



# Phenotypic Variation in Infants, Not Adults, Reflects Genotypic Variation among Chimpanzees and Bonobos

Naoki Morimoto<sup>1\*</sup>, Marcia S. Ponce de León<sup>2</sup>, Christoph P. E. Zollikofer<sup>2\*</sup>

**1** Laboratory of Physical Anthropology, Graduate School of Science, Kyoto University, Kyoto, Japan, **2** Anthropological Institute, University of Zurich, Zurich, Switzerland

## Abstract

Studies comparing phenotypic variation with neutral genetic variation in modern humans have shown that genetic drift is a main factor of evolutionary diversification among populations. The genetic population history of our closest living relatives, the chimpanzees and bonobos, is now equally well documented, but phenotypic variation among these taxa remains relatively unexplored, and phenotype-genotype correlations are not yet documented. Also, while the adult phenotype is typically used as a reference, it remains to be investigated how phenotype-genotype correlations change during development. Here we address these questions by analyzing phenotypic evolutionary and developmental diversification in the species and subspecies of the genus *Pan*. Our analyses focus on the morphology of the femoral diaphysis, which represents a functionally constrained element of the locomotor system. Results show that during infancy phenotypic distances between taxa are largely congruent with non-coding (neutral) genotypic distances. Later during ontogeny, however, phenotypic distances deviate from genotypic distances, mainly as an effect of heterochronic shifts between taxon-specific developmental programs. Early phenotypic differences between *Pan* taxa are thus likely brought about by genetic drift while late differences reflect taxon-specific adaptations.

**Citation:** Morimoto N, Ponce de León MS, Zollikofer CPE (2014) Phenotypic Variation in Infants, Not Adults, Reflects Genotypic Variation among Chimpanzees and Bonobos. PLoS ONE 9(7): e102074. doi:10.1371/journal.pone.0102074

**Editor:** David Caramelli, University of Florence, Italy

**Received:** March 19, 2014; **Accepted:** June 13, 2014; **Published:** July 11, 2014

**Copyright:** © 2014 Morimoto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the Swiss National Science Foundation (no. 3100A0-109344/1) and Japan Society for Promotion of Science Research Fellowship for Young Scientists (no. 251133). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: morimoto@anthro.zool.kyoto-u.ac.jp (NM); zolli@aim.uzh.ch (CPEZ)

## Introduction

The ready accessibility of population-wide genotypic and phenotypic data from humans and our closest relatives, the great apes, has spurred a large number of studies investigating the relationship between patterns of genotypic and phenotypic evolution. One central issue is the relative role of neutral versus adaptive evolutionary processes in shaping genotypic and phenotypic variation. A steadily growing number of studies indicates that variation of cranial morphology among modern human populations, and between modern humans and fossil hominins (species related more closely to modern humans than to great apes) largely reflects the effects of genetic drift, while only a small proportion of variation can be attributed to selection [1,2,3,4,5,6,7,8,9,10]. Fossil hominin aDNA now also permits insights into earlier phases of human population and evolutionary history at an unprecedented level of detail [11,12,13,14,15]. These analyses are limited, however, by the “aDNA preservation horizon”, which is currently around 50,000 years BP for fossil hominin nDNA, and around 400,000 years BP for mtDNA from temperate zones [16].

One possible solution to investigate genotype-phenotype evolution beyond this horizon is to study living great ape species as a model system. The genus *Pan* represents the best model for this purpose, since it is our closest living relative, its species, subspecies and population structure is now genetically well-documented [17,18,19,20], and population history and genetic diversification

are well understood [18,19,21,22,23]. To date, two *Pan* species, *P. troglodytes* (common chimpanzee) and *P. paniscus* (bonobo) are recognized, and *P. troglodytes* is subdivided into four subspecies (*P. t. troglodytes*, *P. t. schweinfurthii*, *P. t. verus* and *P. t. ellioti*) [19]. Also, these *Pan* taxa have been the subject of detailed anatomical [24,25,26,27,28], morphological [29,30,31,32,33], phylogeographic [17,19,23,34], and behavioral [32,35,36,37,38,39,40] studies.

The extant *Pan* taxa are closely related to each other, which represents several advantages for comparative analyses. First, genotypic differences between taxa are small compared to variation within each taxon, such that the number of genes associated with phenotypic differentiation during (sub-) speciation is expected to be comparatively small [41]. Second, diversity among *Pan troglodytes* taxa represents patterns of incipient speciation, which are not yet blurred by long-term processes of taxon-specific specialization and/or convergence [42,43]. Also, we may note that the estimated time frame of *Pan* speciation [19,23] is comparable to that of our own genus *Homo* (ca. 2 million years).

Despite the increasing knowledge about *Pan* taxa, it still remains to be explored how changes at the level of the genotype are linked to changes at the level of the phenotype during speciation. The first aim of this study is thus to provide new phenotypic data documenting the evolutionary divergence of *Pan* taxa, and to relate this new evidence to the well-established body of genotypic evidence. While evolutionary studies traditionally focus on

variation in craniodental features e.g. [44,45], we study here morphological variation of the femoral shaft (= diaphysis). The femur is a functionally highly constrained element of the postcranial skeleton, and can thus be expected to be under strong stabilizing selection.

Most studies exploring genotype-phenotype relationships in great apes and humans have naturally focused on adult morphologies. This is because taxon-specific morphological features are thought to be more clearly expressed in adults than in juveniles. However, there is clear evidence that the phenotypes of early ontogenetic stages, and patterns of developmental change, are highly informative about patterns of evolutionary divergence at the levels of skeletal structure e.g. [46,47,48,49,50,51,52,53], of locomotor behaviors [35,37], and of social interactions [54]. The second aim of this study is thus to expand the scope of genotype-phenotype comparisons by taking into account the perspective of ontogeny. Here we explore how genotype-phenotype relationships change during the development of the femoral diaphysis in the different *Pan* taxa, and relate this information to evolutionary change at the level of the genotype and phenotype. Specifically, we explore when during ontogeny the effects of drift versus selection become evident in taxon-specific phenotypes.

Measuring genotype-phenotype relationships is a complex endeavor, both theoretically and practically, and requires several model assumptions. In the standard model of quantitative population genetics, phenotypic variance  $V_P$  is the combination of genetic variance  $V_G$  and environmental variance  $V_E$ :  $V_P = V_E + V_G$ . Empirical data and theoretical considerations indicate that, for complex traits, phenotypic variance can be approximated by  $V_P = V_E + V_A$ , where  $V_A$  represents additive genetic variation (the portion of phenotypic variation that can be explained by the cumulative effects of allelic variation) [55]. The question of interest here is how  $V_P$  and  $V_A$  evolve in segregating populations. In a constant environment ( $V_E = \text{const.}$ ),  $V_P = V_A$ , such that phenotypic variation reflects additive genotypic variation. Under these basic model assumptions, effects of drift and selection are typically estimated by comparing neutral genotypic distances with non-neutral distances [56,57,58,59,60]. The former distances ( $F_{ST}$ : genetic variation within subpopulation relative to total genetic variation [61,62]) are estimated from non-coding genetic markers thought to evolve under no selection such as STRs (short tandem repeats) and non-coding SNPs (single nucleotide polymorphisms) [63]. The latter distances are typically estimated from continuous quantitative genetic traits ( $Q_{ST}$ : evaluated in analogy to  $F_{ST}$  [64]) assuming additive genetic effects [64]. The question is whether  $Q_{ST}$  is equal to, smaller than, or larger than  $F_{ST}$ , which indicates neutral evolution, uniform or stabilizing selection, and diversifying selection, respectively [65].

$Q_{ST}$  can be estimated from phenotypic distance  $P_{ST}$  [66] using a measure of heritability ( $h^2$ , proportion of additive genetic variance to phenotypic variance,  $V_A/V_P$ ) [66,67,68,69,70]. In wild populations, heritability  $h^2$  is often unknown and needs to be estimated from largely comparable lab studies. Furthermore,  $h^2$  tends to change due to *in-vivo* environmental effects that accumulate during an individual's lifetime, and due to developmental changes in gene activation patterns [71,72,73]. In any case, estimates of  $h^2$  affect the distance measures expressed by  $Q_{ST}$ , such that estimating the relative contribution of additive genetic and *in-vivo* environmental effects to  $P_{ST}$  remains a challenge [74].

