Abstract – Pelvic dimensions differ, on average, among modern human populations. Some recent studies have reported that population differences in pelvic form evolved mainly by neutral processes, without considerable natural selection. This is a surprising claim given the many important functions of the human pelvis. This article re-evaluates this claim through a review of the medical literature and a new analysis of the data from which these results were obtained. We show that variation in modern pelvic form correlates with the risks of obstructed labour, pelvic floor disorders such as incontinence and prolapse, and numerous orthopaedic disorders that impair walking. Comparative population studies also document adaptations of human body form and pelvic dimensions to climatic conditions. However, these various and partly antagonistic selective forces on the human pelvis are not homogeneous across populations. They depend not only on climatic differences, but also on maternal and foetal body size, pelvic floor tissue properties, diet, lifestyle and the resulting metabolic capacities, as well as on obstetric practices and health care, all of which vary across human populations. Given the relatively rapid evolution of polygenic quantitative traits, we propose that pelvic dimensions have evolved different “compromise solutions” in different populations in response to local selective regimes. The results from a reanalysis of a large published global dataset on human pelvic canal dimensions clearly support this view.

Keywords – allometry, childbirth, human evolution, natural selection, pelvis

Did population differences in human pelvic form evolve by drift or selection?

Les différences entre populations de la forme du bassin humain ont-elles évolué par dérive ou par sélection ?

Philipp Mitteroecker 1*, Nicole DS Grunstra 2,3, Ekaterina Stansfield 1, Lukas Waltenberger 1,4, Barbara Fischer 1

1 Department of Evolutionary Biology, Unit for Theoretical Biology, University of Vienna, Vienna, Austria
2 Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria
3 Mammal Collection, Natural History Museum Vienna, Vienna, Austria
4 Austrian Archaeological Institute, Austrian Academy of Sciences, Vienna, Austria

* philipp.mitteroecker@univie.ac.at

Reçu : 27 novembre 2020 ; accepté : 25 février 2021
Bulletin et Mémoires de la Société d’Anthropologie de Paris
dans le cadre de la session “Les 1000 premiers jours de vie dans les populations du présent et du passé”

Mots clés – allométrie, accouchement, évolution humaine, sélection naturelle, pelvis
Introduction

The human bony pelvis has numerous functions. It carries the weight of the upper body and connects the spine to the lower limbs, so that the form, position and orientation of the acetabula determine the stability of the hip joint and mobility of the femur (Lovejoy, 1988; Tannest et al., 2012; Werner et al., 2012; Ruff, 2017). The pelvic canal forms the bony part of the birth canal and anchors the pelvic floor, which has to resist the pressure exerted by the inner organs and the foetus while being flexible enough for parturition and excretion (Schimpi and Tulikangas, 2005; Pavlicev et al., 2020). The coccyx and the lower part of the sacrum as well as the ischial spines protrude into the birth canal and support the pelvic floor (Abitbol, 1988). Finally, the length and orientation of the iliac blades are associated with the dimensions of the thorax, which in turn affect the surface-to-volume ratio of the body and thus, indirectly, also influence thermoregulation (Ruff, 1991; 1994; Betti, 2017).

As the modern human pelvis differs considerably from that of great apes, the overall form of the human pelvis is widely considered to be an adaptation to bipedal locomotion (e.g., Lovejoy, 1988; Wittman and Wall, 2007; Grabowski and Roseman, 2015; Ruff, 2017; Lewis et al., 2017; Gruss et al., 2017). The pelvis is also the most dimorphic region in the human skeleton and the only part that has, on average, larger absolute dimensions in females than in males (e.g., Tague, 1992). This dimorphism is typically considered an evolutionary adaptation to ease parturition of the relatively large human foetuses (e.g., Tague, 1992; Rosenberg and Trevathan, 2005; Weaver and Hublin, 2009; Grabowski and Roseman, 2015; Grunstra et al., 2019; Pavlicev et al., 2020). This is further supported by the observation that pelvic sex differences are larger in magnitude in primates and other mammals that give birth to foetuses with relatively larger heads or bodies compared to species with smaller neonates (Leutenegger, 1974; Ridley, 1995; Tague, 2016; Moffett, 2017; Grunstra et al., 2019).

In addition to massive individual variation, average pelvic dimensions also differ among modern human populations (Betti and Manica, 2018; DelPrete, 2019). Whereas some of these differences have frequently been interpreted as adaptations to climatic and environmental differences (Ruff, 2002; Wells et al., 2012; Gruss et al., 2017; Savell et al., 2016; Ruff, 2017), several recent papers have suggested that global population differences mainly arose from neutral evolutionary processes, i.e., with natural selection playing a lesser role (Betti et al., 2013; Betti, 2017; Betti and Manica, 2018). Given the many biomechanical and reproductive functions of the pelvis, this is a rather surprising result, which we aim to reassess here. Moreover, pelvic form correlates with body proportions and stature (Arsuaga and Carretero, 1994; Kurki, 2013; Fischer and Mitteroecker, 2015; Torres-Tamayo et al., 2018; Torres-Tamayo et al., 2020), which have been repeatedly subject to selection in human populations (Ruff, 2002; Field et al., 2016; Jungers et al., 2016; Savell et al., 2016; Stulp et al., 2011; Guo et al., 2018). Thus, even if the pelvis was not the direct target of natural selection, it would be equally surprising if pelvic form did not respond to indirect selection acting on non-pelvic traits.

