

SUPPLEMENTAL MATERIAL

**Analysis of data.** Because the interpretation of LD measures in terms of conventional population-genetic parameters relies on the assumption of drift-mutation-recombination equilibrium, we restricted our analyses to genomes from single diploid, non-intentionally inbred individuals. The raw reads (i.e., Sanger reads or Illumina short reads) for each genome were downloaded from the NCBI Short Read Archive. The latest repeat-masked reference assembly for each genome was downloaded from NCBI or Ensembl. The Sanger reads were trimmed based on sequence qualities using the program LUCY with default settings (Chou and Holmes 2001). The software BWA 0.7.5 (Li and Durbin 2009; Li and Durbin 2010) was used to map the Sanger raw reads against the reference assembly (bwasw command with default settings) and to align the Illumina short reads (sampe commands with default settings).

To eliminate potential problems that can arise from the mapping process, we applied three strict criteria for selecting sites that go into the final analyses. First, we removed all raw reads that mapped to multiple locations of the genome (which represent, for example, potential duplicate genes and transposons) as well as reads with a mapping quality score lower than 25. Second, sites located within paralogous genes with >90% sequence identity were excluded. Third, our analyses included only sites with coverages >4× but less than twice the genome-wide average coverage.

The software Samtools (Li et al. 2009) was used to generate a pileup of the mapped raw reads for every possible position in the genome. The pileup was then converted to a quartet profile using the software sam2pro (<http://guanine.evolbio.mpg.de/mlRho/>). For each site, the quartet profile describes the numbers of the four different nucleotide reads observed ( $n_A, n_C, n_G, n_T$ ), with their sum ( $n_A + n_C + n_G + n_T$ ) representing the coverage at each site.

The program mlRho 2.0 (Haubold et al. 2010) was used to calculate the maximum-likelihood estimates of the zygosity correlation coefficient ( $\Delta$ ), genome-wide heterozygosity ( $\theta$ ), and sequencing error rate ( $\epsilon$ ). The latter two parameters were estimated using all the available sites in the genomes and were subsequently used in the one-dimensional estimation of  $\Delta$ . We calculated  $\Delta$  for all pairs of sites separated from 1 to 100,000 bp, with an incremental increase in window size for pooled analyses. For distances up to 1000 bp,  $\Delta$  was analyzed for 1-bp distance increment (e.g.,  $\Delta$  was estimated for all pairs of sites that are 1-bp apart, then 2-bp, 3-bp, etc.), whereas for the distances between 1 and 10 kbp, these parameters were estimated with increments of 10 bp, pooling all site pairs within each 10-bp increment; and between 10 and 100 kbp, a window size of 100 bp was used.

### Literature Cited

- Chou, H. H., and M. H. Holmes. 2001. DNA sequence quality trimming and vector removal. *Bioinformatics* 17: 1093-1104.
- Haubold, B., P. Pfaffelhuber, and M. Lynch. 2010. mlRho - a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. *Mol. Ecol.* 19: 277-284.
- Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754-1760.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078-2079.
- Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589-595.

**Table S1** Summary of single diploid genomes analyzed in this study. Theta: maximum-likelihood estimate of heterozygosity in a single genome; epsilon: maximum-likelihood estimate of sequencing error rate. All raw reads were downloaded from NCBI Short Read Archive.