A further challenge of  $F_{ST} - Q_{ST}$  comparisons is the practical difficulty in measuring genotypic and phenotypic distances. Genotypic distances have been typically calculated using population-specific allele frequencies [75] (e.g., in Nei's standard distance  $D_A$  [76] and Cavalli-Sforza and Edwards chord distance  $D_{CH}$

[77]). One problem is that sample sizes of wild populations are often limited, which makes it difficult to estimate population-specific allele frequencies and within-population variation. Complementary methods have thus been proposed, e.g. Principal Components Analysis (PCA) of genetic data [78,79]. While phenotypic distances have traditionally been evaluated from arrays of linear and angular measurements, geometric morphometrics (GM) offers elegant methods to quantify complex patterns of phenotypic variation [80,81,82]. In GM, biological form is typically measured by the spatial configuration (3D geometry) of anatomical points of reference, so-called landmarks [83,84]. Alternatively, various methods of GM have been developed to quantify the shape of landmark-free biological structures such as outlines [85], endocranial cavities [86] and longbone shafts [46,87]. One key feature of all GM methods is that phenotypic variation can simultaneously be represented in physical (three-dimensional) space by means of graphical interpolation and in multivariate space by means of PCA. PCA thus provides an ideal means to compare multivariate genotypic and phenotypic data independent of underlying population models.

## Materials and Methods

Volumetric data of the femora of  $N=146$  *Pan* specimens were acquired with computed tomography (CT) ( $N=50$  *Pan troglodytes troglodytes*,  $N=39$  *P.t. schweinfurthii*,  $N=26$  *P. t. verus*,  $N=31$  *P. paniscus*; see Figs. S1 and S2, Table S1, and Text S1 and S2 for details on sample structure). *P. t. troglodytes* and *P. t. verus* specimens were obtained from the collections of the Anthropological Institute and Museum of the University of Zurich (AIMUZH), *P. t. schweinfurthii* specimens were obtained from the collections of the Royal Africa Museum, Tervuren, Belgium (MRA), and *P. paniscus* specimens were obtained from AIMUZH and MRA (Table S1). Each taxon is represented by four consecutive ontogenetic stages from infancy to adulthood. These were defined according to dental eruption: m2 (second deciduous molar erupted), M1, M2, M3 (first, second, third permanent molars erupted). In *Pan*, m2, M1, M2 and M3 erupt approximately at 0.5–0.83, 3, 7 and 11 years after birth, respectively [88].

Because femoral epiphyses are not yet ossified during the early stages of ontogeny, we focus on diaphyseal morphology. Effects of *in-vivo* bone modification in the femur have been studied in various *Pan* taxa, and it has been shown that ontogenetic changes in femoral morphology reflect an underlying developmental program that is fairly independent of environmental influences [87]. In other words, environmental variance  $V_E$  remains approximately constant throughout ontogeny [31,87,89] (see Text S3), which is an important prerequisite to estimate  $Q_{ST}$  from  $P_{ST}$  [74].

To quantify a specimen's diaphyseal surface morphology the transverse radius of curvature was evaluated for each point of the external (subperiosteal) surface, as specified in ref. [87]. The data of all specimens were then analyzed by means of morphometric mapping (MM) methods [87,90] (Fig. S3 and Text S1). MM is a landmark-free geometric morphometric method that permits dense sampling of data from smooth surfaces. It is thus well suited to quantify even subtle morphological differences in femoral shaft form between different taxa and/or developmental stages [87,91,92,93]. To correct for size differences between specimens, size is normalized by diaphyseal length and the median value of the radius of curvature. Shape variation is then decomposed into statistically independent shape components, which represent multivariate descriptors of the total femoral diaphyseal morphology. Since MM establishes a direct link between femoral geometry and its multivariate representation, patterns of inter- and intra-

group variation can be visualized in multivariate shape space (“morphospace”; Fig. 1) as well as in real (physical) space (Fig. 2). To infer the femoral diaphyseal morphology and its developmental pattern in the last common ancestor (LCA) of *Pan* taxa, the phylogenetic tree of *Pan* taxa was projected onto the morphospace using a model of squared-change parsimony under a Brownian motion model [94] for each ontogenetic stage (Fig. S4) using the software package MorphoJ [95]. Also, MM was used to infer the infant and adult femoral diaphyseal morphology of the LCA (Fig. 2).

Mean femoral diaphyseal shape was calculated for each taxon at each ontogenetic stage  $i$ , and inter-taxon phenotypic (*i.e.*, morphometric) distance matrices  $\mathbf{M}_i$  were calculated for each stage. As a phenotypic distance metric, the Euclidean distance in morphospace was used. Between-taxon quantitative genetic differentiation ( $Q_{ST}$ ) was also estimated for each ontogenetic stage. To this end, pairwise  $Q_{ST}$ s were evaluated from  $P_{ST}$ s with the software RMET 5.0 [96,97], using PC scores (PC1–3) and a standard estimation of heritability  $h^2 = 0.55$ . This procedure resulted in stage-specific distance matrices  $\mathbf{Q}$ .

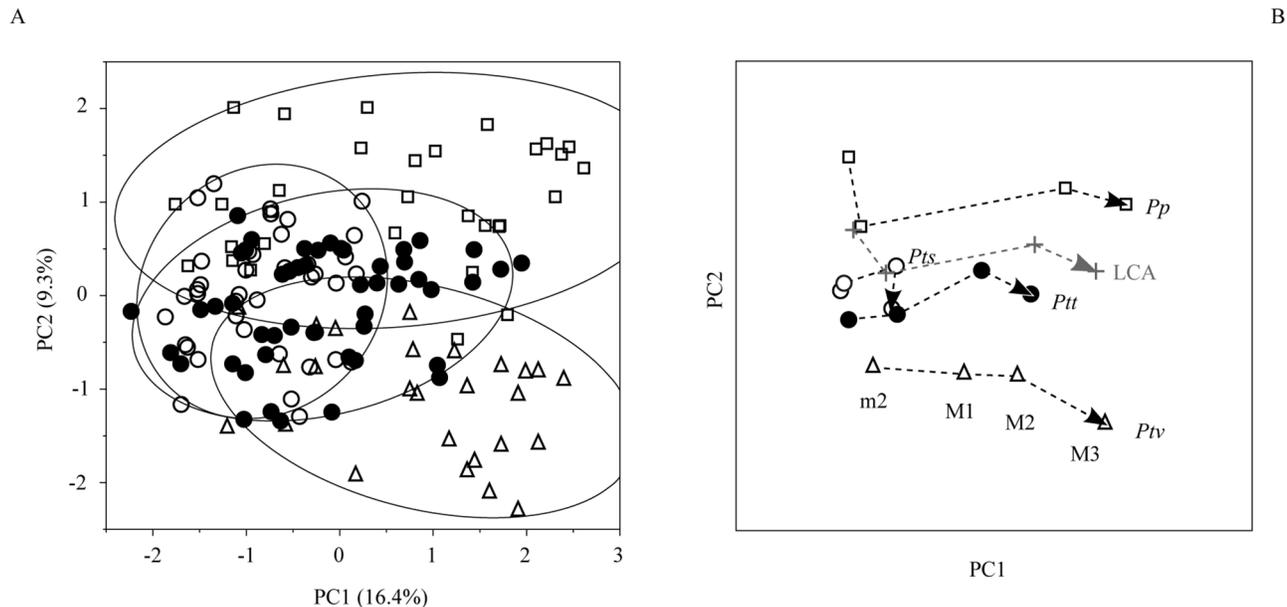
Genotypic distances between *Pan* taxa (matrices  $\mathbf{F}$ ) were calculated from sequence datasets. The sequence data of 150,000 bp on 15 non-coding autosomal regions in  $N = 74$  *Pan* specimens were obtained from GenBank (accession number: JF725992–727161 [22]). Inter-taxon genotypic distances were evaluated with various methods; Nei’s standard distance  $D_a$  [76], Cavalli-Sforza and Edwards chord distance  $D_{CH}$  [77], and Euclidean distances in Patterson’s PC space  $D_{PPC}$  [78,79]. Further,  $F_{ST}$  and  $R_{ST}$  from published sources were also used to construct genotypic distance matrices ([18,19,21,22]; refs. [18] and [19] use the same marker set) (Table S2).

Overall, three kinds of between-taxon distance matrices  $\mathbf{F}$  (genotypic),  $\mathbf{M}$  (phenotypic) and  $\mathbf{Q}$  (quantitative genetic) were evaluated, and these matrices were used for  $\mathbf{F}-\mathbf{M}$  and  $\mathbf{F}-\mathbf{Q}$  ( $F_{ST} - Q_{ST}$  [ $P_{ST}$ ]) comparisons. The similarity between these distance

matrices was evaluated with principal coordinate analysis (PCO), and assessed statistically with the Mantel test and resampling statistics (see Text S1 and Fig. S3 for details on PCO and resampling statistics). In brief, PCO transforms a between-taxon distance matrix into a “taxon constellation” (*i.e.*, locations of taxa relative to each other in multivariate space). To assess the coincidence between genotypic and phenotypic taxon constellations, we used Procrustes analysis. This method superimposes two or more different constellations using a least-squares criterion. The Mantel test was performed using Relethford’s MANTEL 3.1 (software programs RMET and MANTEL are available at <http://employees.oneonta.edu/relethjh/programs/>).