The fields of evolutionary biology and anthropology share an adaptationist tradition. Too often in biology, phenotypic differences have been presented as specific evolutionary adaptations without a mechanistic model or empirical evidence for fitness consequences (Gould and Lewontin, 1979; Orzack et al., 2001; Dunsdworth, 2020). We believe that researchers should be sceptical of claims about adaptive traits and patterns without empirical evidence. Advances in evolutionary modelling have shown that many genetic and phenotypic differences can be explained in terms of neutral evolutionary processes (Lynch and Hill, 1986; Kimura, 1991). Moreover, research on developmental and genetic constraints has documented that traits can also evolve as indirect by-products of natural selection rather than as its direct target (e.g., Beldade et al., 2002; Müller, 2010; Santangelo et al., 2018). However, there are cases where traits are obviously the products of natural selection because variation in these traits is closely connected with variation in fitness (reproductive success). We suggest that multiple aspects of human pelvic morphology are such traits. Here, we briefly review potential selective forces acting on modern human pelvic morphology and discuss ways of inferring evolutionary processes from morphological data. We present a reanalysis of pelvic measurements by Betti and Manica (2018) from 24 human populations. Our findings contradict their interpretation that variation in pelvic morphology among populations is predominantly neutral.

Evidence for selection on human pelvic form

Even if the transformation of the ape-like pelvis into the modern human pelvis was the result of natural selection, it is nonetheless possible that the differences in pelvic form observable today are not associated with any functional differences that influence the probabilities of survival or reproduction. In other words, modern variation in pelvic form could be selectively neutral if all individuals have a pelvis that sufficiently enables walking, continence and parturition. The large body of medical literature documents how variation in pelvic morphology can affect the functionality of the human body. Especially in the fields of obstetrics and orthopaedics, researchers have established a deep understanding of how pelvic form is linked to function.

Most obviously, pelvic form affects the success of the second stage of labour. Cephalopelvic disproportion and shoulder dystocia are still common causes of maternal death (Dolea and AbouZahr, 2003; Neilson et al., 2003). Exact incidences of cephalopelvic disproportion are difficult to estimate and vary widely. In Africa, reported rates range from 1.4 to 8.5% (Dumont et al., 2001). A large US study reported disproportion rates of 2.3% for infants weighing 3.0 to 3.9 kg at birth, and 5.8% for those weighing more (Boulet et al., 2003). Thus, even the most conservative
estimate entails about 40,000 births affected in the United States and about 1.3 million worldwide every year (Pavlíčev et al., 2020). More common still are maternal morbidities, such as fistulas, uterine rupture and injury to the sphincter muscles, resulting from delayed labour because of a tight foetopelvic fit (Arrowsmith et al., 1996). Chronic incontinence can lead to serious infections and social ostracism, as is still common in sub-Saharan Africa (Wall, 1999; 2006). Clearly, the size of the foetus matters, but what determines the success of labour is not the “passenger” alone, but also the maternal “passageway” and the strength of uterine contractions. Clinical studies have confirmed that maternal pelvic dimensions correlate with the risk of cephalopelvic disproportion (Zaretsky et al., 2005; Harper et al., 2013; Franz et al., 2017). Thus, in the absence of modern medical care, childbirth exerts a relatively strong selective pressure on pelvic form. This “obstetric selection” clearly favours a spacious birth canal.

At the same time, it has been suggested that a narrow pelvic canal, including protruding ischial spines and sacrum, would support the pelvic floor in resisting the pressure exerted by the inner organs and the foetus (Abitbol, 1988; Schimpf and Tulikangas, 2005; Brown et al., 2013; Huseynov et al., 2016; Grunstra et al., 2019; Pavlíčev et al., 2020). Several medical studies have shown that women with a wider pelvic canal are more likely to experience pelvic organ prolapse and incontinence (Sze et al., 1999; Handa et al., 2003; Stav et al., 2007; Brown et al., 2013; Berger et al., 2013). Clinically, the form of the pelvic canal is only a minor risk factor for pelvic floor disorders, which presumably is the reason why some studies did not find such an association, but a finite element simulation study confirmed that larger pelvic floors descend more strongly in response to pressure (Stansfield et al., 2021). Therefore, pelvic floor stability imposes a directional selection gradient towards a smaller pelvic canal. Note that even a weak, perhaps clinically insignificant association between a trait value and fitness (i.e., a heritability of most pelvic dimensions \( h^2 \), of 0.5 to 0.8 (Sharma, 2002). The heritability, \( h^2 \), of most pelvic dimensions has been reported to range from 0.5 to 0.8 (Sharma, 2002). Exposed to directional selection of intermediate strength (e.g., \( \beta=0.13 \); cf. Hoekstra et al., 2001; Kingsolver et al., 2001), a mean change of one standard deviation (sd) would require only 10 to 15 generations, depending on the exact heritability. Hence, a quantitative trait can evolve surprisingly fast, if not impeded by developmental constraints, limited genetic variation or antagonistic indirect selection.

At the same time, it has been suggested that a narrow pelvic canal, including protruding ischial spines and sacrum, would support the pelvic floor in resisting the pressure exerted by the inner organs and the foetus (Abitbol, 1988; Schimpf and Tulikangas, 2005; Brown et al., 2013; Huseynov et al., 2016; Grunstra et al., 2019; Pavlíčev et al., 2020). Several medical studies have shown that women with a wider pelvic canal are more likely to experience pelvic organ prolapse and incontinence (Sze et al., 1999; Handa et al., 2003; Stav et al., 2007; Brown et al., 2013; Berger et al., 2013). Clinically, the form of the pelvic canal is only a minor risk factor for pelvic floor disorders, which presumably is the reason why some studies did not find such an association, but a finite element simulation study confirmed that larger pelvic floors descend more strongly in response to pressure (Stansfield et al., 2021). Therefore, pelvic floor stability imposes a directional selection gradient towards a smaller pelvic canal. Note that even a weak, perhaps clinically insignificant association between a trait value and fitness (i.e., a heritability of most pelvic dimensions \( h^2 \), of 0.5 to 0.8 (Sharma, 2002). The heritability, \( h^2 \), of most pelvic dimensions has been reported to range from 0.5 to 0.8 (Sharma, 2002). Exposed to directional selection of intermediate strength (e.g., \( \beta=0.13 \); cf. Hoekstra et al., 2001; Kingsolver et al., 2001), a mean change of one standard deviation (sd) would require only 10 to 15 generations, depending on the exact heritability. Hence, a quantitative trait can evolve surprisingly fast, if not impeded by developmental constraints, limited genetic variation or antagonistic indirect selection.

