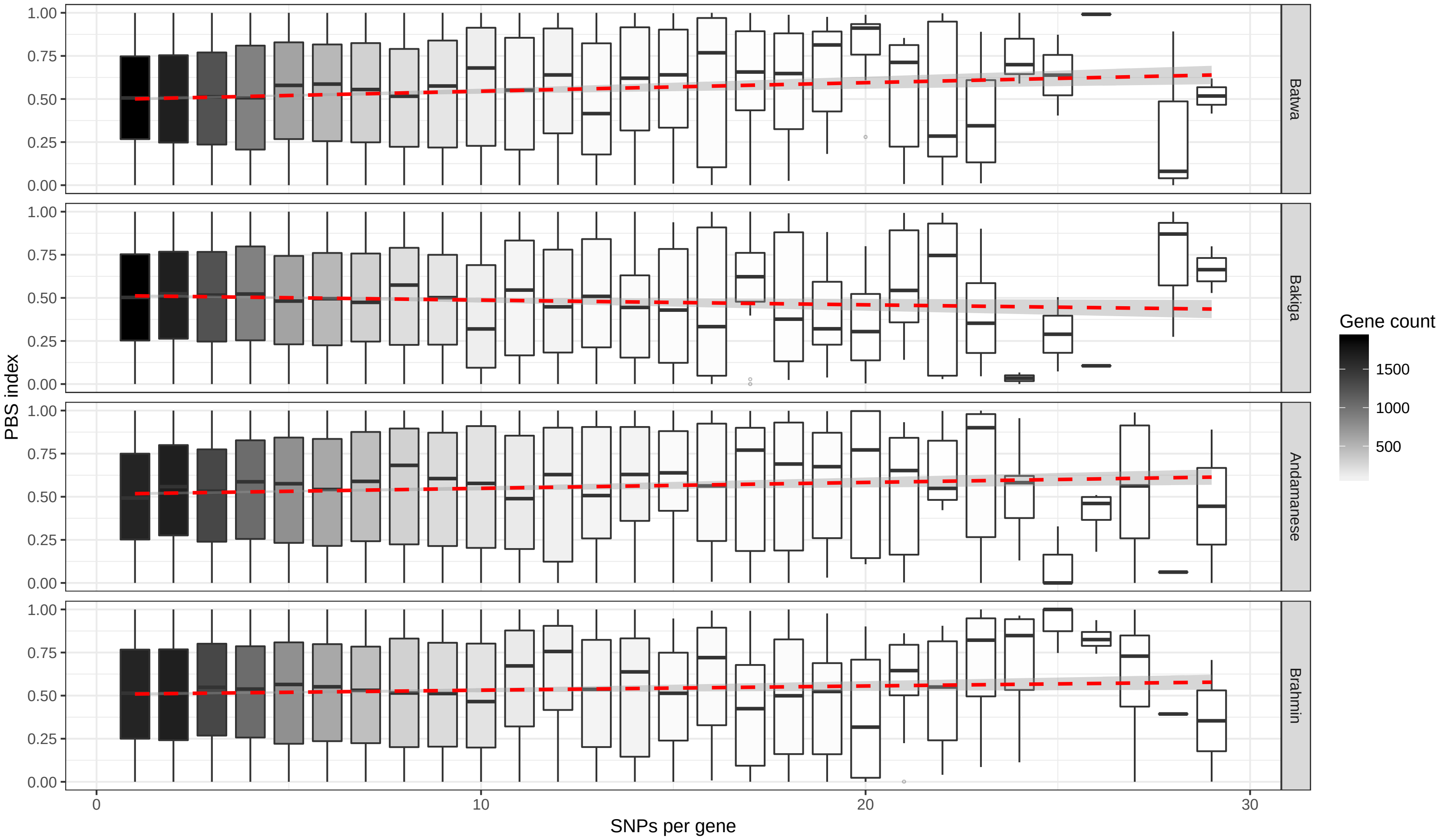


Fig. S3. Population Branch Statistic (PBS) selection index values plotted by number of SNPs in gene. Color indicates number of genes with that SNP count. Only SNP counts from 1 to 30 shown.

Fig. S4: Gene size- and MAF-based corrections' impact on p-value.

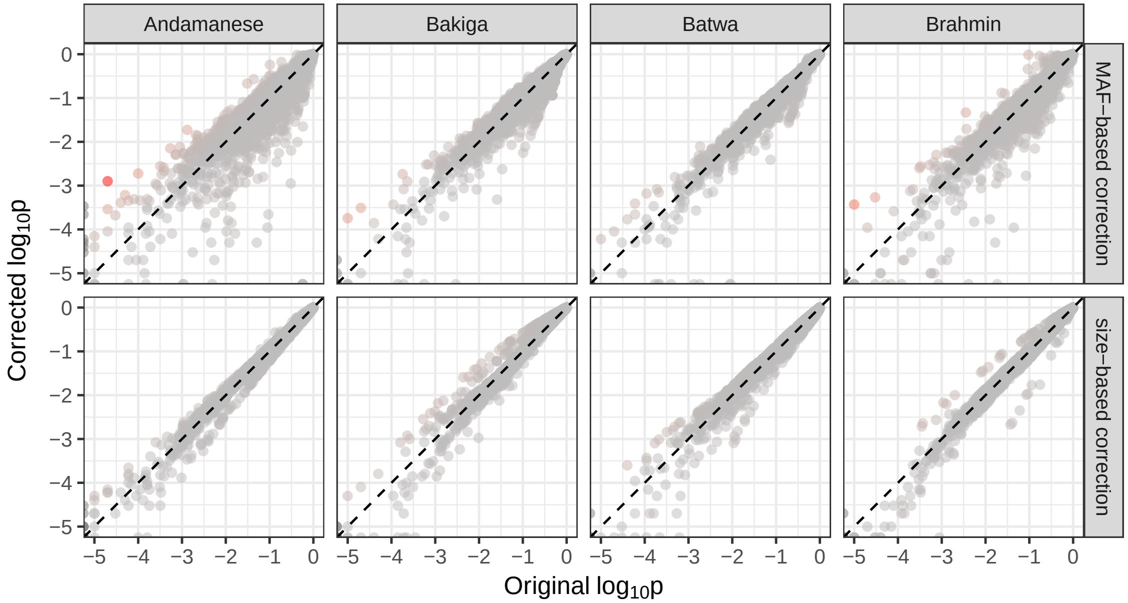
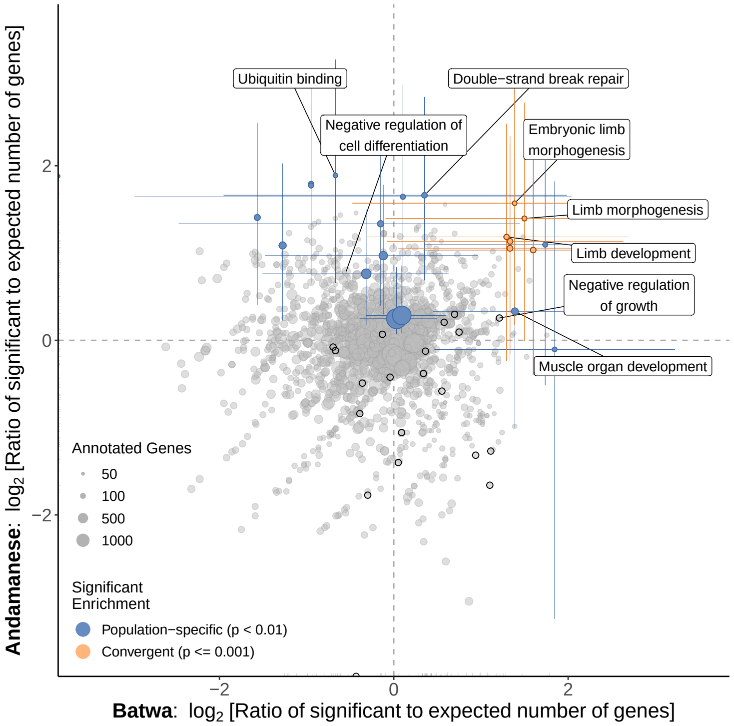


Fig. S4. Plots of PBS selection index values for genes corrected for gene size and MAF shown compared to the original uncorrected values (with both plotted on a logarithmic scale. Red shading indicates higher percent difference from original value.

Fig. S5: Gene size-corrected strong positive selection enrichment results.