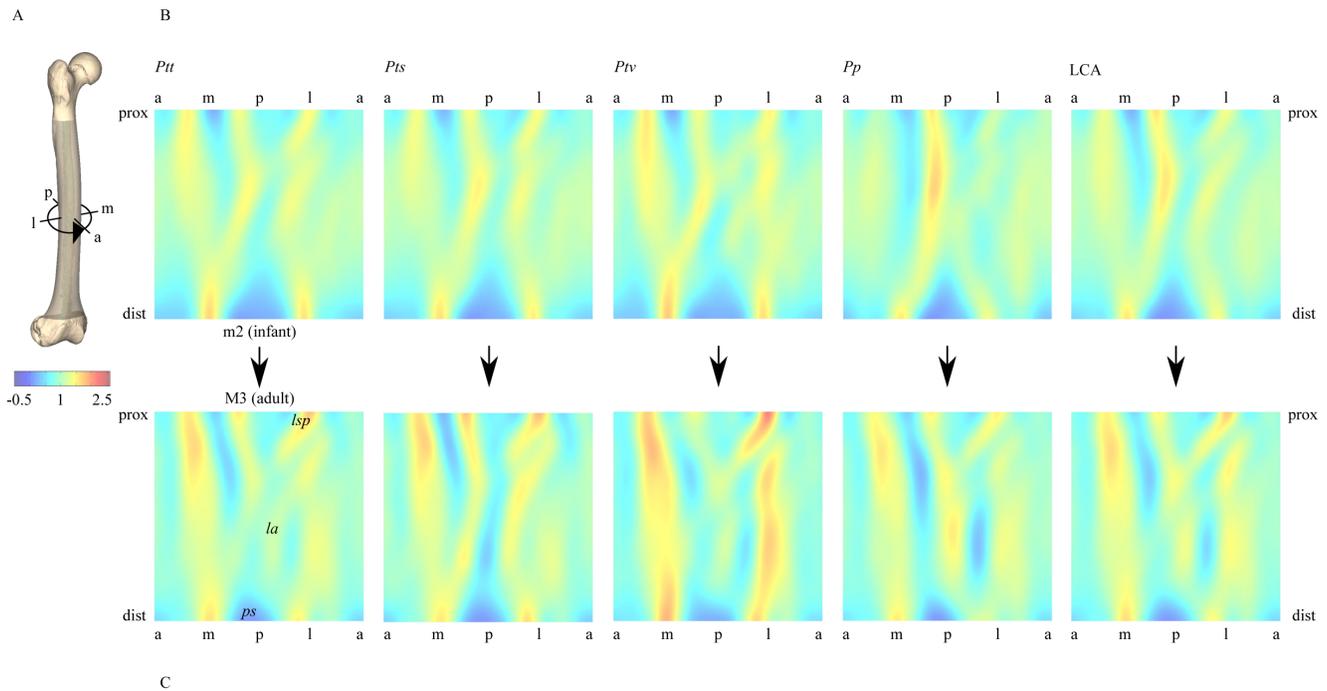
The fact that more than two *Pan* taxa are studied here facilitates rather than complicates  $F_{ST}-Q_{ST}$  comparisons. For  $K = 2$  groups (populations or taxa), one  $F_{ST}$  distance is compared with one  $Q_{ST}$  distance. These need to be scaled appropriately with an estimate of  $h^2$  to permit significant implications on neutral versus adaptive evolution, but  $h^2$  is typically unknown. For  $K > 2$  groups (this study:  $K = 4$ ), the structures of two  $K \times K$  distance matrices ( $F$  and  $Q$ ) are compared, and scaling issues can be addressed with methods of matrix-matrix correlation and multidimensional scaling (MDS) such as the PCO method used here *e.g.* [2,7,98,99,100]. Assuming that  $h^2(i) = \text{const.}$  for all groups at a given ontogenetic stage  $i$ , MDS will thus scale  $P_{ST}$  and  $Q_{ST}$  relative to  $F_{ST}$  even without explicit estimates of  $h^2(i)$  (refs. [10,101]).

These matrix-matrix comparisons permit to assess whether the structure of a phenotypic ( $\mathbf{M}$ ) or quantitative-genetic ( $\mathbf{Q}$ ) distance matrix is similar to, or deviates from, a putatively neutral genotypic distance matrix  $\mathbf{F}$ . Similarity would imply that  $\mathbf{M}$  and  $\mathbf{Q}$  are scaled versions of  $\mathbf{F}$  (scaling factor  $h^2$ ). An important assumption is that the genetic markers to estimate  $F_{ST}$  follow neutral evolution. This is critical to evaluate the relative role of neutral and adaptive processes from phenotypic data. The genetic markers used here to estimate  $F_{ST}$  represent non-coding regions



**Figure 1. Femoral diaphyseal shape variation in an ontogenetic sample of *Pan* taxa.** A: Variation along the first two principal components of shape, PC1 and PC2 (filled circles: *P.t. troglodytes*, open circles: *P.t. schweinfurthii*, open triangles: *P.t. verus*, open squares: *P. paniscus*). Solid outlines show 95%-density ellipses for each taxon. B: plot of mean shapes at consecutive ontogenetic stages. m2: second deciduous molar erupted; M1/M2/M3: permanent molars 1/2/3 erupted. Gray symbols and dashed line indicate the inferred shape at each ontogenetic stage and ontogenetic trajectory of the last common ancestor.

doi:10.1371/journal.pone.0102074.g001



**Figure 2. Taxon-specific femoral diaphyseal shapes.** A: principle of morphometric map generation (anterior [0°] → medial [90°] → posterior [180°] → lateral [270°] → anterior [360°]). B, C: morphometric maps of taxon-specific morphologies at ontogenetic stages m2 (B, infant) and M3 (C, adult) (false-color images of external surface curvature [relative units]). *la*: linea aspera, *lsp*: lateral spiral pilaster, *ps*: popliteal surface. doi:10.1371/journal.pone.0102074.g002

[18,19,21,22], so it is reasonable to assume that variation reflects neutral processes.

## Results

Fig. 1 shows commonalities and differences in femoral diaphyseal shape and shape variation between *Pan* taxa. The first two principal components represented here (PC1 and PC2) account for 25.7% of the total shape variation in the sample. There is substantial overlap between taxon-specific distributions of *P. t. troglodytes* and *P. t. schweinfurthii*, but almost no overlap between *P. paniscus* and *P. t. verus* (Fig. 1A). At each ontogenetic stage, taxon-specific mean shapes are statistically different from each other (Fig. 1B, Table S3). Furthermore, taxon-specific ontogenetic trajectories (see SI and refs. [102,103]) have statistically similar directions through morphospace (Fig. 1B and Table S4). Trajectories differ from each other, however, in their length (mostly along PC1), and in their location in morphospace (mostly along PC2) (Fig. 1B). Trajectories of *P. t. troglodytes* and *P. t. schweinfurthii* are in close vicinity, but the trajectory of the latter taxon is significantly shorter than that of the former. Compared to these taxa, the trajectory of *P. paniscus* is significantly longer (Fig. 1B, Table S5).