To summarise, a broad body of biological and medical literature clearly demonstrates that human pelvic form has been and continues to be subject to natural selection. This does not imply that all aspects of pelvic morphology are subject to selection. Pelvic features that are functionally less relevant may have been free to evolve neutrally as long as they were developmentally and genetically unconstrained. But is it plausible that natural selection has been strong enough to account for the observed population differences?

**Is evolution by natural selection fast enough?**

Quantitative polygenic traits, such as pelvic dimensions, are expected to change in response to natural selection if the trait value shows a consistent correlation with fitness (i.e., the average number of offspring; Lande, 1979; Roff, 2012). Many factors contribute to fitness, including survival, mating success and fecundity. Under certain assumptions, the expected change in the population mean value equals the product of the heritability of the trait and the strength of selection, as expressed by the selection gradient, \( \beta \) (Lande, 1979; Roff, 2012). The heritability, \( h^2 \), of most pelvic dimensions has been reported to range from 0.5 to 0.8 (Sharma, 2002). Exposed to directional selection of intermediate strength (e.g., \( \beta=0.13 \); cf. Hoekstra et al., 2001; Kingsolver et al., 2001), a mean change of one standard deviation (sd) would require only 10 to 15 generations, depending on the exact heritability. Hence, a quantitative trait can evolve surprisingly fast, if not impeded by developmental constraints, limited genetic variation or antagonistic indirect selection.
Betti and Manica (2018) reported measurements of anteroposterior and mediolateral diameters of the pelvic inlet, midplane and outlet for 24 human populations. The maximal pairwise differences between the 24 population means ranged from 2.5 to 3.8 sd for the six variables, but the median population differences ranged only from 0.5 to 1.1 sd. Thus, even the most pronounced pelvic differences between modern human populations could have evolved within less than 60 generations of continual directional selection of intermediate strength. Even weak selection would have sufficed to produce the observed differences after the divergence of the populations.

Can selection explain population divergence in pelvic form?

As discussed above, it is conceivable that some of the population differences in the upper pelvis arose as adaptations to different climates. But how could natural selection have led to population differences in the lower pelvis and the birth canal, given that successful parturition and effective walking are important to all individuals?

The modern human birth canal presumably evolved as a “compromise solution” to the different opposing selective forces (Washburn, 1960; Rosenberg and Trevathan, 2005; Fischer and Mitteroecker, 2015; Wells, 2015; Mitteroecker et al., 2016). In other words, the evolved population distribution “trades-off” the advantage of a large pelvic canal for parturition against its disadvantages for pelvic floor support, locomotion and thermoregulation. The population mean of pelvic canal form is therefore determined by the relative strengths of the antagonistic selection pressures (Mitteroecker et al., 2016; 2017), which may vary between populations. For instance, in a population with relatively large neonates the obstetric selection pressure is increased, eventually leading to the evolution of a slightly more spacious birth canal and, consequently, higher rates of pelvic floor disorders. Conversely, in a population with more vulnerable pelvic floor tissues or a lifestyle imposing more stress on the pelvic floor, selection for a narrower birth canal to support the pelvic floor is increased. This may lead to the evolution of a smaller birth canal and more difficult births or, alternatively, to smaller neonates. Obstetric practices, including caesarean sections, can also reduce selective pressure resulting from childbirth and affect the evolutionary dynamics (Mitteroecker et al., 2016; Mitteroecker, 2019).

We suggest that the evolutionary trade-offs affecting pelvic form and the resulting compromise solutions shift in response to biological, environmental, and sociocultural transitions. We refer to this as “shifting trade-off dynamics”. The factors contributing to the trade-off indeed vary between human populations. Mean birth weight varies considerably from about 2.7 to 3.6 kg across human populations (Meredith, 1970; Mikolajczyk et al., 2011; Blencowe et al., 2019), imposing variable magnitudes of obstetric selection. Maternal stature is a well-known obstetric risk factor because, on average, shorter women have harder births and a higher risk of cephalopelvic disproportion (Camilleri, 1981; Dougherty and Jones, 1988; Stulp et al., 2011), but average female stature varies considerably among human populations (country-specific female average stature ranges from 149 to 170 cm; NCD-RisC 2016). Also, mean neonatal head circumference and gestation length vary globally (Patel et al., 2004; Steer, 2005; Janssen et al., 2007; Villar et al., 2014). Likewise, age-specific prevalences of pelvic organ prolapse and incontinence vary across countries as well as by ethnicity and sociocultural background (Walker and Gunasekera, 2011; Vergeldt et al., 2015; Islam et al., 2019), as does the length of labour (Hass et al., 1980; Greenberg et al., 2006). It has been reported that both adult body proportions and birth weight differ between populations living at low versus high altitudes (Hass et al., 1980; Moore, 2017), perhaps linking the selective pressures on the upper and lower pelvis. Furthermore, physical activities, subsistence strategies and diet differ among populations and cultures, thus exposing the pelvis and the pelvic floor to different physical stresses (e.g., Walker and Gunasekera, 2011) and providing different metabolic capacities during pregnancy (Dunsworth et al., 2012). It has also been suggested that transitions in environmental and socioeconomic conditions can affect the relationship between foetal and maternal size, thus influencing the difficulty of labour (Wells, 2015; Zaffarini and Mitteroecker, 2019).