A. Rainforest hunter-gatherers



B. Agriculturalists

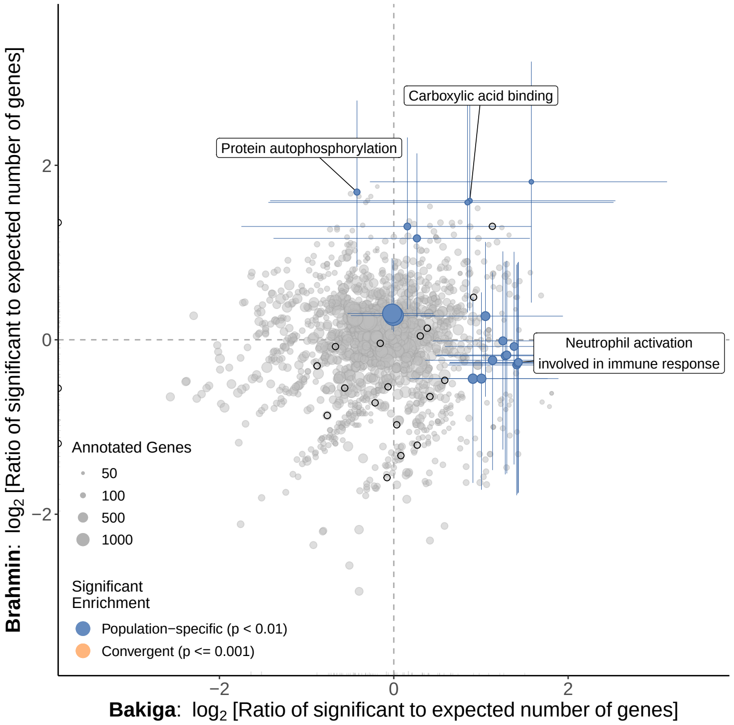
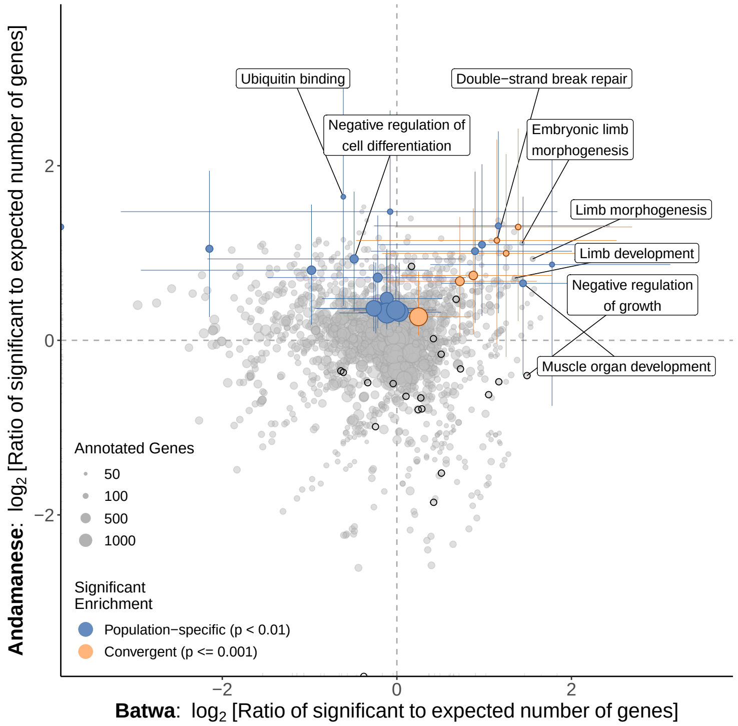


Fig. S5. After gene size-based correction, Gene Ontology (GO) functional categories' ratios of expected to observed counts of outlier genes (with PBS selection index < 0.01) in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly overrepresented for genes under positive selection (Fisher $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

Fig. S6: MAF-corrected strong positive selection enrichment results.

A. Rainforest hunter-gatherers



B. Agriculturalists

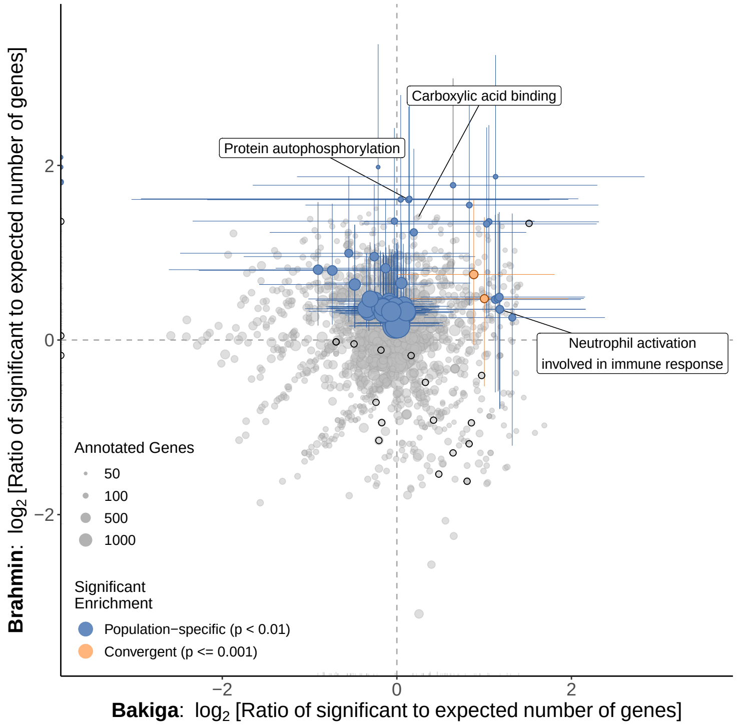
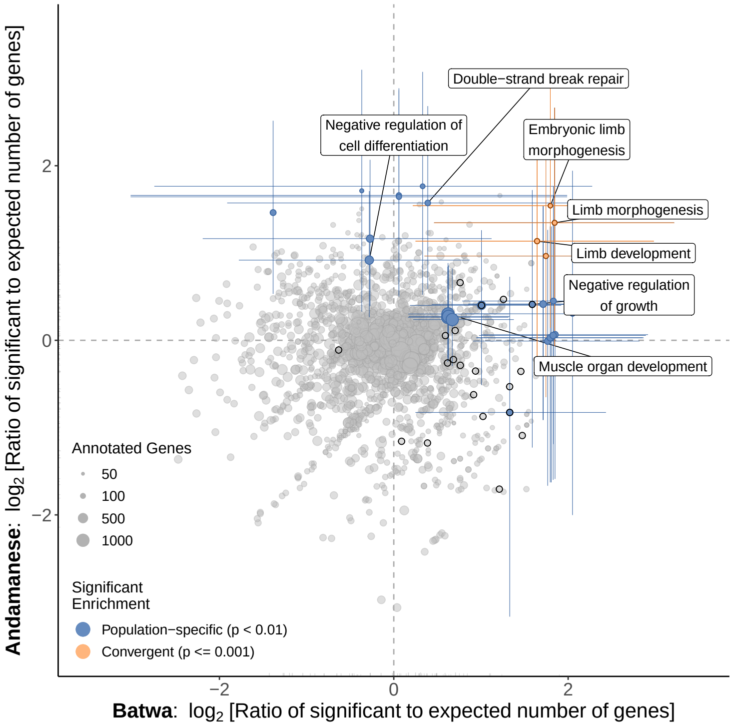


Fig. S6. After MAF-based correction, Gene Ontology (GO) functional categories' ratios of expected to observed counts of outlier genes (with PBS selection index < 0.01) in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly overrepresented for genes under positive selection (Fisher $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

Fig. S7: Strong positive selection enrichment results after subsampling to $\leq$ 30 individuals per population.

A. Rainforest hunter-gatherers



B. Agriculturalists

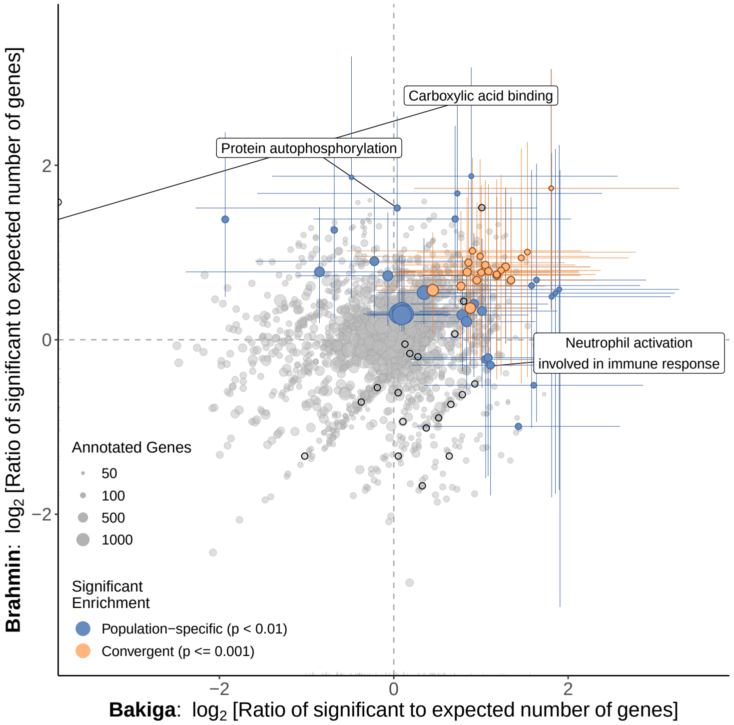
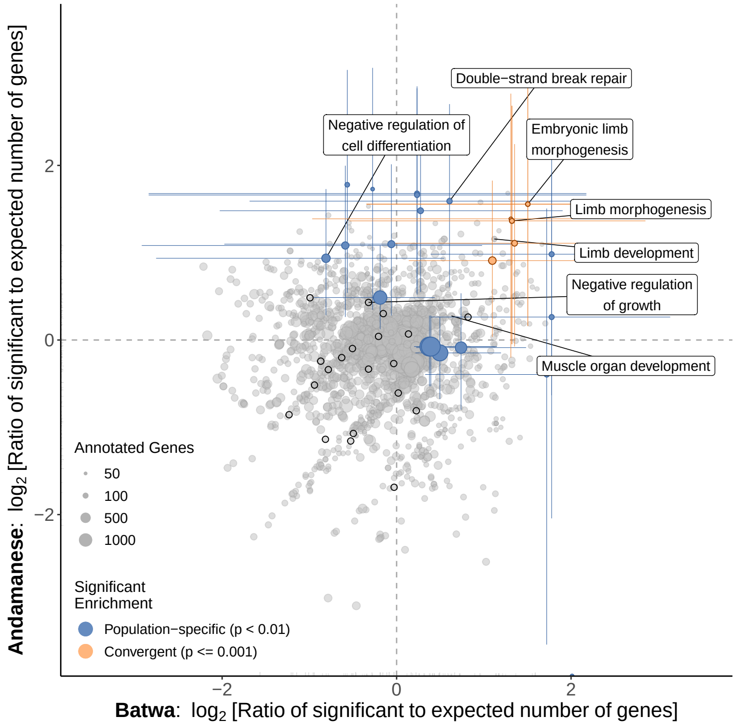