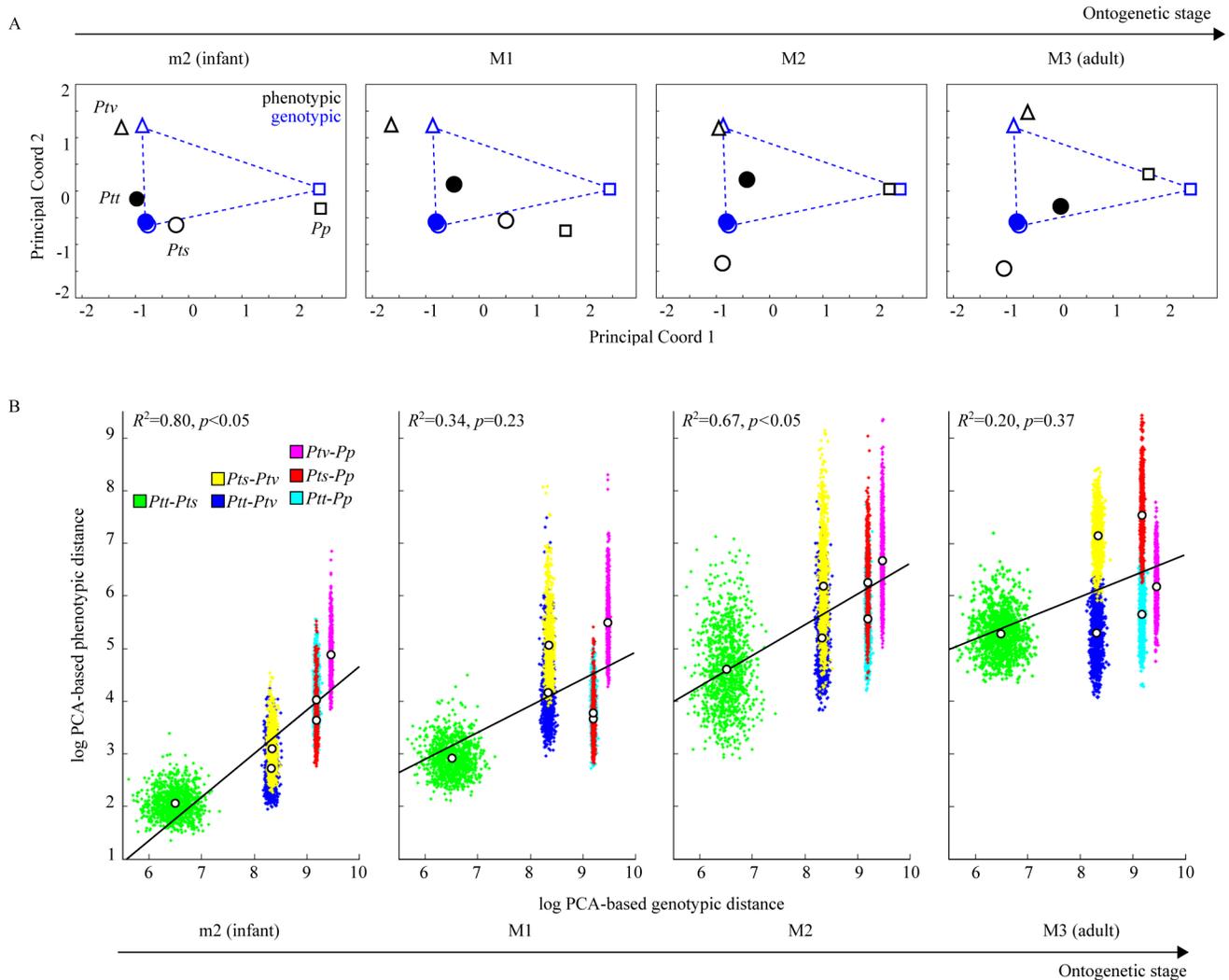
Differences between trajectories are already present at the m2 (infant) stage, indicating that taxon-specific femoral shape is established early during ontogeny. The differences in trajectory length indicate that the shape differences between *Pan* taxa increase toward adulthood. Longer trajectories indicate a larger total amount of femoral shape change during ontogeny, and possibly higher rates of shape change. Fig. 2 visualizes the corresponding real-space patterns of femoral diaphyseal shape change from infant to adult for each taxon. Each stage- and taxon-specific diaphyseal shape is represented here with a morphometric map (MM), which represents surface structures around (x-axis) and

along (y-axis) the femoral diaphysis. MMs visually confirm that taxon-specific femoral shape is present already at the m2 (infant) stage, and that taxon-specific features become more pronounced toward the M3 (adult) stage.

Using methods of squared-change parsimony [94], it is possible to infer the ontogenetic trajectory of the LCA of *Pan* taxa. The LCA trajectory lies between the trajectory of *P. paniscus* and the average trajectory of *P. troglodytes* taxa (Figs. 1, 2, S4). The length of the LCA trajectory is comparable to that of *P. t. troglodytes*, *P. t. verus*, and *P. paniscus*, but is longer than that of *P. t. schweinfurthii*.

All measures of genotypic distances ( $F_{ST}$ ,  $D_a$ ,  $D_{CH}$ ,  $D_{PPC}$ ) are highly correlated with each other (Table S6; Mantel test). Genotypic distances ( $F_{ST}$  and  $R_{ST}$ ) evaluated from different marker sets [18,19,21,22] (Table S2) are also concordant with each other (Fig. S5), indicating that potential noise due to the small sample sizes of these studies does not greatly affect the results [104]. In all further comparative analyses we use  $D_{PPC}$  because evaluation of this distance measure does not presuppose estimation of within-group variance.

To assess the congruence between genotypic and phenotypic distance matrices, we projected the genotypic and phenotypic PCO data into the same multidimensional space and aligned them with Procrustes Analysis. Patterns of phenotypic similarity among *Pan* taxa ( $P_{ST}$ ) are overall congruent with patterns of genetic similarity ( $D_{PPC}$ ,  $F_{ST}$ ) (Figs. 3A, S5, Tables 1, S6, S7). Figs. 3A and S5 show that the match between genotypic and phenotypic data is closest at the m2 (infant) stage (Table 1;  $p < 0.05$ , Mantel test). While taxa advance along their ontogenetic trajectories, patterns of phenotypic variation tend to deviate from the pattern of genetic variation (Fig. 3A, S5). These results are statistically supported by a resampling test (Fig. 3B). **F-M** correlation is highest at the m2 (infant) stage ( $R^2 = 0.80$ ,  $p = 0.02$ ), and is lowest at the M3 (adult) stage ( $R^2 = 0.20$ ,  $p = 0.37$ ). Likewise, the **F-M** correlation between genotypic and phenotypic distances evaluated by a Mantel test is



**Figure 3. Comparison of genotypic and phenotypic distances between *Pan* taxa.** A: Principal Coordinates Analysis (PCO) permits representation of genotypic and phenotypic distance data in the same multivariate space. The four subgraphs show phenotypic data (black dots) for consecutive ontogenetic stages m2, M1, M2, and M3, and genotypic data (same blue dots for all stages). For graphical clarity genotypic data points, which are independent of ontogenetic stage, are connected with dashed lines. Note that during ontogeny the phenotypic distance configuration departs from the neutral genetic distance configuration (see also Fig. S5). B: Correlation between phenotypic and neutral genetic distances between taxa. Each point cloud consists of 1000 randomly sampled phenotypic and genotypic distances between individuals belonging to different *Pan* taxa (resampling procedures are explained in Text S1). Correlation of phenotypic and neutral genetic distances is highest at the m2 (infant) stage and declines towards adulthood (M3). Genetic and phenotypic distances are normalized by their respective median values. Note overall increase of phenotypic distance between taxa toward adulthood. doi:10.1371/journal.pone.0102074.g003

highest at the m2 stage (Tables 1 and S7). **F–M** correlation is also significant at the M2 stage, but to a lesser extent than at the m2 stage. The decline in **F–M** correlation from infancy to adulthood thus follows a non-monotonous pattern.

The results of **F–Q** comparisons (i.e., standard  $F_{ST}-Q_{ST}$  tests) are similar to the results obtained with PCA/PCO analyses (Table 1). The correlation between  $F_{ST}$  and  $Q_{ST}$  [ $P_{ST}$ ] is highest at the m2 (infant) stage ( $R^2=0.72, p<0.01$ ), and lowest at the M3 (adult) stage ( $R^2=0.10, p=0.35$ ). The finding that correlation between genotypic and phenotypic markers decreases during ontogeny is thus independent of the method of genotypic and phenotypic distance measurement.

## Discussion

Investigating the evolutionary divergence between populations and/or closely related taxa at the level of genes and phenes, and inferring underlying processes of selection and drift, has become an important research topic in primatology and anthropology [1,2,3,4,5,6,7,8,9,10]. Progress in this field is fostered by the availability of ever-increasing volumes of genomic and phenomic data, and sophisticated analytical tools to compare patterns of genotypic and phenotypic variation. While DNA sequence data provide *static* structural information about the genome, data at any level above the DNA (from the transcriptome to morphology) provide *dynamic* structural information about the phenotype, which changes during ontogeny. Interestingly, the effect of ontogenetic time on correlations between genotypic and phenotypic variation is still relatively unexplored. For example, ontogenetic time does

**Table 1.** Correlation between genotypic and phenotypic distance matrices.

		m2 (infant)	M1	M2	M3 (adult)
genotypic distance <sup>1</sup> –phenotypic distance <sup>2</sup> (Mantel <sup>3</sup> )	$R^2$	<b>0.84</b>	0.15	<b>0.64</b>	0.18
	$p$	<b>&lt;0.01</b>	0.2609	<b>&lt;0.01</b>	0.3478
genotypic distance–phenotypic distance (resampling <sup>4</sup> )	$R^2$	<b>0.80</b>	0.34	<b>0.67</b>	0.20
	$p$	<b>0.015</b>	0.23	<b>0.045</b>	0.37
$F_{ST}$ – $Q_{ST}$ test <sup>5</sup> (Mantel)	$R^2$	<b>0.72</b>	0.40	<b>0.67</b>	0.10
	$p$	<b>&lt;0.01</b>	0.087	<b>&lt;0.01</b>	0.3478

<sup>1</sup>Euclidean distance in Patterson's PC space.