To summarise, the anatomical, physiological, environmental and sociocultural factors affecting childbirth, locomotion patterns and pelvic floor function vary between modern human populations. It is therefore likely that these differences have influenced the evolutionary trade-off dynamics and given rise to different pelvic compromise solutions in human populations.

How to interpret differences in pelvic dimensions?

Previous claims about the neutral evolution of pelvic form were not based on medical, functional or experimental studies as reviewed above. For instance, in a worldwide sample of human skeletons, Betti and Manica (2018) measured anteroposterior (AP) and mediolateral (ML) diameters of the inlet, midplane and outlet of the pelvic canal and proposed to test the “null hypothesis that geographic patterns [in these dimensions] reflect past population history; significant deviations above and beyond this simple expectation can be taken as potentially representing the effects of natural selection” (p. 4). They reported a decline in phenotypic diversity with increasing distance from Africa and a modest correlation ($R^2=0.24$) between phenotypic distances and genome-wide genetic distances across 24 human populations. Furthermore, they reported weak or absent correlations between phenotypic distances and differences in temperature across populations. They also found that pelvic canal shape (the ratio AP/ML) varied more in their sample than other indices of body shape. Finally, they did not confirm the findings of Fischer and Mitteroecker (2015) that pelvic canal shape correlates with stature. From these findings,
Betti and Manica inferred that “shape differences between main geographical regions have likely arisen from a stochastic drift towards different average shapes” (p. 7).

We propose an alternative evolutionary scenario to neutral pelvic evolution and derive a set of hypotheses against which to evaluate the statistical results. The medical, functional, and comparative literature reviewed above documents that the human pelvis has been under strong selection pressure, even in recent history, but this selection is not homogeneous across populations. Given the relatively rapid evolution of polygenic quantitative traits, we expect pelvic dimensions to have evolved different distributions in different populations in response to local trade-off dynamics. Therefore, we expect pelvic traits to differ, on average, between populations (H1). We expect these population differences in pelvic form to correlate with biometric, sociocultural and environmental factors that influence the evolutionary dynamics (H2). Body shape, physiology, environment and culture do not vary randomly but often are more similar in geographically close populations. We therefore expect differences in pelvic form to correlate positively with geographic distance among populations (H3). Selection pressures are likely to differ more between more distant populations, but many neutral genetic markers also show an isolation-by-distance pattern. Therefore, overall genetic distances are expected to correlate with pelvic form differences (H4). We thereby expect correlations between pelvic form and certain biometric, environmental and sociocultural variables to be stronger than correlations between pelvic form differences and geographic or genetic distances, as the latter are mainly a causal consequence of the former (H5).

Note that, contrary to Betti and Manica’s (2018) premise, a correlation between phenotypic dissimilarity and geographic distance or “population history” is not necessarily indicative of neutral evolution, because environmental variation that affects animal physiology (e.g., haemoglobin levels) or anatomy (e.g., body form) is itself often geographically structured. Only after statistically accounting for all these biological, environmental and sociocultural factors that differ across populations would a correlation between neutral genetic divergence and quantitative genetic variation of pelvic form be congruent with, but still not exclusive evidence of, a neutral mode of pelvic evolution.

All these hypotheses are about differences in average pelvic form between populations. But pelvic dimensions have also been shown to correlate with several aspects of body form within populations. Such correlations result from variation in pleiotropic genes and common developmental factors, such as circulating growth hormones and steroid hormones. For instance, found that women with large heads tend to have a birth canal that can better accommodate foetuses with large heads because they have a sacrum that protrudes less into the birth canal, thereby enlarging the pelvic outlet (Fischer and Mitteroecker, 2015). Head size and pelvic dimensions are heritable traits, implying that mothers with large heads tend to give birth to foetuses with large heads.

Fischer and Mitteroecker (2015) therefore interpreted these within-population correlations as adaptive because they ease birth. The same associations between maternal pelvic dimensions and foetal head size have subsequently been identified for rhesus macaques, which also have a tight cephalopelvic fit (Kawada et al., 2020). Many studies have consistently found that pelvic shape also covaries with stature. On average, shorter women have a rounder pelvic inlet, also called a gynecoid pelvis, than taller women, who have a more anthropoid pelvis with an anteroposteriorly oval pelvic inlet (Ince and Young, 1940; Bernard, 1952; Holland et al., 1982; Tague, 2000; Fischer and Mitteroecker, 2015; Mitteroecker et al., 2017). The obstetric literature suggests that gynecoid pelvises are more suitable for parturition than other pelvis shapes (anthropoid, platypelloid pelvis shapes; Caldwell and Moloy, 1939; Maharaj, 2010). Shorter women tend to have more difficult births and are therefore exposed to stronger obstetric selection, suggesting that this within-population correlation between stature and pelvic shape is also adaptive (Fischer and Mitteroecker, 2015).

In addition to pleiotropy, phenotypic correlations can also result from the linkage disequilibrium of genes affecting each trait separately, as opposed to variation in genes affecting both traits jointly. However, in an interbreeding population linkage disequilibrium decays rapidly unless correlational selection is strong and persistent (Sinervo and Svensson, 2002). For instance, obstetric selection may favour alleles for a gynecoid pelvis in short women and alleles for an anthropoid pelvis in tall women, but after random mating the alleles for pelvis shape and stature are recombined, thus reducing the correlation. But whenever gene flow is reduced, for example by a certain degree of geographic isolation or by assortative mating, phenotypic correlations across populations can be shaped more easily by correlational selection. For instance, in populations with a short average stature the frequency of alleles for a gynecoid pelvis may increase, whereas in taller populations the alleles for an anthropoid pelvis may become more common. We thus expect obstetric selection to have led to a correlation between pelvic form and stature both within and between populations, but the magnitude and pattern of this association across populations may differ from those within populations (H6).