Fig. S7. After subsampling to $\leq$ 30 individuals per population, Gene Ontology (GO) functional categories' ratios of expected to observed counts of outlier genes (with PBS selection index $<$ 0.01) in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly overrepresented for genes under positive selection (Fisher $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

Fig. S8: Strong positive selection enrichment results after subsampling Batwa and Bakiga to 10 individuals per population.

A. Rainforest hunter-gatherers



B. Agriculturalists

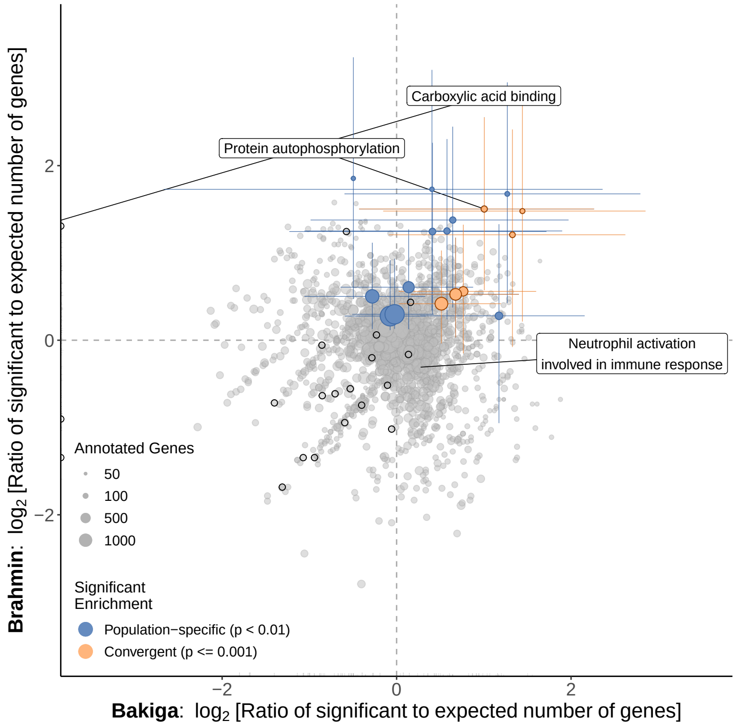
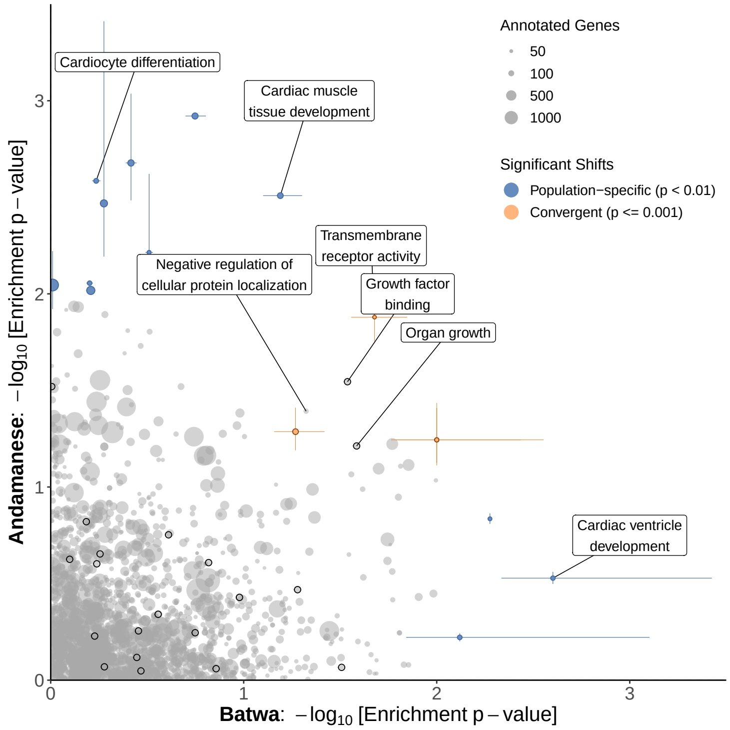


Fig. S8. After subsampling Batwa and Bakiga to 10 individuals per population, Gene Ontology (GO) functional categories' ratios of expected to observed counts of outlier genes (with PBS selection index < 0.01) in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly overrepresented for genes under positive selection (Fisher $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

Fig. S9: Gene size-corrected polygenic distribution shift test results.

A. Rainforest hunter-gatherers



B. Agriculturalists

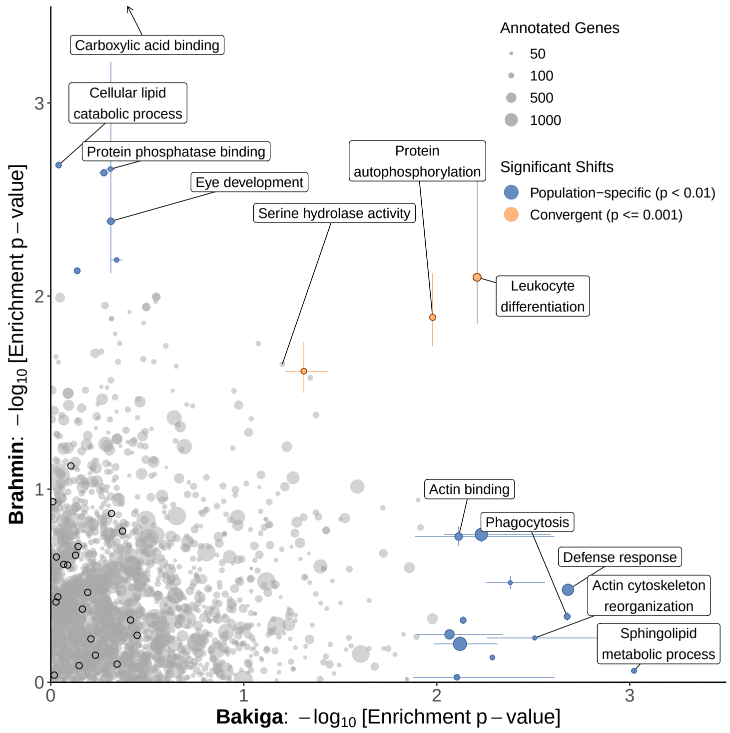
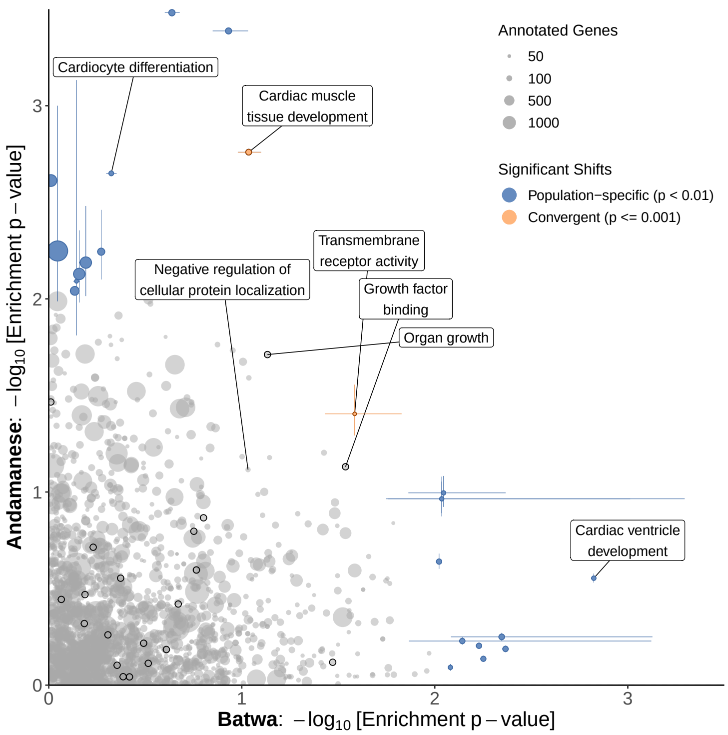


Fig. S9. After gene size-based correction, Gene Ontology (GO) functional categories' distribution shift test p-values, indicating a shift in the PBS selection index values for genes, in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly enriched for genes under positive selection (Kolmogorov-Smirnov $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories. One GO molecular function, "carboxylic acid binding" (GO:0031406; Brahmin $p = 7.3 \times 10^{-5}$; $q = 0.0157$) not shown.