<sup>2</sup>Euclidean distance in morphospace (shape PCs).

<sup>3</sup>correlation ( $R^2$ ) and significance levels ( $p$ ) evaluated with Mantel test (1000 permutations).

<sup>4</sup>evaluated with resampling statistics (see methods; Fig. S3C).

<sup>5</sup>estimate of heritability  $h^2$ : 0.55.

doi:10.1371/journal.pone.0102074.t001

not appear as an explicit variable in the standard equations relating  $V_P$  to  $V_A$ , nor is it typically considered explicitly in  $F_{ST}$ – $Q_{ST}$  comparisons.

To fill this gap, we studied femoral diaphyseal shape change in the genus *Pan* and compared patterns of phenotypic divergence (both during development and evolution) with patterns of genotypic divergence. The results presented here yield several new insights into evolutionary and developmental links between genotypic and phenotypic diversification in *Pan*. Before any general inferences can be drawn, it should be reminded, however, that the genotypic and phenotypic data sets studied here represent subsets of the total genotypic/phenotypic evidence that is potentially available for such studies.

The close correspondence between genotypic and phenotypic distances at the earliest ontogenetic stage analyzed here (the m2 stage) gives rise to two alternative hypotheses; H0: if the molecular markers of refs. [18,19,21,22] track neutral evolution then the observed pattern of phenotypic evolution is “neutral-like” within the constraints imposed by stabilizing selection (often described as “wandering around an adaptive optimum” [105,106,107]); H1: if the pattern of phenotypic distances between taxa is the result of selection and adaptation, then the molecular markers are non-neutral and carry an adaptive signal. Given the good evidence for neutrality in the molecular markers [108] used here, hypothesis H1 is less likely. Also, the congruence of the genotypic distance patterns evaluated from different marker types (Fig. S5) suggests that H1 is less likely, since one would expect that selection acts differently on different marker types. Our data thus support hypothesis H0, which implies that morphological variation of the femoral diaphysis in infant *Pan* reflects neutral evolutionary diversification between taxa rather than taxon-specific adaptation.

While phenotypic distances between *Pan* taxa at the m2 stage are in good concordance with genotypic distances ( $R^2=0.8$ ; Fig. 3B), correlations are lower at later ontogenetic stages, and reach a value of  $R^2=0.2$  at adulthood (Figs. 3, S5; Table 1). As already reported in earlier studies [74,109,110], correlations between molecular and phenotypic markers are typically low, and this has been interpreted in two ways: (1) that (non-coding) molecular marker variation does not adequately represent the quantitative genetic variation of coding genes that becomes manifest in the phenotype, and (2) that environmental variation has a significant influence on  $V_P$ , and hence on  $Q_{ST}$ .

The ontogenetic data presented in this study provide an empirical basis to test these hypotheses. The high correlation ( $R^2=0.80$ ) between inter-taxon molecular and phenotypic varia-

tion at the m2 stage (Fig. 3B) indicates that, during early ontogeny, molecular marker variation indeed represents quantitative genetic variation. Departure from genotypic-phenotypic correspondence during later ontogenetic stages might indicate *in-vivo* modification of the femoral shaft morphology, indicating an increasing contribution of  $V_E$  to  $V_P$  over ontogenetic time. Given the evidence from earlier studies investigating *in-vivo* effects on femoral shaft morphology [31,87,89,111], however, this interpretation is unlikely, and  $V_E$  remains fairly constant from infancy to adulthood [87]. Another possible explanation is size allometry, implying that the observed pattern of phenotypic divergence reflects differences in adult body mass among *Pan* taxa. Since direct data on body mass are available for only few specimens in this study, we use the taxon-specific body masses reported in the literature [112] to test this hypothesis. Taxon-specific means of PC scores at adulthood are not correlated with adult body masses of *Pan* taxa (Fig. S6, Table S8). It is thus unlikely that the observed pattern of divergence is due to allometry.

After excluding major environmental and allometric effects, it appears most likely that phenotypic divergence is caused by genetically determined taxon-specific developmental programs. This implies that the genetic variance  $V_G$  changes during ontogenetic time  $t$ :  $V_P(t) = V_E + V_G(t)$ . In the present case, it is not known whether  $V_G(t)$  can be approximated by additive genetic variance  $V_A(t)$  alone, or whether non-additive effects have to be taken into account. Several alternative hypotheses must thus be considered to explain the observed pattern of phenotypic divergence. Under the additive genetic variance model [ $V_P(t) = V_E + V_A(t)$ ], our hypothesis is that the genes mediating early ontogeny (up to the m2 stage) evolved by neutral processes ( $Q_{ST} \sim F_{ST}$ ), whereas the genes mediating late ontogeny (from m2 to adulthood) evolved under selection ( $Q_{ST} > F_{ST}$ ), probably as an adaptation to taxon-specific locomotor regimes. An alternative hypothesis is that non-additive effects  $V_N$  are a function of developmental time:  $V_P(t) = V_E + V_A(t) + V_N(t)$ . With the currently available empirical evidence, we cannot decide between these hypotheses. In any case, the molecular markers used here to estimate  $V_G$  are unlikely to represent variation in the actual coding genes that cause  $V_P$  to increase over ontogenetic time [110].

In spite of these uncertainties, our data permit inferences on the developmental mechanisms that cause taxon-specific differences in femoral diaphyseal shape, and to speculate on their genetic basis. As shown in Fig. 1B, taxon-specific ontogenetic trajectories set out at similar locations along PC1, but differ in their length. This pattern indicates differences in taxon-specific *rates* of development

from the m2 stage onward, resulting in significant differences between adult morphologies. Evolutionary divergence via differential developmental rates is well-known as heterochrony. It thus appears that heterochronic shifts played a major role in the development of the adult femoral morphologies of *Pan* taxa. Such shifts might be effected by changes in a small number of developmental genes [113,114], which are difficult to trace with standard molecular markers, but might be further investigated with whole-genome comparisons [23].

It has been shown that a marked paedomorphic pattern is expressed in the skull relative to the postcranial skeleton in bonobos (*P. paniscus*) compared to common chimpanzees (*P. troglodytes*) [33,115,116]. The present study shows that the femur also exhibits heterochronic variation among *Pan* taxa. It is interesting to note that the femoral diaphysis of bonobos exhibits peramorphic development compared to common chimpanzees. This mosaic structure of evolutionary developmental modification is in concordance with the observation made earlier that *P. paniscus* is not just a paedomorphic chimpanzee [116,117]. It remains to be elucidated whether cranial and postcranial ontogenies are governed by the same set of “heterochrony genes”, which have different local effects, or whether different sets of heterochrony genes are expressed locally [113,118].

Currently, we can only speculate about the adaptive significance of taxon-specific heterochronic modifications of femoral development, since more comparative field data are necessary to specify the diversity of locomotor behaviors and their ontogeny in all *Pan* taxa. The inferred femoral diaphyseal morphology and developmental trajectory of the *Pan* LCA indicates that the peramorphic pattern as in *P. paniscus*, *P. t. troglodytes* and *P. t. verus* represents the primitive state whereas the paedomorphic (rate hypomorphic) pattern as in *P. t. schweinfurthii* represents the derived state. The inferred femoral diaphyseal morphology of the LCA at the adult stage is relatively close to the morphology of adult *P. paniscus* and *P. t. troglodytes*. The locomotor repertoire of the LCA might thus have been close to that of adult *P. paniscus* and *P. t. troglodytes*.

The data presented here provide empirical insights into the role of neutral and adaptive evolutionary mechanisms at the level of genes and phenes. In the system studied here, it appears that – among the closely related *Pan* taxa – early developmental genes evolve mostly neutrally and produce neutral taxon-specific phenotypes, while selection acts on late developmental genes (most likely on those involved in the regulation of developmental rates) and produces adaptive phenotypes.