Because of the out-of-Africa migration pattern, modern human genetic variation and also certain aspects of phenotypic variation have been shown to decrease with the geographic distance from Africa, a so-called serial founder effect (e.g., Manica et al., 2007; Henn et al., 2012). Regardless of whether pelvis-specific genetic alleles are affected by this kind of genetic drift, a decrease of within-population variance in pelvic form is expected to accompany a reduction of variance in overall body form (H7) because pelvic form and body form correlate within populations. Furthermore, reduced variation in maternal and foetal body form entails a smaller range of “optimal” pelvic forms, i.e., stronger stabilizing selection, which triggers an evolutionary reduction of pelvic form variance.
Reanalysis of the Betti and Manica (2018) data

Data

Betti and Manica (2018) measured the recomposed bony pelvis of 348 female skeletons from 24 populations following Kurki (2007), which they kindly made available on DRYAD (Betti and Manica, 2018a). They determined sex from non-metric pelvic features (Phenice 1969; Sutherland and Suchey 1991) and inferred body mass from femur diameter (Auerbach and Ruff, 2004). They calculated migration distances for these populations (shortest distances on land between the location of the population and the assumed origin of human migration in Africa). As environmental variables, the average minimum temperature of the coldest month and the average maximum temperature of the warmest month for each population were collected from WorldClim (www.worldclim.org; Hijmans et al., 2005). Based on 594,924 genome-wide SNPs collected from the Human Origins Database (Lazaridis et al., 2004), they calculated pairwise population $F_{ST}$ values. They compared variation in six anteroposterior (AP) and mediolateral (ML) diameters as well as in three pelvic shape indices (AP/ML for inlet, midplane and outlet) to the variation of non-pelvic body dimensions for the same populations from the Goldman dataset (Auerbach and Ruff, 2006; http://web.utk.edu/~auerbach/GOLD.htm).

Methods

We reanalysed these data by computing different variance ratios for the AP and ML diameters and shape indices of the pelvis as well as for the lengths and shape indices of long bones (see below for details). We also regressed these variables on minimum and maximum temperature using ordinary least-squares regression. To study allometric relationships, we regressed the natural logarithm of canal size (approximated by the square root of inlet AP×ML) and canal shape (AP/ML) on log body mass. Type I error rates were estimated by permutation tests (10,000 random permutations). A multivariate test of proportionality of the within-population and between-population covariance matrices was performed using the maximum-likelihood test described in Mardia et al. 1979.

To estimate the strength of divergent selection, we calculated global and pairwise $P_{ST}$ values for the different measurements and compared them with the reported $F_{ST}$ values (Leinonen et al. 2006; 2013; Brommer, 2011; Walsh and Lynch 2018). We calculated the $P_{ST}$ values following Brommer (2011):

$$P_{ST} = \frac{c \sigma^2_h}{c \sigma^2_h + 2h^2 \sigma^2_q}$$

where $\sigma^2_h$ denotes the phenotypic variance between populations, $\sigma^2_q$ the phenotypic variance within populations, and $h^2$ the heritability (the fraction of phenotypic variance owing to additive genetic effects). The variance components were computed by a one-way ANOVA following Storz et al. (2002). The scalar $c$ expresses the proportion of the total variance that is presumed to result from additive genetic effects between populations. As a conservative estimate, we set $c=h^2$ (cf. Brommer 2011), which implies that they cancel. Different values for $c$ produced very similar results.

We computed the total variances of the six pelvic diameters in the different populations as the sum of the six variances (trace of the covariance matrix). Likewise, we computed the total variances of the three shape indices. Confidence intervals were computed using bootstrap estimates (10,000 random bootstraps).

Results

Population differences

Betti and Manica (2018) reported significant and geographically patterned differences in pelvic canal dimensions. They also found that the magnitude of shape variation (as measured by the coefficient of variation of the three AP/ML ratios in their full sample) was consistently higher for the three pelvic planes than for the other body proportion indices. Both findings would be consistent with our hypothesis H1 and are not specific to a drift scenario. However, the coefficient of variation (CV, the standard deviation divided by the mean) only applies to variables with a natural zero value (i.e., with a ratio scale), not to shape variables (e.g., Houle et al., 2011; Pélabon et al., 2020). The finding that the three AP/ML ratios have a higher CV than the pelvic diameters and the long bone lengths simply derives from the fact that the length measurements have much larger mean values than the AP/ML ratios. Likewise, the long bones are considerably longer than the pelvic diameters, which challenges the comparison of the CVs (Polly, 1998).

Importantly, phenotypic variances in a multi-population sample are not interpretable per se in an evolutionary context. Instead, evolutionary theory predicts that, under pure genetic drift, the quantitative genetic variance between population means is equal to the quantitative genetic variance within the ancestral population multiplied by the number of generations, $t$, since population divergence, and divided by the average effective population size, $N_e$ (e.g., Roff, 2012; Walsh and Lynch, 2018). In other words, traits with large heritable variation within a population are also expected to undergo a higher degree of drift compared to traits that have little heritable variation. Deviations from this expectation indicate divergent or stabilizing selection. As we are comparing different variables across the same set of populations (with the same $t$ and $N_e$), the ratios of between-population variance to within-population variance can be compared among variables as rough estimates of the relative strength of divergent selection. This unitless variance ratio does not need to be corrected for mean differences (to yield a CV) as these would cancel anyway. Without reliable estimates of $t$ and $N_e$, however, only the relative magnitudes of these ratios are interpretable. For example,
a trait with a high variance ratio is likely to have undergone more divergent selection or less stabilizing selection than a trait with a lower variance ratio.