Fig. S10: Gene size-corrected polygenic distribution shift test results.

A. Rainforest hunter-gatherers



B. Agriculturalists

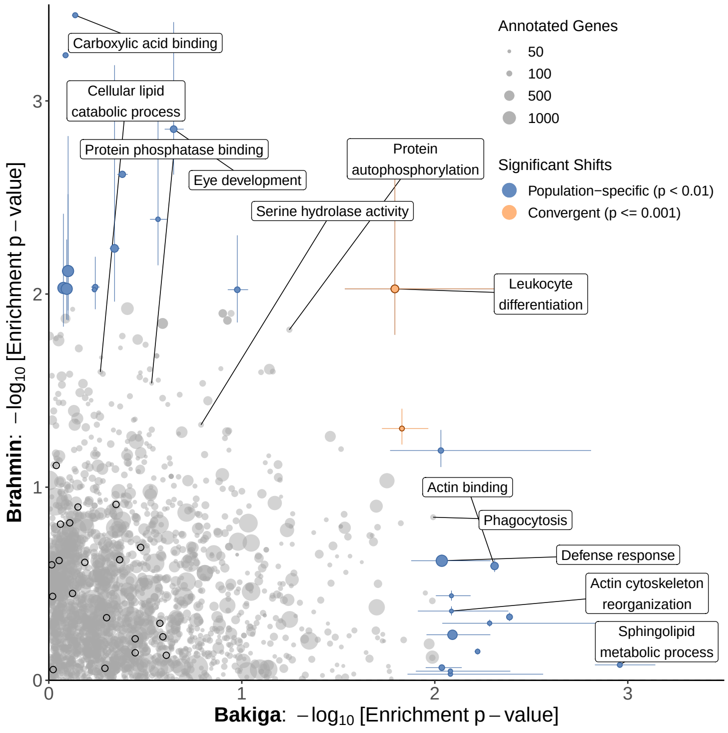
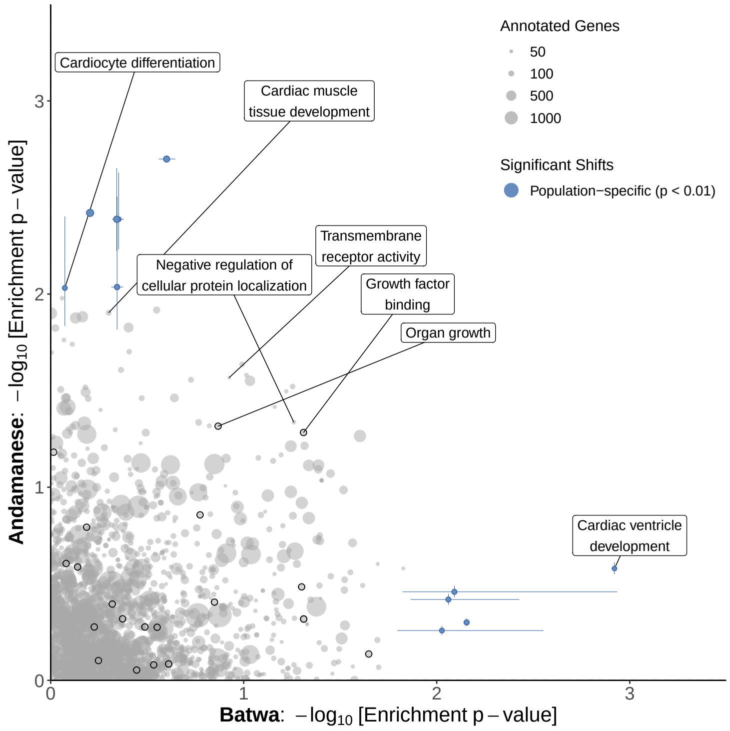


Fig. S10. After MAF-based correction, Gene Ontology (GO) functional categories' distribution shift test p-values, indicating a shift in the PBS selection index values for genes, in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly enriched for genes under positive selection (Kolmogorov-Smirnov $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

Fig. S11: Gene size-corrected polygenic distribution shift test results after subsampling to $\leq$ 30 individuals per population.

A. Rainforest hunter-gatherers



B. Agriculturalists

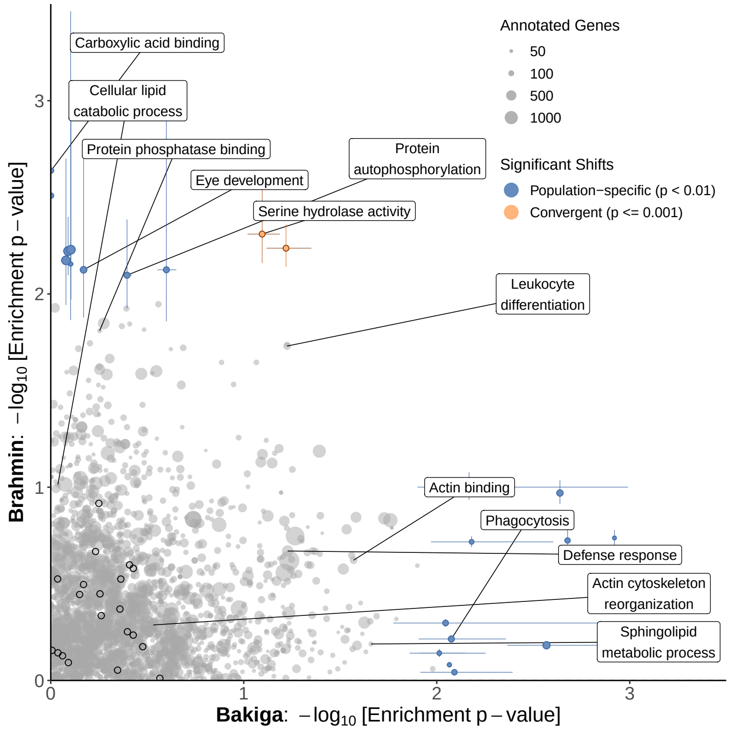
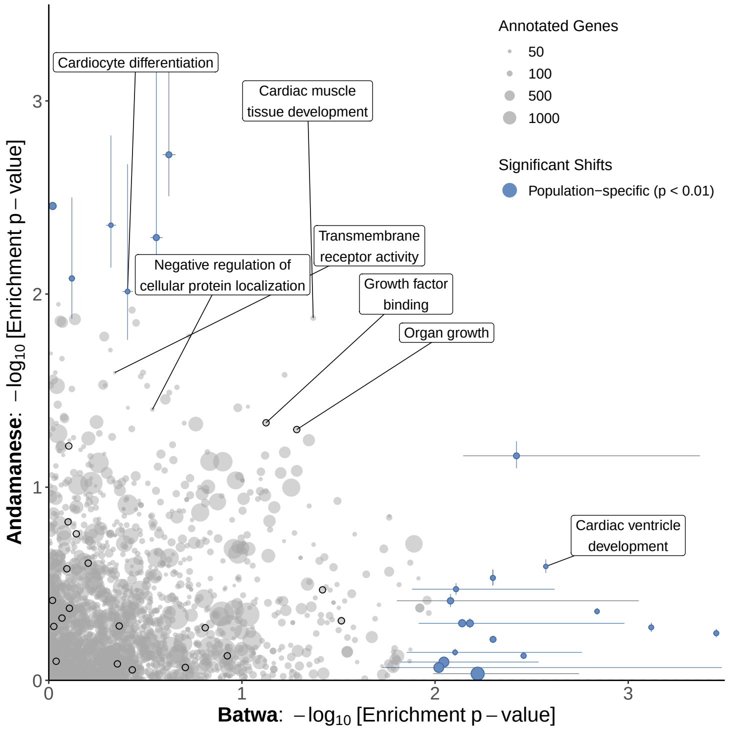


Fig. S11. After subsampling to $\leq$ 30 individuals per population, Gene Ontology (GO) functional categories' distribution shift test p-values, indicating a shift in the PBS selection index values for genes, in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly enriched for genes under positive selection (Kolmogorov-Smirnov $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

Fig. S12: Gene size-corrected polygenic distribution shift test results after subsampling Batwa and Bakiga to 10 individuals per population.