Evidence for this pattern of evolution has also been found in the hominin clade. For example, the pattern of genotypic and phenotypic divergence between *Homo sapiens* and *H. neanderthalensis* is concordant with a model of neutral evolution by mutation and drift [6,8]. Also, parallel ontogenetic trajectories and heterochronic divergence during late ontogeny are reported for *Homo sapiens* and *H. neanderthalensis* [51]. Likewise, it appears that genetic and phenotypic divergence in early *Homo* and between modern human populations is governed to a large extent by neutral processes [1,3,5,10,119,120]. Our data indicate that this pattern of evolution might be more general than currently thought and characteristic not only for *Homo* but also for the taxa descending from the last common ancestor of humans and chimpanzees. It remains to be tested whether the observed patterns of developmental diversification in *Pan* also characterize the developmental diversification in other great ape taxa.

As a general outcome of this study, we may state that the phenotype of early developmental stages conveys a better neutral phylogenetic signal than the adult phenotype. This finding is in contrast with the traditional notion that the fully-developed adult

phenotype is most significant for taxonomy and phyletic inference. The close match between patterns of neutral molecular and phenotypic variation during early ontogeny, however, indicates that immature individuals are of special relevance to infer phylogenetic relationships, although taxon-specific features are less expressed in early stages of ontogeny (Fig. 2B) compared to late stages (Fig. 2C). Femoral diaphyseal morphology of hominoids provides a good example. While adult-based studies often show similarities of femoral diaphyseal morphology among great apes to the exclusion of humans e.g. [121,122,123], at an early developmental stage humans and chimpanzees are grouped together to the exclusion of gorillas [46]. Furthermore, our data may explain why previous meta-analyses showed a generally low correlation of  $F_{ST}$  and  $Q_{ST}$  in adult phenotypes [74,110,124]. Generalizing our findings to hominoid (and hominin) evolution, the comparison of immature and adult phenotypes will permit a better discrimination between phyletic and adaptive signals in the phenotype.

## Supporting Information

### Figure S1 Geographical distribution and taxonomy of *Pan* (modified from ref. [22]).

(TIF)

### Figure S2 Sample structure by taxon and age class.

A, distribution of femoral diaphyseal length (measured as the linear distance between proximal and distal epiphyseal lines). B: distribution of femoral diaphyseal cross-sectional area (measured as the median of cross-sectional areas between proximal and distal epiphyses). Filled circles: *P.t. troglodytes*, open circles: *P.t. schweinfurthii*, open triangles: *P.t. verus*, open squares: *P. paniscus*. Age classes: m2: second deciduous molar erupted; M1/M2/M3: permanent molars 1/2/3 erupted. Each symbol represents a specimen; black lines/whiskers indicate mean and range; red boxes and whiskers indicate first/third quartiles and median.

(TIF)

### Figure S3 Principle of morphometric mapping.

A, 3D representation of the right femur. B, principle of cylindrical projection (anterior [0°] → medial [90°] → posterior [180°] → lateral [270°] → anterior [0°]).

(TIF)

### Figure S4 Phylogenetic tree in morphospace.

The phylogenetic tree (blue lines; diamonds indicate the inferred state of last common ancestor at each ontogenetic stage) of the genus *Pan* is projected onto the shape space using a model of squared-change parsimony. A: m2 (infant), B: M1, C: M2, D: M3 (adult) stage. Gray symbols and line indicate the inferred ontogenetic trajectory of the last common ancestor.

(TIF)

### Figure S5 Phenetic and genetic similarity between *Pan* taxa.

Principal Coordinates Analysis (PCO) of phenetic and genetic distance data. Phenetic data (black) are given for consecutive ontogenetic stages (connected with dashed lines). Genetic data (color) are from ref. [18] (blue), ref. [19] (green), ref. [21] (red), and ref. [22] (magenta). Note that during ontogeny the phenetic distance configuration departs from the genetic distance configuration.

(TIF)

### Figure S6 Correlation of taxon-specific means of adult body weight and PC scores.

Taxon-specific means of adult body weight was calculated as a mean of male and female body weight taken from the literature [112].

(TIF)

**Table S1 Specimen list.** The following specimens are used in this study. AIMUZH: Anthropological Institute and Museum of University of Zurich. MRA: Royal Africa Museum, Tervuren, Belgium.  
(DOCX)

**Table S2 Genetic distances between *Pan* taxa ( $F_{ST}$  and  $R_{ST}$ ).**  
(DOCX)

**Table S3 Phenetic distances between taxon-specific mean shapes.**  
(DOCX)

**Table S4 Divergence of ontogenetic vector.**  
(DOCX)

**Table S5 F-test on taxon-specific variance along PC1.**  
(DOCX)

**Table S6 Correlation of genetic and phenetic distances.**  
(DOCX)

**Table S7 Correlation between phenetic and genetic distance matrices.**  
(DOCX)

## References

- Ackermann RR, Cheverud JM (2004) Detecting genetic drift versus selection in human evolution. *Proc Natl Acad Sci U S A* 101: 17946–17951.
- Roseman CC (2004) Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. *Proceedings of the National Academy of Sciences of the United States of America* 101: 12824–12829.
- Roseman CC, Weaver TD (2004) Multivariate apportionment of global human craniometric diversity. *American Journal of Physical Anthropology* 125: 257–263.
- Harvati K, Weaver TD (2006) Human cranial anatomy and the differential preservation of population history and climate signatures. *Anat Rec A Discov Mol Cell Evol Biol* 288: 1225–1233.
- Roseman CC, Weaver TD (2007) Molecules versus morphology? Not for the human cranium. *Bioessays* 29: 1185–1188.
- Weaver TD, Roseman CC, Stringer CB (2007) Were neandertal and modern human cranial differences produced by natural selection or genetic drift? *Journal of Human Evolution* 53: 135–145.
- Smith HF, Terhune CE, Lockwood CA (2007) Genetic, geographic, and environmental correlates of human temporal bone variation. *American Journal of Physical Anthropology* 134: 312–322.
- Weaver TD, Roseman CC, Stringer CB (2008) Close correspondence between quantitative- and molecular-genetic divergence times for Neandertals and modern humans. *Proc Natl Acad Sci U S A* 105: 4645–4649.
- von Cramon-Taubadel N, Weaver TD (2009) Insights from a quantitative genetic approach to human morphological evolution. *Evolutionary Anthropology* 18: 237–240.
- Betti L, Balloux F, Hanihara T, Manica A (2010) The relative role of drift and selection in shaping the human skull. *American Journal of Physical Anthropology* 141: 76–82.
- Reich D, Green RE, Kircher M, Krause J, Patterson N, et al. (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468: 1053–1060.
- Krause J, Fu Q, Good JM, Viola B, Shunkov MV, et al. (2010) The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature* 464: 894–897.
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, et al. (2010) A draft sequence of the Neandertal genome. *Science* 328: 710–722.
- Hawks J (2013) Significance of Neandertal and Denisovan genomes in human evolution. *Annual Review of Anthropology* 42: 433–449.
- Sankararaman S, Patterson N, Li H, Pääbo S, Reich D (2012) The date of interbreeding between Neandertals and modern humans. *PLoS Genetics* 8: e1002947.
- Meyer M, Fu Q, Aximu-Petri A, Glocke I, Nickel B, et al. (2014) A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* 505: 403–406.
- Gonder MK, Disotell TR, Oates JF (2006) New genetic evidence on the evolution of chimpanzee populations and implications for taxonomy. *International Journal of Primatology* 27: 1103–1127.
- Becquet C, Patterson N, Stone AC, Przeworski M, Reich D (2007) Genetic structure of chimpanzee populations. *PLoS Genetics* 3: e66.
- Gonder MK, Locatelli S, Ghobrial L, Mitchell MW, Kujawski JT, et al. (2011) Evidence from Cameroon reveals differences in the genetic structure and histories of chimpanzee populations. *Proceedings of the National Academy of Sciences of the United States of America* 108: 4766–4771.
- Auton A, Fedel-Alon A, Pfeifer S, Venn O, Ségurel L, et al. (2012) A fine-scale chimpanzee genetic map from population sequencing. *Science* 336: 193–198.
- Fischer A, Pollack J, Thalmann O, Nickel B, Pääbo S (2006) Demographic history and genetic differentiation in apes. *Current Biology* 16: 1133–1138.
- Fischer A, Prüfer K, Good JM, Halbwax M, Wiebe V, et al. (2011) Bonobos fall within the genomic variation of chimpanzees. *PLoS ONE* 6: e21605.
- Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, et al. (2013) Great ape genetic diversity and population history. *Nature* 499: 471–475.
- Champneys F (1871) On the muscles and nerves of a chimpanzee (*Troglodytes niger*) and a *Cynocepalus anubis*. *J Anat Lond* 6: 176–211.
- Crass E (1952) Musculature of the hip and thigh of the chimpanzee: a comparison to man and other primates. PhD thesis, Univ Wisconsin.
- Signon BA (1974) A functional analysis of pongid hip and thigh musculature. *Journal of Human Evolution* 3: 161–185.
- Stern JT (1972) Anatomical and functional specializations of human gluteus maximus. *American Journal of Physical Anthropology* 36: 315–338.
- Morimoto N, Zollikofer CPE, Ponce de León MS (2011) Femoral morphology and femoropelvic musculoskeletal anatomy of humans and great apes: a comparative virtopsy study. *Anatomical Record* 294: 1433–1445.
- Bourne GH, editor (1969) *The Chimpanzee, Vol. 1: Anatomy, Behavior, and Diseases of Chimpanzees*. Basel/New York: Karger.
- Bourne GH, editor (1971) *The Chimpanzee, Vol. 4: Behavior, Growth, and Pathology of Chimpanzees*. Basel/New York: Karger.
- Carlson KJ, Doran-Sheehy DM, Hunt KD, Nishida T, Yamanaka A, et al. (2006) Locomotor behavior and long bone morphology in individual free-ranging chimpanzees. *Journal of Human Evolution* 50: 394–404.
- Doran DM (1993) Comparative locomotor behavior of chimpanzees and bonobos - the influence of morphology on locomotion. *American Journal of Physical Anthropology* 91: 83–98.
- Lieberman DE, Carlo J, Ponce de León MS, Zollikofer CP (2007) A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *J Hum Evol* 52: 647–662.
- Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, et al. (1994) Kin selection, social-structure, gene flow, and the evolution of chimpanzees. *Science* 265: 1193–1201.
- Doran DM (1992) The ontogeny of chimpanzee and pygmy chimpanzee locomotor behavior: a case-study of paedomorphism and its behavioral-correlates. *Journal of Human Evolution* 23: 139–157.
- Doran DM (1996) Comparative positional behavior of the African apes. In: McGrew MC, Marchant LF, Nishida T, editors. *Great Ape Societies*. Cambridge: Cambridge University Press.