We reanalysed these data as well as the reference data for non-pelvic measurements. We computed the ratios of the between-population variance to the pooled within-population variance (as a proxy for the ancestral population variance) for the different length and shape variables. Similarly to the brachial and crural indices, all three pelvic shape variables showed relatively little divergence between populations as compared with the long bones (figure 1a). Interestingly, the AP diameters of the three pelvic planes also showed little population divergence, whereas the ML diameters showed divergences similar to that of the long bones. Compared to the midplane and outlet, the ML diameter and also the shape (AP/ML) of the inlet had the highest ratio of between-to-within-population variance. The most pronounced population divergence among all measurements was for bi-iliac breadth. To summarise, this suggests, relatively conserved AP pelvic dimensions and pelvic canal shape when compared with the long bones. The width of the upper pelvis and the pelvic inlet showed the greatest divergence between populations relative to within-population variance among all variables, suggesting that upper pelvic width may have been subject to divergent selection.

**Climate**

Betti and Manica (2018) reported “no significant correlation between overall canal shape differences and temperature differences between populations, before or after correcting for genetic distance. When the three planes were analysed separately, however, a significant correlation was found between inlet shape differences and temperature differences.”

![Figure 1](image1.png)

**Figure 1.** (a) Ratios of between-population variance to within-population variance for different length and shape variables. The blue bars correspond to variables from Betti and Manica (2018) and the white bars to variables from the Goldman data (Auerbach and Ruff, 2006). (b) $P_{ST}$ values (measures of population divergence) for the same length and shape variables.

![Figure 2](image2.png)

**Figure 2.** (a) Mean mediolateral diameter of the pelvic inlet plotted against minimum temperature for each of the 24 populations measured by Betti and Manica (2018). (b) Pelvic inlet shape (AP/ML) versus minimum temperature. (c) Body mass, as estimated from the acetabular diameter, versus minimum temperature.
Their multivariate test was based on a Mantel test, but in the presence of a directional statistical trend (such as Bergmann’s rule), a Mantel test has considerably less power than a multivariate linear regression of pelvic dimensions directly on temperature. We found that multivariate regressions of the six pelvic measurements on minimum temperature and also on maximum temperature were statistically significant (both at $p<0.0001$). Likewise, the multivariate regressions of the three shape ratios on minimum and maximum temperature were significant ($p<0.001$, $p=0.018$, respectively). When considered separately, all ML diameters were considerably and significantly correlated with minimum temperature (inlet: $r=0.54$, $p=0.007$; midplane: $r=0.65$, $p<0.001$; outlet: $r=-0.51$, $p=0.010$) and, to a slightly lesser extent, also with maximum temperature. Hence, the pelvic canal clearly tends to be relatively and absolutely wider in colder environments than in warmer ones (figure 2), as found by many previous studies and expected in our H2. Also, body mass was negatively related to minimum and maximum temperature ($p<0.0001$, $p=0.0269$). When correcting these associations for geographic distance (we did not have the genetic distances) by a weighted least squares approach, the correlations dropped considerably. This is no surprise as climate has a strong spatial pattern; removing this pattern also removes a large part of the likely adaptive signal.

Tests for neutral evolution

Betti and Manica (2018) reported a statistically significant correlation ($R^2=0.24$) between multivariate phenotypic distances (based on the three canal shape indices) and genome-wide genetic distances across the 24 human populations. This is in agreement with hypothesis H4, because the selective factors also show a spatial pattern. The explained variance between genetic and phenotypic distances was $R^2=0.24$, whereas the correlation between inlet shape and minimum temperature was considerably stronger ($r=0.61$, $R^2=0.37$), as we expected in our H5.

Of course, this does not preclude that certain aspects of pelvic form were influenced by neutral evolutionary processes. However, correlations between supposedly neutral genetic distances and multivariate phenotypic distances are not straightforward to interpret (e.g., Miller et al., 2008; Pujol et al., 2008), even in the absence of geographically patterned selective forces. For instance, their proportionality under a model of pure genetic drift rests on the assumption of equal heritability of all assessed traits, but Sharma (2002) showed that heritability differs among pelvic measurements. Moreover, in complex anatomical structures, such as the cranium or the pelvis, different aspects of the morphology can show very different functional and evolutionary dynamics (Caumul and Polly, 2005; Harvati and Weaver, 2006; Cardini and Elton, 2008; Smith, 2009; Bookstein and Mitteroecker, 2014; Grunstra et al., 2018; Grunstra et al., in press; Mitteroecker et al., 2020). Correlations between genetic distances and multivariate phenotypic distances lump all these different evolutionary signals together. The resulting lack of statistical power to identify a signature of selection should not be falsely interpreted as evidence for drift.

As mentioned above, quantitative genetic theory predicts that under pure genetic drift the quantitative genetic variance between population means is proportional to the quantitative genetic variance within the ancestral population. For multiple traits this translates into proportional genetic covariance matrices within and between populations (Lande, 1979). In the absence of reliable estimates of genetic covariance matrices, many studies have interpreted deviations between the corresponding phenotypic covariance matrices as indications of natural selection, but this rests on the assumption that the traits have similar heritability (e.g., Cheverud, 1988; Roff, 1995; Marroig et al., 2009; Bookstein and Mitteroecker, 2014; Grabowski and Rosenman, 2015). We plotted these ratios of between-population variance to pooled within-population variance separately for each pelvic variable in figure 1a. Under a pure drift model, they should all be equal, but they clearly are not. ML pelvic canal dimensions diverged more than twice as much as AP dimensions (even more than three times as much in the inlet). Similarly, pelvic inlet shape diverged more than twice as much as midplane shape. A multivariate test of proportionality of the within-population and between-population covariance matrices rejected the null hypothesis at $p<0.001$. Because of the numerous underlying assumptions and idealizations, these results should be interpreted with caution, but the data deviate strongly from the pattern expected under neutral evolution.