A. Rainforest hunter-gatherers



B. Agriculturalists

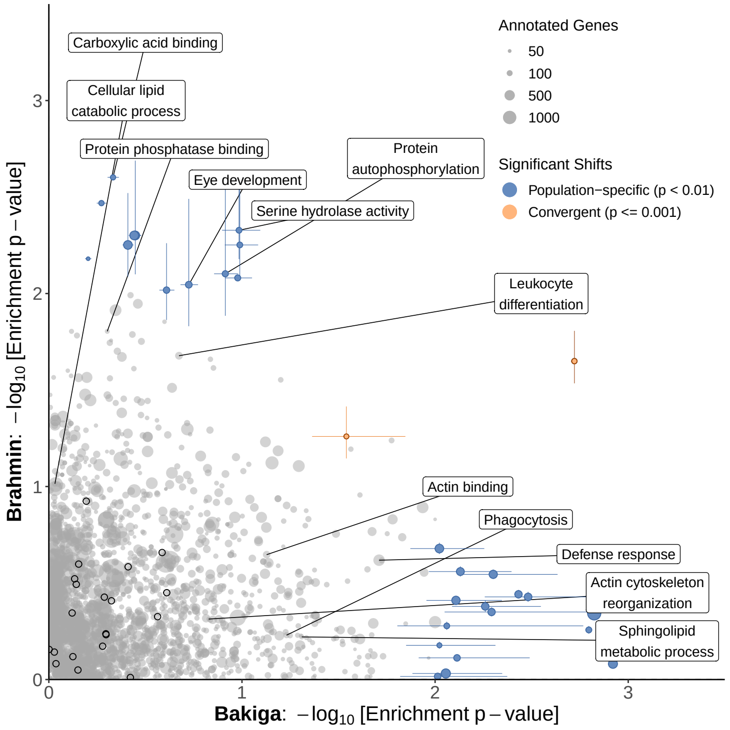


Fig. S12. After subsampling Batwa and Bakiga to 10 individuals per population, Gene Ontology (GO) functional categories' distribution shift test p-values, indicating a shift in the PBS selection index values for genes, in the Batwa and Andamanese rainforest hunter-gatherers (A) and Bakiga and Brahmin agriculturalist control (B). Results shown for GO biological processes and molecular functions. Point size is scaled to number of annotated genes in category. Terms that are significantly enriched for genes under positive selection (Kolmogorov-Smirnov $p < 0.01$) in either population shown in blue and for both populations convergently (empirical permutation-based $p < 0.005$) shown in orange. Colored lines represent 95% CI for significant categories estimated by bootstrapping genes within pathways. Black outlines indicate growth-associated terms: the 'growth' biological process (GO:0040007) and its descendant terms, or the molecular functions 'growth factor binding,' 'growth factor receptor binding,' 'growth hormone receptor activity,' and 'growth factor activity' and their sub-categories.

1. Tables

Table S1. Gene Ontology (GO) biological processes with evidence of convergent enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. No molecular functions were found to be convergently enriched. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

			Batwa:				*Andamanese:*			
	GO Biological Process	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0030326	embryonic limb morphogenesis	0.000	1.47	4	0.0584	1	2.01	6	0.0147	0.901
GO:0035107	appendage morphogenesis	0.000	1.69	5	0.0267	1	2.27	6	0.0254	0.901
GO:0035108	limb morphogenesis	0.000	1.69	5	0.0267	1	2.27	6	0.0254	0.901
GO:0035113	embryonic appendage morphogenesis	0.000	1.47	4	0.0584	1	2.01	6	0.0147	0.901
GO:0048736	appendage development	0.001	1.95	5	0.0448	1	2.62	6	0.047	1.000
GO:0060173	limb development	0.001	1.95	5	0.0448	1	2.62	6	0.047	1.000
GO:0030048	actin filament-based movement	0.002	1.72	4	0.0927	1	2.33	5	0.0834	1.000
GO:0048705	skeletal system morphogenesis	0.002	2.20	5	0.0688	1	2.95	6	0.0743	1.000
GO:0007034	vacuolar transport	0.003	1.37	4	0.0470	1	1.75	4	0.0968	1.000

Table S2. Gene Ontology (GO) biological processes and molecular functions with evidence of population-specific enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Results with $p < 0.01$ shown.

GO		Exp.	Obs.	p	adj. p
Batwa RHG - Biological Processes:					
GO:0007517	muscle organ development	10	4.02	0.0069	0.708
GO:0045926	negative regulation of growth	7	2.48	0.0118	0.708
Andamanese RHG - Biological Processes:					
GO:0006302	double-strand break repair	10	3.14	0.0011	0.171
GO:0070085	glycosylation	12	4.34	0.0013	0.171
GO:0000723	telomere maintenance	8	2.30	0.0020	0.175
GO:0033365	protein localization to organelle	25	14.09	0.0036	0.189
GO:1903827	regulation of cellular protein localization	19	9.66	0.0036	0.189
GO:0007569	cell aging	6	1.62	0.0052	0.208
GO:0009101	glycoprotein biosynthetic process	12	5.28	0.0065	0.208
GO:0034613	cellular protein localization	41	27.87	0.0067	0.208
GO:0051179	localization	116	97.30	0.0079	0.208
GO:0060249	anatomical structure homeostasis	13	6.09	0.0079	0.208
GO:0045596	negative regulation of cell differentiation	18	9.79	0.0090	0.215
Batwa RHG - Molecular Functions:					
GO:0003723	RNA binding	26	17.66	0.028	0.732
GO:0043167	ion binding	83	70.68	0.034	0.732
Andamanese RHG - Molecular Functions:					
GO:0043130	ubiquitin binding	7	1.89	0.0026	0.177
GO:0008233	peptidase activity	17	9.58	0.0153	0.383

Table S3. Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of convergent distribution shifts in PBS selection index values in the hunter-gatherer populations. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

				Batwa:		*Andamanese:*	
		GO	Joint p	p	adj. p	p	adj. p
BP	GO:0035265	organ growth	0.001	0.0275	0.997	0.04509	1.000
	GO:0048738	cardiac muscle tissue development	0.001	0.0461	0.997	0.00265	1.000
	GO:1903828	negative regulation of cellular protein localization	0.001	0.0360	0.997	0.04275	1.000
	GO:0016202	regulation of striated muscle tissue development	0.002	0.0135	0.997	0.04406	1.000
	GO:1901861	regulation of muscle tissue development	0.002	0.0135	0.997	0.04406	1.000
	GO:0045444	fat cell differentiation	0.004	0.0573	0.997	0.04058	1.000
MF	GO:0019199	transmembrane receptor protein kinase activity	0.000	0.027	0.817	0.0261	0.784
	GO:0019838	growth factor binding	0.000	0.021	0.817	0.0269	0.784
	GO:0032559	adenyl ribonucleotide binding	0.003	0.020	0.817	0.0579	0.877
	GO:0030554	adenyl nucleotide binding	0.004	0.017	0.817	0.0755	0.877

Table S4. Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of population-specific distribution shifts in PBS selection index values in the hunter-gatherer populations. No molecular functions were found to be significantly shifted for the Batwa. Results with $p < 0.01$ are shown.

	GO	p	adj. p
Batwa RHG - Biological Processes:			
GO:0003231	cardiac ventricle development	0.001	0.302
GO:0061351	neural precursor cell proliferation	0.007	0.348
GO:0034976	response to endoplasmic reticulum stress	0.009	0.348
Andamanese RHG - Biological Processes:			
GO:0016579	protein deubiquitination	0.001	0.232
GO:0035051	cardiocyte differentiation	0.002	0.232
GO:0048738	cardiac muscle tissue development	0.003	0.232
GO:1901800	positive regulation of proteasomal protein catabolic process	0.004	0.262
GO:0006508	proteolysis	0.009	0.453
Andamanese RHG - Molecular Functions:			
GO:0005085	guanyl-nucleotide exchange factor activity	0.005	0.278

Table S5. After gene size-based correction, Gene Ontology (GO) biological processes and molecular functions with evidence of convergent enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

			Batwa:				*Andamanese:*			
	GO Biological Process	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0035107	appendage morphogenesis	0.000	1.77	5	0.0315	1	2.28	6	0.0258	0.966
GO:0035108	limb morphogenesis	0.000	1.77	5	0.0315	1	2.28	6	0.0258	0.966
GO:0048736	appendage development	0.000	2.04	5	0.0524	1	2.64	6	0.0478	1.000
GO:0060173	limb development	0.000	2.04	5	0.0524	1	2.64	6	0.0478	1.000
GO:1901617	organic hydroxy compound biosynthetic process	0.000	2.38	6	0.0314	1	3.19	7	0.0401	0.980
GO:0030326	embryonic limb morphogenesis	0.001	1.53	4	0.0665	1	2.02	6	0.0150	0.966
GO:0035113	embryonic appendage morphogenesis	0.001	1.53	4	0.0665	1	2.02	6	0.0150	0.966
GO:0030048	actin filament-based movement	0.002	1.80	6	0.0089	1	2.34	5	0.0845	1.000
GO:0007034	vacuolar transport	0.004	1.43	4	0.0537	1	1.76	4	0.0979	1.000

			Batwa:				*Andamanese:*			
	GO Molecular Function	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0008514	organic anion transmembrane transporter activity	0.000	2.78	7	0.0212	0.6853	3.37	7	0.0515	0.902
GO:0015081	sodium ion transmembrane transporter activity	0.000	2.31	7	0.0081	0.6853	2.93	6	0.0726	0.902