## Table S8 Correlation between PC scores and taxon-specific adult body masses.

(DOCX)

## Text S1 Materials and methods.

(DOCX)

## Text S2 Habitats of *Pan* taxa.

(DOCX)

## Text S3 *In-vivo* bone modification in the femur of *Pan* taxa.

(DOCX)

## Acknowledgments

We thank P. Jans, E Gillissen and W Coudyzer for help with sample preparation and CT scanning. The comments of H. Bagheri, M. Kobayashi, and T. Marques-Bonet are greatly acknowledged. We are also grateful to the anonymous reviewers for their valuable comments and suggestions.

## Author Contributions

Conceived and designed the experiments: NM MSPDL CPEZ. Performed the experiments: NM. Analyzed the data: NM. Contributed reagents/materials/analysis tools: NM MSPDL CPEZ. Contributed to the writing of the manuscript: NM MSPDL CPEZ.

37. Doran DM (1997) Ontogeny of locomotion in mountain gorillas and chimpanzees. *J Hum Evol* 32: 323–344.
38. Doran DM, Jungers WL, Sugiyama Y, Fleagle J, Heesy C (2002) Multivariate and phylogenetic approaches to understanding chimpanzee and bonobo behavioral diversity. In: Boesch C, Hohmann G, Marchant LF, editors. *Behavioural diversity in Chimpanzees and Bonobos*. Cambridge: Cambridge University Press. 14–34.
39. Goodall J (1986) *The Chimpanzees of Gombe*. Cambridge: Harvard University Press.
40. McGrew MC, Marchant LF, Nishida T, editors (1996) *Great Ape Societies*. Cambridge: Cambridge University Press.
41. Nei M (2007) The new mutation theory of phenotypic evolution. *Proc Natl Acad Sci U S A* 104: 12235–12242.
42. West-Eberhard MJ (2005) Developmental plasticity and the origin of species differences. *Proc Natl Acad Sci U S A* 102 Suppl 1: 6543–6549.
43. Shaw KL, Mullen SP (2011) Genes versus phenotypes in the study of speciation. *Genetica* 139: 649–661.
44. Collard M, Wood B (2000) How reliable are human phylogenetic hypotheses? *Proceedings of the National Academy of Sciences* 97: 5003–5006.
45. Strait DS, Grine FE (2004) Inferring hominoid and early hominid phylogeny using craniodental characters: the role of fossil taxa. *Journal of Human Evolution* 47: 399–452.
46. Morimoto N, Zollikofer CPE, Ponce de León MS (2012) Shared human-chimpanzee pattern of perinatal femoral shaft morphology and its implications for the evolution of hominin locomotor adaptations. *PLoS ONE* 7: e41980.
47. Geiger M, Forasiepi AM, Koyabu D, Sanchez-Villagra MR (2013) Heterochrony and post-natal growth in mammals - an examination of growth plates in limbs. *Journal of Evolutionary Biology* 20: 12279.
48. Koyabu D, Endo H, Mitgutsch C, Suwa G, Catania KC, et al. (2011) Heterochrony and developmental modularity of cranial osteogenesis in lipotyphlan mammals. *EvoDevo* 2: 21.
49. Wilson LAB, Sánchez-Villagra MR (2011) Evolution and phylogenetic signal of growth trajectories: the case of chelid turtles. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 316B: 50–60.
50. Sánchez M (2012) *Embryos in Deep Time: The Rock Record of Biological Development*. University of California Press. 265 p.
51. Ponce de León MS, Zollikofer CPE (2001) Neanderthal cranial ontogeny and its implications for late hominid diversity. *Nature* 412: 534–538.
52. Ackermann RR (2005) Ontogenetic integration of the hominoid face. *Journal of Human Evolution* 48: 175–197.
53. Gunz P, Neubauer S, Golovanova L, Doronichev V, Maureille B, et al. (2012) A uniquely modern human pattern of endocranial development. Insights from a new cranial reconstruction of the Neanderthal newborn from Mezmaiskaya. *Journal of Human Evolution* 62: 300–313.
54. Palagi E, Cordoni G (2012) The right time to happen: play developmental divergence in the two *Pan* species. *PLoS ONE* 7: e52767.
55. Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* 4: e1000008.
56. Morgan TJ, Evans MA, Garland T Jr, Swallow JG, Carter PA (2005) Molecular and quantitative genetic divergence among populations of house mice with known evolutionary histories. *Heredity* 94: 518–525.
57. Whitlock MC (2008) Evolutionary inference from  $Q_{ST}$ . *Molecular Ecology* 17: 1885–1896.
58. Smith HF (2009) Which cranial regions reflect molecular distances reliably in humans? Evidence from three-dimensional morphology. *American Journal of Human Biology* 21: 36–47.
59. Brommer JE (2011) Whither  $P_{ST}$ ? The approximation of  $Q_{ST}$  by  $P_{ST}$  in evolutionary and conservation biology. *Journal of Evolutionary Biology* 24: 1160–1168.
60. Edelaar PIM, Burraco P, Gomez-Mestre I (2011) Comparisons between  $Q_{ST}$  and  $F_{ST}$ —how wrong have we been? *Molecular Ecology* 20: 4830–4839.
61. Wright S (1969) *Evolution and the Genetics of Populations, Vol. II. The Theory of Gene Frequencies*. Chicago: University of Chicago Press.
62. Wright S (1978) *Evolution and the Genetics of Populations, Vol. IV. Variability Within and Among Natural Populations*. Chicago: University of Chicago Press.
63. Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting  $F_{ST}$ . *Nature Reviews Genetics* 10: 639–650.
64. Spitze K (1993) Population-structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135: 367–374.
65. Leinonen T, McCairns RJS, O'Hara RB, Merilä J (2013)  $Q_{ST}$ – $F_{ST}$  comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics* 14: 179–190.
66. Leinonen T, Cano JM, MÄKinen H, Merilä J (2006) Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology* 19: 1803–1812.
67. Merilä J, Björklund M, Baker AJ (1997) Historical demography and present day population structure of the greenfinch, *carduelis chloris*—an analysis of mtDNA control-region sequences. *Evolution* 51: 946–956.
68. Storz JF (2002) Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. *Molecular Ecology* 11: 2537–2551.
69. Saint-Laurent R, Legault M, Bernatchez L (2003) Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchell). *Molecular Ecology* 12: 315–330.
70. Slate J (2013) From beavis to beak color: a simulation study to examine how much qtl mapping can reveal about the genetic architecture of quantitative traits. *Evolution* 67: 1251–1262.
71. Atchley WR (1984) Ontogeny, timing of development, and genetic variance-covariance structure. *American Naturalist* 123: 519–540.
72. Charmantier A, Perrins C, McCleery RH, Sheldon BC (2006) Age-dependent genetic variance in a life-history trait in the mute swan. *Proceedings of the Royal Society B-Biological Sciences* 273: 225–232.
73. Lesser KJ, Paiusi IC, Leips J (2006) Naturally occurring genetic variation in the age-specific immune response of *Drosophila melanogaster*. *Aging Cell* 5: 293–295.
74. Pujol B, Wilson AJ, Ross RIC, Pannell JR (2008) Are  $Q_{ST}$ – $F_{ST}$  comparisons for natural populations meaningful? *Molecular Ecology* 17: 4782–4785.
75. Kalinowski ST (2002) Evolutionary and statistical properties of three genetic distances. *Molecular Ecology* 11: 1263–1273.
76. Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* 19: 153–170.
77. Cavalli-Sforza LL, Edwards AW (1967) Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet* 19: 233–257.
78. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genetics* 2: e190.
79. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
80. Slice DE (2007) Geometric morphometrics. *Annual Review of Anthropology* 36: 261–281.
81. Mitteroecker P, Gunz P (2009) Advances in geometric morphometrics. *Evolutionary Biology* 36: 235–247.
82. Zollikofer CPE, Ponce de León MS (2005) *Virtual Reconstruction. A Primer in Computer-Assisted Paleontology and Biomechanics*. Hoboken NJ: John Wiley & Sons.
83. Bookstein F (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge University Press.
84. Gunz P, Mitteroecker P, Bookstein FL (2005) Semilandmarks in Three Dimensions. In: Slice DE, editor. *Developments in Primatology: Progress and Prospects*. New York: Springer.
85. Kuhl F, Giardina C (1982) Elliptic Fourier features of a closed contour. *Computer graphics and image processing* 18: 236–258.
86. Specht M, Lebrun R, Zollikofer CPE (2007) Visualizing shape transformation between chimpanzee and human brains. *Visual Computer* 23: 743–751.
87. Morimoto N, Zollikofer CPE, Ponce de León MS (2011) Exploring femoral diaphyseal shape variation in wild and captive chimpanzees by means of morphometric mapping: a test of Wolff's Law. *Anatomical Record* 294: 589–609.
88. Bolter DR, Zihlman AL (2011) Brief communication: dental development timing in captive *Pan paniscus* with comparisons to *Pan troglodytes*. *American Journal of Physical Anthropology* 145: 647–652.
89. Carlson K, Sumner D, Morbeck M, Nishida T, Yamanaka A, et al. (2008) Role of nonbehavioral factors in adjusting long bone saphyseal structure in free-ranging *Pan troglodytes*. *International Journal of Primatology* 29: 1401–1420.
90. Zollikofer CPE, Ponce de León MS (2001) Computer-assisted morphometry of hominoid fossils: the role of morphometric maps. In: De Bonis L, Koufos G, Andrews P, editors. *Phylogeny of the Neogene Hominoid Primates of Eurasia*. Cambridge: Cambridge University Press. 50–59.
91. Bondioli L, Bayle P, Dean C, Mazurier A, Puymerail L, et al. (2010) Technical note: Morphometric maps of long bone shafts and dental roots for imaging topographic thickness variation. *American Journal of Physical Anthropology* 142: 328–334.
92. Puymerail L (2013) The functionally-related signatures characterizing the endostructural organisation of the femoral shaft in modern humans and chimpanzee. *Comptes Rendus Palevol*.
93. Puymerail L, Ruff CB, Bondioli L, Widiyanto H, Trinkaus E, et al. (2012) Structural analysis of the Kresna 11 *Homo erectus* femoral shaft (Sangiran, Java). *Journal of Human Evolution* 63: 741–749.
94. Maddison WP (1991) Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. *Systematic Biology* 40: 304–314.
95. Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources* 11: 353–357.
96. Relethford JH, Blangero J (1990) Detection of differential gene flow from patterns of quantitative variation. *Human Biology* 62: 5–25.
97. Relethford JH, Crawford MH, Blangero J (1997) Genetic drift and gene flow in post-famine Ireland. *Human Biology* 69: 443–465.
98. Sæther SA, Fiske P, Kålås JA, Kuresoo A, Luiquijõe L, et al. (2007) Inferring local adaptation from  $Q_{ST}$ – $F_{ST}$  comparisons: neutral genetic and quantitative trait variation in European populations of great snipe. *Journal of Evolutionary Biology* 20: 1563–1576.
99. Chapuis E, Martin G, Goudet J (2008) Effects of selection and drift on G matrix evolution in a heterogeneous environment: a multivariate  $Q_{ST}$ – $F_{ST}$  test with the freshwater snail *Galba truncatula*. *Genetics* 180: 2151–2161.