Species	Coverage	No. of sites ( $10^7$ )	Theta	Epsilon	Reference	DNA source
<i>Ailuropoda melanoleuca</i>	26	10.798	0.00318	0.0150	Li <i>et al.</i> (2010)	a single non-inbred female
<i>Anolis carolinensis</i>	8	106.446	0.00260	0.0034	Alfoldi <i>et al.</i> (2011)	A single female
<i>Canis familiaris</i>	10	120.495	0.00094	0.0033	Lindblad-Toh <i>et al.</i> 2005	A female Boxer
Denisovan	28	165.960	0.00042	0.0014	Meyer <i>et al.</i> 2012	DNA from phalanx of a Denisovan individual
<i>Fugu rubripes</i>	6	7.535	0.00320	0.0014	Aparicio <i>et al.</i> 2002	A single fish
<i>Gorilla gorilla</i>	32	122.182	0.00314	0.0090	Scally <i>et al.</i> 2012	A single female
<i>Homo sapiens</i> (African)	18	162.557	0.00357	0.0066	Schuster <i>et al.</i> 2010	A single male (KB1)
<i>Homo sapiens</i> (Chinese)	31	152.247	0.00164	0.0119	Wang <i>et al.</i> 2008	A single male
<i>Homo sapiens</i> (Watson)	7	147.303	0.00103	0.0004	Wheeler <i>et al.</i> 2008	A single male
<i>Loxodonta africana</i>	7	98.194	0.00134	0.0020	<a href="http://www.broadinstitute.org/">http://www.broadinstitute.org/</a>	A single individual
<i>Macaca mulatta</i>	8	81.724	0.00261	0.0033	Gibbs <i>et al.</i> 2007	A single female
<i>Ornithorhynchus anatinus</i>	10	59.061	0.00131	0.0033	Warren <i>et al.</i> 2008	A single female
<i>Pan troglodytes</i>	8	88.106	0.00101	0.0042	Chimpanzee Sequencing and Analysis Consortium 2005	A single male
<i>Petromyzon marinus</i>	12	22.753	0.00443	0.0031	Smith <i>et al.</i> 2013	A single female
<i>Pongo abelii</i>	4	27.204	0.00527	0.0087	Locke <i>et al.</i> 2011	A single female (SB550)
<i>Pongo pygmaeus</i>	6	56.747	0.00272	0.0044	Locke <i>et al.</i> 2011	A single male (KB4204)

## Literature Cited

- Chimpanzee Sequencing and Analysis Consortium, 2005 Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437: 69-87.
- Alfoldi, J., F. Di Palma, M. Grabherr, C. Williams, L. Kong *et al.*, 2011 The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* 477: 587-591.
- Aparicio, S., J. Chapman, E. Stupka, N. Putnam, J. Chia *et al.*, 2002 Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* 297: 1301-1310.
- Gibbs, R. A., J. Rogers, M. G. Katze, R. Bumgarner, G. M. Weinstock *et al.*, 2007 Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316: 222-234.
- Li, R., W. Fan, G. Tian, H. Zhu, L. He *et al.*, 2010 The sequence and de novo assembly of the giant panda genome. *Nature* 463: 311-317.
- Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe *et al.*, 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438: 803-819.
- Locke, D. P., L. W. Hillier, W. C. Warren, K. C. Worley, L. V. Nazareth *et al.*, 2011 Comparative and demographic analysis of orangutan genomes. *Nature* 469: 529-533.
- Meyer, M., M. Kircher, M.-T. Gansauge, H. Li, F. Racimo *et al.*, 2012 A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338: 222-226.
- Scally, A., J. Y. Dutheil, L. W. Hillier, G. E. Jordan, I. Goodhead *et al.*, 2012 Insights into hominid evolution from the gorilla genome sequence. *Nature* 483: 169-175.
- Schuster, S. C., W. Miller, A. Ratan, L. P. Tomsho, B. Giardine *et al.*, 2010 Complete Khoisan and Bantu genomes from southern Africa. *Nature* 463: 943-947.
- Smith, J. J., S. Kuraku, C. Holt, T. Sauka-Spengler, N. Jiang *et al.*, 2013 Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat. Genet.* 45: 415-421.
- Wang, J., W. Wang, R. Q. Li, Y. R. Li, G. Tian *et al.*, 2008 The diploid genome sequence of an Asian individual. *Nature* 456: 60-65.
- Warren, W. C., L. W. Hillier, J. a. M. Graves, E. Birney, C. P. Ponting *et al.*, 2008 Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453: 175-183.
- Wheeler, D. A., M. Srinivasan, M. Egholm, Y. Shen, L. Chen *et al.*, 2008 The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452: 872-875.