Table S6. After MAF-based correction, Gene Ontology (GO) biological processes and molecular functions with evidence of convergent enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

			Batwa:				*Andamanese:*			
	GO Biological Process	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0048522	positive regulation of cellular process	0.000	55.57	66	0.0534	1	87.93	106	0.01153	0.946
GO:1901617	organic hydroxy compound biosynthetic process	0.000	2.29	6	0.0267	1	3.66	9	0.01069	0.946
GO:0006302	double-strand break repair	0.001	2.26	5	0.0757	1	3.62	8	0.02808	0.946
GO:0006897	endocytosis	0.001	8.49	14	0.0447	1	13.76	22	0.01948	0.946
GO:0048738	cardiac muscle tissue development	0.001	2.52	6	0.0399	1	4.01	8	0.04694	0.946
GO:0080135	regulation of cellular response to stress	0.001	7.63	14	0.0203	1	11.95	20	0.01624	0.946
GO:0014706	striated muscle tissue development	0.002	4.07	8	0.0509	1	6.55	14	0.00585	0.946
GO:0070302	regulation of stress-activated protein kinase signaling cascade	0.002	2.52	6	0.0399	1	3.93	7	0.09815	0.946
GO:1901615	organic hydroxy compound metabolic process	0.002	5.70	9	0.1168	1	8.87	14	0.06046	0.946
GO:2001020	regulation of response to DNA damage stimulus	0.002	2.09	5	0.0572	1	3.35	7	0.04997	0.946
GO:0051592	response to calcium ion	0.003	1.75	6	0.0079	1	2.74	5	0.13789	0.946
GO:0015718	monocarboxylic acid transport	0.004	1.63	4	0.0793	1	2.39	5	0.08980	0.946
GO:0030048	actin filament-based movement	0.004	1.73	4	0.0942	1	2.78	7	0.02039	0.946
GO:0030326	embryonic limb morphogenesis	0.004	1.48	4	0.0594	1	2.31	5	0.08052	0.946
GO:0031098	stress-activated protein kinase signaling	0.004	3.15	7	0.0383	1	4.97	8	0.12454	0.946
GO:0035113	embryonic appendage morphogenesis	0.004	1.48	4	0.0594	1	2.31	5	0.08052	0.946
GO:0060537	muscle tissue development	0.004	4.30	8	0.0660	1	6.90	14	0.00910	0.946
			Batwa:				*Andamanese:*			
	GO Molecular Function	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0008514	organic anion transmembrane transporter activity	0.002	2.68	6	0.051	0.798	4.03	10	0.0068	0.959
GO:0005342	organic acid transmembrane transporter activity	0.003	1.97	5	0.046	0.798	2.96	7	0.0282	1.000
GO:0015081	sodium ion transmembrane transporter activity	0.003	2.22	5	0.071	0.821	3.50	7	0.0600	1.000
GO:0046943	carboxylic acid transmembrane transporter activity	0.004	1.79	4	0.103	0.964	2.70	7	0.0177	1.000

Table S7. After subsampling to ≤ 30 individuals per population, Gene Ontology (GO) biological processes and molecular functions with evidence of convergent enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

			Batwa:				*Andamanese:*			
	GO Biological Process	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0030326	embryonic limb morphogenesis	0	1.44	5	0.0142	1	2.06	6	0.0165	1
GO:0035107	appendage morphogenesis	0	1.67	6	0.0062	1	2.36	6	0.0302	1
GO:0035108	limb morphogenesis	0	1.67	6	0.0062	1	2.36	6	0.0302	1
GO:0035113	embryonic appendage morphogenesis	0	1.44	5	0.0142	1	2.06	6	0.0165	1
GO:0048736	appendage development	0	1.92	6	0.012	1	2.73	6	0.0548	1
GO:0060173	limb development	0	1.92	6	0.012	1	2.73	6	0.0548	1
GO:0030048	actin filament-based movement	0.001	1.79	6	0.0088	1	2.56	5	0.1127	1
GO:1901617	organic hydroxy compound biosynthetic process	0.003	2.71	6	0.0539	1	3.92	7	0.0978	1
GO:0006282	regulation of DNA repair	0.004	1.29	3	0.1387	1	1.9	6	0.0112	1
GO:0006898	receptor-mediated endocytosis	0.004	3.41	7	0.0543	1	4.99	8	0.1266	1
			Batwa:				*Andamanese:*			
	GO Molecular Function	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0003779	actin binding	0.004	5.44	9	0.09466	0.643688	7	13	0.023	1
GO:0005342	organic acid transmembrane transporter...	0.004	2.01	6	0.01482	0.25194	2.71	5	0.134	1
GO:0046943	carboxylic acid transmembrane transporter...	0.004	2.01	6	0.01482	0.25194	2.71	5	0.134	1

Table S8. After subsampling Batwa and Bakiga to 10 individuals per population, Gene Ontology (GO) biological processes and molecular functions with evidence of convergent enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

			Batwa:				*Andamanese:*			
	GO Biological Process	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0030326	embryonic limb morphogenesis	0.000	1.41	4	0.0516	1	2.04	6	0.0157	1
GO:0035107	appendage morphogenesis	0.000	1.6	4	0.0757	1	2.33	6	0.0287	1
GO:0035108	limb morphogenesis	0.000	1.6	4	0.0757	1	2.33	6	0.0287	1
GO:0035113	embryonic appendage morphogenesis	0.000	1.41	4	0.0516	1	2.04	6	0.0157	1
GO:0071103	DNA conformation change	0.000	2.35	6	0.0301	1	3.71	8	0.0322	1
GO:0002822	regulation of adaptive immune response...	0.001	1.21	3	0.1206	1	1.91	5	0.0411	1
GO:0002700	regulation of production of molecular mediator of immune response	0.002	1.26	4	0.0368	1	1.87	4	0.1168	1
GO:1901617	organic hydroxy compound biosynthetic process	0.002	2.57	6	0.0436	1	3.88	7	0.0932	1
GO:0006282	regulation of DNA repair	0.003	1.24	3	0.126	1	1.87	6	0.0106	1
GO:0006457	protein folding	0.004	2.6	6	0.0453	1	3.88	7	0.0932	1

			Batwa:				*Andamanese:*			
	GO Molecular Function	Joint p	Exp.	Obs.	p	adj. p	Exp.	Obs.	p	adj. p
GO:0003779	actin binding	0.000	5.14	11	0.0137	0.73158	6.92	13	0.021	1
GO:0003729	mRNA binding	0.002	2.45	6	0.0354	1	3.24	6	0.105	1
GO:0051015	actin filament binding	0.002	2.15	5	0.0638	1	3.02	6	0.081	1

Table S9. After gene size-based correction, Gene Ontology (GO) biological processes with evidence of population-specific enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Results with $p < 0.01$ shown.

GO		Exp.	Obs.	p	adj. p
Batwa RHG - Biological Processes:					
GO:0007517	muscle organ development	11	4.20	0.0032	0.5650
GO:1903825	organic acid transmembrane transport	6	1.67	0.0061	0.5652
GO:0030048	actin filament-based movement	6	1.80	0.0061	0.5652
Andamanese RHG - Biological Processes:					
GO:0006302	double-strand break repair	10	3.16	0.0012	0.231
GO:0000723	telomere maintenance	8	2.31	0.0020	0.231
GO:1903827	regulation of cellular protein localization	19	9.70	0.0038	0.231
GO:0070085	glycosylation	11	4.36	0.0042	0.231
GO:1900180	regulation of protein localization to nucleus	10	3.77	0.0044	0.231
GO:0007569	cell aging	6	1.63	0.0053	0.232
GO:0060249	anatomical structure homeostasis	13	6.12	0.0082	0.299
GO:0051234	establishment of localization	99	81.24	0.0091	0.299
Batwa RHG - Molecular Functions:					
GO:0015081	sodium ion transmembrane transporter activity	7	2.31	0.0081	0.33245
Andamanese RHG - Molecular Functions:					
GO:0043130	ubiquitin binding	7	1.89	0.0026	0.177

Table S10. After MAF-based correction, Gene Ontology (GO) biological processes with evidence of population-specific enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. No molecular functions were found to be significantly shifted for the Batwa. Results with $p < 0.01$ shown.

GO		Exp.	Obs.	p	adj. p
Batwa RHG - Biological Processes:					
GO:0007517	muscle organ development	11	4.04	0.0023	0.492
GO:0051592	response to calcium ion	6	1.75	0.0079	0.492
Andamanese RHG - Biological Processes:					
GO:0051179	localization	142	114.34	0.00044	0.115
GO:0045596	negative regulation of cell differentiation	22	11.53	0.0026	0.266
GO:0071229	cellular response to acid chemical	9	3.24	0.0048	0.266
GO:0002460	adaptive immune response based on somatic recombination...	11	4.47	0.0050	0.266
GO:0014706	striated muscle tissue development	14	6.55	0.0059	0.266
GO:0048584	positive regulation of response to stimulus	53	38.05	0.0067	0.266
GO:0016337	single organismal cell-cell adhesion	22	12.61	0.0076	0.266
Andamanese RHG - Molecular Functions:					
GO:0043130	ubiquitin binding	7	2.24	0.0067	0.221
GO:0008514	organic anion transmembrane transporter activity	10	4.03	0.0068	0.221

Table S11. After subsampling to ≤ 30 individuals per population, Gene Ontology (GO) biological processes with evidence of population-specific enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. No molecular functions were found to be significantly shifted for the Andamanese. Results with $p < 0.01$ shown.