100. Martin G, Chapuis E, Goudet J (2008) Multivariate  $Q_{st}$ - $F_{st}$  comparisons: a neutrality test for the evolution of the G matrix in structured populations. *Genetics* 180: 2135–2149.
101. Relethford JH (2004) Boas and beyond: Migration and craniometric variation. *American Journal of Human Biology* 16: 379–386.
102. Zollikofer CPE, Ponce de León MS (2006) Neanderthals and modern humans - chimps and bonobos: similarities and differences in development and evolution. In: Harvati K, Harrison T, editors. *Neanderthals Revisited: New Approaches and Perspectives*. New York: Springer. 71–88.
103. Penin X, Berge C, Baylac M (2002) Ontogenetic study of the skull in modern humans and the common chimpanzees: neotenic hypothesis reconsidered with a tridimensional Procrustes analysis. *Am J Phys Anthropol* 118: 50–62.
104. Willing E-M, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by *Fst* do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE* 7: e42649.
105. Hunt G (2007) The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proceedings of the National Academy of Sciences of the United States of America* 104: 18404–18408.
106. Haller BC, Hendry AP (2014) Solving the paradox of stasis: squashed stabilizing selection and the limits of detection. *Evolution* 68: 483–500.
107. Wagner GP (1996) Apparent stabilizing selection and the maintenance of neutral genetic variation. *Genetics* 143: 617–619.
108. Kirk H, Freeland JR (2011) Applications and implications of neutral versus non-neutral markers in molecular ecology. *Int J Mol Sci* 12: 3966–3988.
109. McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution* 17: 285–291.
110. Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55: 1095–1103.
111. Carlson KJ, Lublinsky S, Judex S (2008) Do different locomotor modes during growth modulate trabecular architecture in the murine hind limb? *Integrative and Comparative Biology* 48: 385–393.
112. Smith RJ, Jungers WL (1997) Body mass in comparative primatology. *Journal of Human Evolution* 32: 523–559.
113. Somel M, Franz H, Yan Z, Lorenc A, Guo S, et al. (2009) Transcriptional neoteny in the human brain. *Proceedings of the National Academy of Sciences of the United States of America* 106: 5743–5748.
114. Somel M, Liu X, Tang L, Yan Z, Hu H, et al. (2011) MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS Biology* 9: e1001214.
115. Mitteroecker P, Gunz P, Bookstein FL (2005) Heterochrony and geometric morphometrics: a comparison of cranial growth in *Pan paniscus* versus *Pan troglodytes*. *Evolution & Development* 7: 244–258.
116. Shea BT (1983) Pedomorphosis and neoteny in the pygmy chimpanzee. *Science* 222: 521–522.
117. Shea BT (1983) Allometry and heterochrony in the African apes. *American Journal of Physical Anthropology* 62: 275–289.
118. Khaitovich P, Enard W, Lachmann M, Pääbo S (2006) Evolution of primate gene expression. *Nature Reviews Genetics* 7: 693–702.
119. Lynch M (1989) Phylogenetic hypotheses under the assumption of neutral quantitative-genetic variation. *Evolution* 43: 1–17.
120. Roseman CC (2004) Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. *Proc Natl Acad Sci U S A* 101: 12824–12829.
121. Lovejoy CO, Meindl RS, Ohman JC, Heiple KG, White TD (2002) The Maka femur and its bearing on the antiquity of human walking: Applying contemporary concepts of morphogenesis to the human fossil record. *American Journal of Physical Anthropology* 119: 97–133.
122. Ruff CB (2002) Long bone articular and diaphyseal structure in old world monkeys and apes. I: Locomotor effects. *American Journal of Physical Anthropology* 119: 305–342.
123. Carlson KJ (2005) Investigating the form-function interface in African apes: Relationships between principal moments of area and positional behaviors in femoral and humeral diaphyses. *Am J Phys Anthropol* 127: 312–334.
124. Leinonen T, O'Hara RB, Cano JM, Merilä J (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology* 21: 1–17.