Another common statistical approach to study selection of quantitative traits is the comparison of neutral genetic divergence ($F_{ST}$) and quantitative genetic trait variation ($Q_{ST}$) across populations. In the standard quantitative genetic models, $Q_{ST}=F_{ST}$ under pure genetic drift, whereas $Q_{ST} > F_{ST}$ (i.e., more heritable phenotypic divergence than neutral genetic divergence) indicates divergent selection and $Q_{ST} < F_{ST}$ stabilizing selection (Leinonen et al., 2013; Walsh and Lynch 2018). In the absence of reliable estimates of additive genetic variance, researchers have also used a measure of phenotypic divergence, $P_{ST}$, as a rough proxy of $Q_{ST}$ (Leinonen et al., 2006; Saether et al., 2007; Brommer, 2011). We calculated $P_{ST}$ values for all the pelvic variables and the reference variables (figure 1b). Betti and Manica (2018) did not present overall $F_{ST}$ values, but they computed pairwise $F_{ST}$ values for the 24 populations, which we extracted from their figure (p. 6) and plotted as a histogram in figure 3a. They ranged from 0.002 to 0.256 with a mean of 0.105, well below the overall $P_{ST}$ values for inlet shape (0.165) and the three ML pelvic diameters (0.318, 0.200, 0.258). The histograms of pairwise $P_{ST}$ values for inlet shape and inlet ML show that the largest pairwise $P_{ST}$ values exceed the largest $F_{ST}$ values by factors of 2 and 4, respectively (figure 3b, c). This indicates that the mediolateral pelvic dimensions diverged considerably more between human populations than expected under genetic drift.
Serial founder effect

In our reanalysis, within-population variance of canal shape did indeed decrease with increasing distance from Africa, but the variance of the diameters themselves did not (figure 4). Note, however, that the relatively small number of specimens per population ($5 \leq N \leq 25$) imply large standard errors in these variance estimates. This decline in pelvic shape variance may indicate a loss of genetic variation resulting from a serial founder effect, as suggested by Betti and Manica (2018), but it could also result from reduced variance in other aspects of body form that correlate genetically or epigenetically with pelvic shape. Similarly, it could reflect local adaptations to the reduced variance in maternal or foetal body size (H7). As the geographic distances in figure 4 are distances from central sub-Saharan Africa, the reduction of pelvic shape variance far away from this part of Africa could also be due to stronger stabilizing selection in regions with a harsher, colder climate. Overall, it is difficult to interpret these results as they are not consistent between pelvic shape and form variables and because many evolutionary and environmental scenarios can explain them; they are not specific to drift.

Figure 3. (a) Histogram of the $F_{ST}$ values between all pairs of populations published by Betti and Manica (2018). (b) Histogram of pairwise $P_{ST}$ values for inlet shape (AP/ML). (c) Histogram of pairwise $P_{ST}$ values for the mediolateral diameter of the inlet.

(a) Histogramme des valeurs $F_{ST}$ entre chaque paire de population tel que publié par Betty et Manica (2018). (b) Histogramme des valeurs de $P_{ST}$ pour la conformation de la cavité pelvienne (AP/ML). (c) Histogramme des valeurs de $P_{ST}$ pour le diamètre médiolatéral de la cavité pelvienne

Figure 4. (a) Total variance of the six pelvic diameters in the populations measured by Betti and Manica (2018) versus geographic distance from central sub-Saharan Africa ($8^\circ S 25^\circ E$), estimated as the shortest distance on land avoiding long sea crossings and high mountains. (b) Total variance of the three canal shape indices versus geographic distance from central sub-Saharan Africa. The grey lines indicate the 90% confidence intervals of the population variances.

(a) Variance totale des six diamètres pelviens des différentes populations tels que mesurés par Betti et Manica (2018) comparé à la distance géographique de l’Afrique sub-saharienne centrale ($8^\circ S 25^\circ E$), estimé comme la distance la plus courte par terre en évitant les longues traversées de bras de mer et de chaînes montagneuses. (b) Variance totale des trois indices de conformation du canal pelvien comparé à la distance géographique de l’Afrique sub-saharienne centrale. Les lignes grises représentent les intervalles de confiance à 90% de la variance des populations.
Pelvic canal form and stature

Numerous earlier studies have reported that both pelvic size and shape show an allometric relationship with body size, where shorter women tend to have a relatively larger and more gynecoid pelvis with a round inlet, compared with taller women (e.g., Ince and Young, 1940; Bernard, 1952; Holland et al., 1982; Wood and Chamberlain, 1986; Tague, 2000; Kurki, 2007; Fischer and Mitteroecker, 2015; Fischer and Mitteroecker, 2017). Betti and Manica did not confirm these findings in their analysis. They fitted a linear regression to birth canal size and body mass (estimated from the acetabular diameter) and found that “residuals of the linear regression do not show any remaining effect of body mass on canal size; moreover, adding a quadratic and cubic term to the regression model does not improve the fit. The results, therefore, fail to support the hypothesis that smaller women have a larger than expected birth canal” (p. 5). This is not the standard approach to assess allometric relationships, and the residuals of a linear regression are, by definition, exactly uncorrelated with the predictor variable. We reanalysed their data by regressing log canal size on log body mass, which yielded an allometric exponent close to 1/3 (figure 5a). As body mass scales approximately with the square of body height (Heymsfield et al., 2007; Hood et al., 2019), this result translates into an allometric coefficient of ~ 2/3 for pelvic canal size on stature (i.e., negative allometry), which is close to those reported earlier (e.g., 0.7-0.8 in Fischer and Mitteroecker, 2017). In other words, also across the 24 populations measured by Betti and Manica (2018), shorter women tend to have a relatively larger birth canal than taller women.