GO		Exp.	Obs.	p	adj. p
Batwa RHG - Biological Processes:					
GO:0045926	negative regulation of growth	2.66	8	0.005	0.497
GO:0035107	appendage morphogenesis	1.67	6	0.0062	0.497
GO:0030048	actin filament-based movement	1.79	6	0.0088	0.497
Batwa RHG - Molecular Functions:					
GO:0005525	GTP binding	3.33	12	0.00012	0.00828
GO:0019001	guanyl nucleotide binding	3.66	12	0.00028	0.00966
GO:0003924	GTPase activity	2.53	9	0.00092	0.0212
GO:0035639	purine ribonucleoside triphosphate binding	21.36	34	0.00369	0.0637
Andamanese RHG - Biological Processes:					
GO:0006302	double-strand break repair	3.36	10	0.0019	0.216
GO:0002460	adaptive immune response based on somatic...	3.99	11	0.0021	0.216
GO:0060249	anatomical structure homeostasis	6.69	15	0.0028	0.216
GO:0000723	telomere maintenance	2.53	8	0.0036	0.216
GO:0051053	negative regulation of DNA metabolic process	2.06	7	0.0043	0.216
GO:0045596	negative regulation of cell differentiation	10.58	20	0.0045	0.216
GO:0007569	cell aging	1.83	6	0.0095	0.346

Table S12. After subsampling Batwa and Bakiga to 10 individuals per population, Gene Ontology (GO) biological processes with evidence of population-specific enrichment for strong positive selection in the hunter-gatherer populations, as measured by outlier Population Branch Statistic (PBS) values. Results with $p < 0.01$ shown.

GO		Exp.	Obs.	p	adj. p
Batwa RHG - Biological Processes:					
GO:0006354	DNA-templated transcription, elongation	1.24	5	0.0076	0.604
GO:0000077	DNA damage checkpoint	1.75	6	0.0078	0.604
GO:0030048	actin filament-based movement	1.75	6	0.0078	0.604
GO:0016125	sterol metabolic process	1.82	6	0.0094	0.604
Batwa RHG - Molecular Functions:					
GO:1901363	heterocyclic compound binding	60.18	79	0.0016	0.106
Andamanese RHG - Biological Processes:					
GO:0006302	double-strand break repair	3.32	10	0.0017	0.230
GO:0002460	adaptive immune response based on somatic...	3.94	11	0.0019	0.230
GO:0000723	telomere maintenance	2.5	8	0.0033	0.230
GO:0045596	negative regulation of cell differentiation	10.45	20	0.0039	0.230
GO:0051053	negative regulation of DNA metabolic process	2.04	7	0.004	0.230
GO:0060249	anatomical structure homeostasis	6.61	14	0.0063	0.302
GO:0007569	cell aging	1.81	6	0.0089	0.323
Andamanese RHG - Molecular Functions:					
GO:0140096	catalytic activity, acting on a protein	35.68	50	0.007	0.380

Table S13. After gene size-based correction, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of convergent distribution shifts in PBS selection index values in the hunter-gatherer populations. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

				Batwa:		*Andamanese:*	
		GO	Joint p	p	adj. p	p	adj. p
BP	GO:0016202	regulation of striated muscle tissue development	0.000	0.0100	0.994	0.0570	1.000
	GO:1901861	regulation of muscle tissue development	0.000	0.0100	0.994	0.0570	1.000
	GO:0045444	fat cell differentiation	0.001	0.0539	0.994	0.0517	1.000
	GO:0048634	regulation of muscle organ development	0.002	0.0101	0.994	0.0924	1.000
	GO:0035265	organ growth	0.003	0.0260	0.994	0.0613	1.000
	GO:0048738	cardiac muscle tissue development	0.003	0.0646	0.994	0.0031	1.000
	GO:0051147	regulation of muscle cell differentiation	0.003	0.0242	0.994	0.1026	1.000
	GO:1903828	negative regulation of cellular protein localization	0.003	0.0475	0.994	0.0405	1.000
	GO:0046434	organophosphate catabolic process	0.004	0.0154	0.994	0.0780	1.000
MF	GO:0019199	transmembrane receptor protein kinase activity	0.000	0.021	0.698	0.0132	0.736
	GO:0019838	growth factor binding	0.002	0.029	0.750	0.0285	0.736
	GO:0030554	adenyl nucleotide binding	0.002	0.014	0.698	0.0769	0.877
	GO:0032559	adenyl ribonucleotide binding	0.002	0.017	0.698	0.0599	0.877

Table S14. After MAF-based correction, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of convergent distribution shifts in PBS selection index values in the hunter-gatherer populations. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

				Batwa:		*Andamanese:*	
		GO	Joint p	p	adj. p	p	adj. p
BP	GO:0048738	cardiac muscle tissue development	0.000	0.0921	0.993	0.00174	0.878
	GO:0033002	muscle cell proliferation	0.002	0.0919	0.993	0.02565	1.000
	GO:0035265	organ growth	0.003	0.0736	0.993	0.01943	1.000
	GO:0003007	heart morphogenesis	0.004	0.0375	0.993	0.06262	1.000
	GO:0016579	protein deubiquitination	0.004	0.1171	0.993	0.00041	0.369
MF	GO:0019199	transmembrane receptor protein kinase activity	0.001	0.026	0.675	0.039	0.874

Table S15. After subsampling to ≤ 30 individuals per population, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of convergent distribution shifts in PBS selection index values in the hunter-gatherer populations. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

				Batwa:		*Andamanese:*	
		GO	Joint p	p	adj. p	p	adj. p
BP	GO:0034655	nucleobase-containing compound catabolic...	0.002	0.0484	0.999	0.0122	0.0669
BP	GO:0051250	negative regulation of lymphocyte activation	0.002	0.0602	0.999	0.00917	0.0290
BP	GO:1903828	negative regulation of cellular protein...	0.002	0.055	0.999	0.00990	0.0390
BP	GO:0006282	regulation of DNA repair	0.003	0.0691	0.999	0.00393	0.0340
BP	GO:0046700	heterocycle catabolic process	0.003	0.0355	0.999	0.0175	0.0870
BP	GO:0044270	cellular nitrogen compound catabolic process	0.004	0.0411	0.999	0.0129	0.0789
MF	GO:0004702	signal transducer, downstream of receptor...	0.002	0.064	0.915	0.00977	0.0509
MF	GO:0008144	drug binding	0.002	0.025	0.915	0.00722	0.0579
MF	GO:0019838	growth factor binding	0.003	0.049	0.915	0.00927	0.0260

Table S16. After subsampling Batwa and Bakiga to 10 individuals per population, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of convergent distribution shifts in PBS selection index values in the hunter-gatherer populations. Joint p-values were computed via a permutation-based method, and those with joint empirical $p < 0.005$ are shown.

				Batwa:		*Andamanese:*	
		GO	Joint p	p	adj. p	p	adj. p
BP	GO:0003007	heart morphogenesis	0.002	0.00379	0.868	0.0110	0.0579
BP	GO:0016197	endosomal transport	0.002	0.0601	0.992	0.0105	0.0240
BP	GO:0035265	organ growth	0.002	0.05203	0.988	0.0131	0.0579
BP	GO:0007498	mesoderm development	0.003	0.0931	0.992	0.00951	0.0320
BP	GO:0048738	cardiac muscle tissue development	0.003	0.0428	0.977	0.00943	0.0160
MF	GO:0019838	growth factor binding	0.003	0.075	0.829	0.0131	0.0450

Table S17. After gene size-based correction, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of population-specific distribution shifts in PBS selection index values in the hunter-gatherer populations. No molecular functions were found to be significantly shifted for the Batwa. Results with $p < 0.01$ are shown.

	GO	*p*	adj. *p*
Batwa RHG - Biological Processes:			
GO:0003231	cardiac ventricle development	0.0025	0.371
GO:0061351	neural precursor cell proliferation	0.0080	0.371
Andamanese RHG - Biological Processes:			
GO:0016579	protein deubiquitination	0.001	0.273
GO:0035051	cardiocyte differentiation	0.003	0.273
GO:0048738	cardiac muscle tissue development	0.003	0.273
GO:1901800	positive regulation of proteasomal protein catabolic process	0.006	0.322
GO:0001936	regulation of endothelial cell proliferation	0.006	0.322
GO:0006508	proteolysis	0.009	0.396
Andamanese RHG - Molecular Functions:			
GO:0005085	guanyl-nucleotide exchange factor activity	0.0034	0.224
GO:0008134	transcription factor binding	0.0096	0.224

Table S18. After MAF-based correction, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of population-specific distribution shifts in PBS selection index values in the hunter-gatherer populations. No molecular functions were found to be significantly shifted for the Batwa. Results with $p < 0.01$ are shown.