Similarly, Betti and Manica (2018) did not find any association between canal shape and body mass. We regressed inlet shape on body mass (figure 5b), which yielded an AP/ML ratio just below 1.0 for the smallest individuals, along with a small negative slope for the 330 individuals and a more pronounced negative slope for the 24 population means (p=0.05 and p=0.02, respectively). We also found that the AP/ML ratios of the pelvic midplane and outlet show negative slopes across the population means (p=0.04, p=0.05). In other words, across populations, short women had, on average, a relatively rounder (gynecoid) pelvic canal, whereas taller women had a mediolaterally oval (platypelloid) pelvic canal. This differs from previous findings on within-population allometry, where tall individuals had an anteroposteriorly oval (anthropoid) pelvic canal. As outlined in H6, within-population correlations usually result from the variation of pleiotropic genetic and developmental factors, whereas between-population correlations can also result from correlational selection as well as independent selection pressures that co-occur geographically. The between-population association of body size and canal shape in these data was largely driven by the differences between populations living in hot versus cold climates (compare the labelled populations in figure 5b). Cold-adapted populations tend to have both a higher body mass as well as a wider pelvis than populations living in warm environments (Ruff, 1994; Jungers et al., 2016). Hence, the negative association of body mass and canal shape (AP/ML) across populations is likely to result from the adaptation of average stature and average body width to the thermal environment, whereas the positive within-population association may have evolved in response to obstetric selection.

Discussion

Neutral evolution of phenotypic traits is not easy to demonstrate, especially at an intra-specific or low taxonomic level, because of the highly idealised and non-exclusive null models. In practice, conclusions about neutral evolution are often a “diagnosis of exclusion,” reached by the failure
to identify signatures of selection. A correlation between neutral genetic distances and phenotypic distances can have multiple origins, such as geographically patterned selective factors in the biotic and abiotic environment, adaptive divergence, or neutral evolution. In most cases, however, past adaptive evolution can also be inferred only indirectly, if at all, because fitness advantages of certain trait expressions are difficult to prove for extinct species. Historically, the abundance of adaptive traits produced by natural selection has been overestimated by 20th century biologists. On the other hand, if certain heritable traits are correlated with survival or reproductive success, evolutionary change is inevitable when not inhibited by indirect antagonistic selection, developmental constraints or a lack of genetic variation.

Fitness consequences of variation in modern human pelvic traits are more straightforward to study than traits in non-human species thanks to the exhaustive medical literature. We have reviewed ample evidence that individual pelvic differences occurring today are still related to maternal and foetal mortality and morbidity, pelvic floor disorders and the ability to walk. Only in the last few decades has modern medicine ameliorated these effects in many, but far from all, parts of the world. We also showed that many of the selective factors (e.g., climatic conditions, maternal and foetal body size, pelvic floor function, diet, lifestyle, obstetric practices, health care) vary across populations, thus leading to different evolutionary “optima” in pelvic form. The evolution of quantitative traits occurs rapidly enough to have produced modern population differences in average pelvic form even under weak directional selection pressure. It is therefore likely that at least some, but not necessarily all, of the population differences in pelvic form are the result of natural selection. The correlations presented between pelvic width and temperature support this claim, as well as the high $P_{ST}$ values and the clear difference between the covariance patterns within and across populations. We showed that especially the mediolateral dimensions of the upper pelvis and the inlet differ substantially among human populations, more so than expected under pure genetic drift, and that they also clearly correlate with temperature. It remains to be shown that aspects of pelvic form also correlate with other biomedical and sociocultural factors, including maternal and foetal body size, incidences of cephalopelvic disproportion and pelvic floor disorders.

Pelvic dimensions are highly heritable but they are also affected by nutritional and environmental factors. At the individual level, for instance, undernutrition and malnutrition often lead to short stature and early menarche, which in turn can lead to a relatively small pelvis and narrow birth canal with an increased risk of cephalopelvic disproportion (Abitbol et al., 1997; Neilson et al., 2003; Wells, 2015; 2017; Zaffarini and Mitteroecker, 2019). Some of the population differences in pelvic form may thus be a consequence of phenotypic plasticity rather than evolved genetic differences. At the population level, however, phenotypic plasticity and selection go hand in hand: environmental differences between populations may induce differences in individual development (phenotypic plasticity) but they can also impose a selective gradient and thereby lead to heritable change. Phenotypic plasticity can stimulate evolution as a consequence of environmental change (e.g., Ghalmambor et al., 2007) and may even pave the way for genetic change, because environmentally induced phenotypes can sometimes become genetically encoded (Lande, 2009; Pigliucci et al., 2006; Schlichting and Wund, 2014). Hence, differences in nutrition, lifestyle, and environment between populations do not preclude evolved genetic differences in human body form. In fact, environmentally induced differences may act as triggers of evolutionary change and influence developmental trajectories as well as complex evolutionary trade-off dynamics.

Acknowledgments

This work was supported by an Elise Richter grant (V 826-B) and a Lise Meitner grant (M-2772-B) from the Austrian Science Fund (FWF) to Barbara Fischer and Ekaterina Stansfield, respectively.

References

Bergmann C (1848) Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Größe. Göttingen: Vandenhoeck und Ruprecht


Betti L, Manica A (2018b) Data from: Human variation in the shape of the birth canal is significant and geometrically structured, Dryad, Dataset [https://doi.org/10.5061/dryad.1gk3014]


Grunstra NDS, Zachos FE, Herdina AN et al (2019) Humans as inverted bats: A comparative approach to the obstetric...