GO		p	adj. p
Batwa RHG - Biological Processes:			
GO:0003231	cardiac ventricle development	0.0015	0.346
GO:0034284	response to monosaccharide	0.0043	0.346
GO:0008217	regulation of blood pressure	0.0056	0.346
GO:0050864	regulation of B cell activation	0.0083	0.346
GO:0048634	regulation of muscle organ development	0.0090	0.346
GO:1901861	regulation of muscle tissue development	0.0092	0.346
Andamanese RHG - Biological Processes:			
GO:0070646	protein modification by small protein removal	0.0003	0.087
GO:0048738	cardiac muscle tissue development	0.0017	0.160
GO:0035051	cardiocyte differentiation	0.0022	0.160
GO:0006508	proteolysis	0.0024	0.156
GO:0071840	cellular component organization or biogenesis	0.0057	0.283
GO:0007155	cell adhesion	0.0065	0.283
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	0.0091	0.298
Andamanese RHG - Molecular Functions:			
GO:0005085	guanyl-nucleotide exchange factor activity	0.0057	0.198
GO:0019783	ubiquitin-like protein-specific protease activity	0.0081	0.198

Table S19. After subsampling to ≤ 30 individuals per population, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of population-specific distribution shifts in PBS selection index values in the hunter-gatherer populations. No molecular functions were found to be significantly shifted for the Batwa. Results with $p < 0.01$ are shown.

GO		p	adj. p
Batwa RHG - Biological Processes:			
GO:0003231	cardiac ventricle development	0.0012	0.329
GO:0016125	sterol metabolic process	0.007	0.436
GO:1902652	secondary alcohol metabolic process	0.0081	0.436
GO:0061351	neural precursor cell proliferation	0.0094	0.436
Andamanese RHG - Biological Processes:			
GO:0016579	protein deubiquitination	0.002	0.342
GO:0032092	positive regulation of protein binding	0.0034	0.343
GO:0042737	drug catabolic process	0.0041	0.343
GO:1901800	positive regulation of proteasomal protein...	0.0061	0.343
GO:0007623	circadian rhythm	0.0092	0.343
GO:0035051	cardiocyte differentiation	0.0093	0.343
Andamanese RHG - Molecular Functions:			
GO:0005085	guanyl-nucleotide exchange factor activity	0.0038	0.289

Table S20. After subsampling Batwa and Bakiga to 10 individuals per population, Gene Ontology (GO) biological processes (BP) and molecular functions (MF) with evidence of population-specific distribution shifts in PBS selection index values in the hunter-gatherer populations. Results with $p < 0.01$ are shown.

	GO	p	adj. p
Batwa RHG - Biological Processes:			
GO:0030323	respiratory tube development	0.000350	0.0931
GO:0030010	establishment of cell polarity	0.00145	0.19285
GO:0003231	cardiac ventricle development	0.00267	0.214
GO:0060541	respiratory system development	0.00348	0.214
GO:2001234	negative regulation of apoptotic signaling pathway	0.00501	0.214
GO:0071772	response to BMP	0.00502	0.214
GO:0051649	establishment of localization in cell	0.00602	0.214
GO:0007389	pattern specification process	0.0066	0.214
GO:0048568	embryonic organ development	0.00724	0.214
GO:0051270	regulation of cellular component movement	0.00957	0.236
Batwa RHG - Molecular Functions:			
GO:0008092	cytoskeletal protein binding	0.009	0.363
Andamanese RHG - Biological Processes:			
GO:0016579	protein deubiquitination	0.0019	0.338
GO:0032092	positive regulation of protein binding	0.0035	0.338
GO:0042737	drug catabolic process	0.0044	0.338
GO:1901800	positive regulation of proteasomal protein...	0.0057	0.338
GO:0007623	circadian rhythm	0.0083	0.338
GO:0035051	cardiocyte differentiation	0.0097	0.338
Andamanese RHG - Molecular Functions:			
GO:0005085	guanyl-nucleotide exchange factor activity	0.0035	0.266

Table S21. Comparison of results of two methods for computing empirical test for convergence in strong outlier selection in both the Batwa and Andamanese RHGs. In the original method, genes and PBS selection index values are permuted to create an empirical null distribution. In the modified case, genes and their Gene Ontology (GO) annotations are instead permuted to create the null distribution. Biological processes (BP) with empirical test for convergence $p < 0.005$ in either method shown. No molecular functions were found to be significantly convergently enriched in both RHG populations.

	GO Biological Process	Original convergence p	Modified convergence p
GO:0035107	appendage morphogenesis	0.000	0.000
GO:0035108	limb morphogenesis	0.000	0.000
GO:0030326	embryonic limb morphogenesis	0.000	0.002
GO:0035113	embryonic appendage morphogenesis	0.000	0.002
GO:0048736	appendage development	0.001	0.003
GO:0060173	limb development	0.001	0.003
GO:0048705	skeletal system morphogenesis	0.002	0.003
GO:0048522	positive regulation of cellular process	0.006	0.003
GO:0080135	regulation of cellular response to stress	0.018	0.004
GO:0030048	actin filament-based movement	0.002	0.006
GO:0007034	vacuolar transport	0.003	-

Table S22. Comparison of results of two methods for computing empirical test for convergence in PBS selection index shift in both the Batwa and Andamanese RHGs. In the original method, genes and PBS selection index values are permuted to create an empirical null distribution. In the modified case, genes and their Gene Ontology (GO) annotations are instead permuted to create the null distribution. Biological processes (BP) and molecular functions (MF) with empirical test for convergence $p < 0.005$ in either method shown.

	GO Biological Process	Original convergence p	Modified convergence p
GO:0016202	regulation of striated muscle tissue development	0.002	0.000
GO:1901861	regulation of muscle tissue development	0.002	0.000
GO:0048738	cardiac muscle tissue development	0.001	0.001
GO:0048634	regulation of muscle organ development	0.005	0.001
GO:0045444	fat cell differentiation	0.004	0.002
GO:0003007	heart morphogenesis	0.006	0.002
GO:0035265	organ growth	0.001	0.004
GO:0046434	organophosphate catabolic process	0.007	0.004
GO:1903828	negative regulation of cellular protein localization	0.001	0.006
	GO Molecular Function	Original convergence p	Modified convergence p
GO:0019838	growth factor binding	0.000	0.001
GO:0019199	transmembrane receptor protein kinase activity	0.000	0.002
GO:0032559	adenyl ribonucleotide binding	0.003	0.004
GO:0030554	adenyl nucleotide binding	0.004	0.004
GO:0005524	ATP binding	0.005	0.004

Additional data table S1 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1a. SNPs with highest PBS values

Additional data table S2 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1b. SNPs with convergently low PBS selection index p-values between all pairs of populations

Additional data table S3 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1c. Genes that are outliers for PBS selection index

Additional data table S4 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1d. Genes with convergently low PBS selection index values between two populations

Additional data table S5 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1e. GO categories with evidence of convergent enrichment for strong positive selection (outlier-based test) in the agriculturalist populations

Additional data table S6 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1f. SNPs with highest Bayenv BF values

Additional data table S7 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1g. Genes that are outliers for Bayenv BF index value

Additional data table S8 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1h. GO categories with convergent evidence of enrichment (outlier-based test) for outlier Bayenv BF index values

Additional data table S9 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1i. Results of Fisher overrepresentation test results (outlier-based test) for *a priori* lists of growth associated genes

Additional data table S10 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1j. GO categories with population-specific enrichment for strong positive selection (outlier-based test) in the agriculturalist populations

Additional data table S11 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1k. Results of Kolmogorov-Smirnov test results (shift-based test) for *a priori* lists of growth associated genes

Additional data table S12 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1l. GO categories with convergent shift in PBS selection index value distributions (shift-based test) in the agriculturalist populations

Additional data table S13 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1m. GO categories with convergent shift in Bayenv BF index value distribution (shift-based test)

Additional data table S14 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1n. Genes in cardiac GO categories and associated PBS selection index values for each population

Additional data table S15 (Bergey_batwa_andaman_convergence.ExtendedData.xlsx)

Dataset S1o. GO categories with population-specific shifts in the PBS selection index distribution (shift-based test)

References

1. Blake JA, et al. (2017) Mouse Genome Database (MGD)-2017: Community knowledge resource for the laboratory mouse. *Nucleic Acids Research* 45(Database Issue):D723–D729.
2. Wood AR, et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* 46(11):1173–1186.
3. Lui JC, et al. (2012) Synthesizing genome-wide association studies and expression microarray reveals novel genes that act in the human growth plate to modulate height. *Human Molecular Genetics* 21(23):5193–5201.
4. Perry GH, et al. (2014) Adaptive, convergent origins of the pygmy phenotype in African rainforest hunter-gatherers. *Proceedings of the National Academy of Sciences* 111(35):E3596–E3603.
5. Brown KR, Jurisica I (2005) Online predicted human interaction database. *Bioinformatics* 21(9):2076–2082.