The morphological affinity of the early Pleistocene footprints from Happisburgh, England with other tracks of Pliocene, Pleistocene and Holocene age.

Abstract

Fossil hominin footprints provide a direct source of evidence of locomotor behavior and allow inference of other biological data such as anthropometrics. Many recent comparative analyses of hominin footprints have employed 3D analytical methods to assess their morphological affinities, comparing tracks from different locations and/or time periods. However, environmental conditions can sometimes preclude 3D digital capture, as was the case at Happisburgh (England) in 2013. Consequently, we use here a 2D geometric morphometric approach to investigate the evolutionary context of the Happisburgh tracks. The comparative sample of hominin tracks comes from eight localities that span a broad temporal range from the Pliocene to late Holocene.

Results show disparity in the shapes of tracks ascribed to hominins from the Pliocene (presumably *Australopithecus afarensis*), Pleistocene (presumably *Homo erectus* and *Homo antecessor*) and Holocene (*Homo sapiens*). Three distinct morphological differences are apparent between time samples: changes in adduction of the hallux, changes in the shape and position of the medial longitudinal arch impression, and apparent changes in foot proportions. Linear dimensions classified the potential *Homo antecessor* tracks from Happisburgh as being most similar to the presumed *Homo erectus* prints from Ileret.

We demonstrate using 2D geometric morphometric methods and linear dimensions that the Happisburgh tracks are morphologically similar to other presumed *Homo* tracks, and differ from the Laetoli footprints. The probable functional implications of these results fit well with previous comparative analyses of hominin tracks at other sites.

Keywords

Hominins, fossilised footprints, geometric morphometrics, foot anatomy, functional morphology.

Abbreviations.

Ma – millions of years ago. Ka – thousands of years ago. AMH – anatomically modern humans.

1. Introduction

Fossil hominin tracks are known from the Pliocene, Pleistocene and Holocene (Bennett and Morse, 2014) and more contentiously from the Miocene (Gierlinski et al., 2017; Crompton, 2017; Meldrum and Sarmiento, 2018), and can provide evidence of locomotor behavior, and offer inference of other biological data including anthropometrics (Webb, 2007; Webb et al., 2007; Tuttle, 2008; Vaughan et al., 2008; Bennett et al., 2009; D'Août et al., 2010; Crompton et al., 2012; Morse et al., 2013; Bennett and Morse, 2014; Masao et al., 2016; Hatala et al., 2016a; Hatala et al., 2016c; Bennett et al., 2016a; Raichlen and Gordon, 2017). The development of 3D modelling for fossil tracks has been pivotal in pioneering a revolution in the study of such tracks (Remondino et al., 2010; Falkingham, 2012; Bennett et al., 2016b; Falkingham et al., 2018), permitting reconstructions of behavior, kinematics and body size metrics from the shapes and dimensions of fossil hominin tracks (e.g., Hatala et al., 2016b; Raichlen and Gordon, 2017). Digitization has advanced scientific research while simultaneously enhancing the flexibility of analyses and availability of data to numerous research teams (Belvedere et al., 2011; Falkingham, 2012; Falkingham et al., 2018). The advantages of digital data are particularly pertinent for fossil track sites where excavation can be damaging and where tracks are susceptible to erosional processes (Bates et al., 2008; Wiseman and De Groote, 2018; Zimmer et al., 2018). However, the digital 3D capture of tracks can be challenging in certain environmental conditions, especially where tracks are exposed for only a brief period (Wiseman and De Groote, 2018).

This was the case at Happisburgh, England (Fig.1) where high quality 3D data could not unfortunately be captured before the fossil tracks were destroyed by marine erosion in May 2013 just two weeks after exposure/discovery (Ashton et al., 2014). Marine erosion at Happisburgh exposed a sediment bed dated to 950-850 Ka that contained 152 small (c.50 mm-320 mm) hollows, 49 of which were identified as potentially hominin tracks tentatively ascribed to *Homo antecessor*. Of these, only 12 were included in the original analyses due to the severe erosion of many of the prints (Ashton et al., 2014). No tracks could be associated as belonging to a common trackway; rather, the sediment bed is a mixture of singular prints.

The prints were recorded using a handheld DSLR camera with the intention of creating photogrammetric models, yet 3D reconstructions were later deemed to be of a low resolution. The likely cause was the wetness of the fossil bed. The prints were rapidly infilled with water due to poor weather conditions during data capture. Water is a reflective material which impedes photogrammetric reconstruction. This resulted in sparsely reconstructed 3D models (i.e., the track outlines were well-defined, but the internal features of all prints were not

captured). Poor weather conditions combined with marine erosion caused the bed to be destroyed in just two weeks. Consequently, high quality 3D data was not captured prior to the loss of the prints (Ashton et al., 2014). This has led to the necessary exclusion of these tracks from many of the recent studies that have applied 3D analyses (e.g., Hatala et al., 2016b; Bennett et al., 2016a).

It is no longer possible to re-capture the Happisburgh prints in 3D, meaning that we must now work with the available 2D data. The loss of the third dimension in the Happisburgh tracks is problematic because it potentially limits the information that can be gained from such an important set of fossils. Tracks are representative of the dynamic motion of the foot and the way that the foot has interacted with the underlying substrate – concepts which are preserved three-dimensionally (e.g., Falkingham and Gatesy, 2017). With the loss of the third dimension it has been necessary to identify another methodology to quantitatively and/or qualitatively analyse these important fossils solely from 2D images.

It has been over 170 years since the first publication that used 2D methods to comparatively assess different sets of tracks belonging to different species (Hitchcock 1858), and the use of 2D approaches still continues today (e.g., Costa-Perez et al., 2019; Duveau et al., 2019). Simple 2D linear measurements – while informative – have also been combined in recent years with applications of 2D geometric morphometrics (e.g., Bennett et al., 2009). This shape analysis approach uses a landmark-based identification of homologous anatomically/geographically-defined points to statistically compare the outline shapes of tracks found at different sites. This use of geometric morphometrics has been reliably used in ichnotaxonomic classifications where intra-species variation is minimal (i.e., where little size differences exist between sexes). However, this method is less reliable if stark differences in speed and substrate are present (Costa-Perez et al., 2019). If a selection of tracks belonging to different species are collected and speed estimates and substrate are found to be comparable, a 2D geometric morphometric approach can be reliably implemented (Costa-Perez, 2019). However, the effect of substrate on print morphology must be carefully considered. Differences in substrate typologies can cause similar anatomical and kinematic patterns to generate different print shapes (Bates et al., 2013b; Morse et al., 2013). In general, these differences manifest in the three-dimensional topologies of footprints, which leads to the question: if we remove depth from comparative assessments, can we partially circumvent the "substrate-effect" issue? If so, this would permit the first comparative assessment of the Happisburgh tracks with other tracks belonging to Pliocene, Pleistocene and Holocene hominins.

One final set of considerations concerns the quality of the 2D data from Happisburgh. The tracks have no discernible internal features, except for print-8 (Ashton et al., 2014) (Fig.1a). Our approach must then focus solely on quantifying external features, specifically the footprint outlines. While defining the outline of a track can be somewhat subjective (Falkingham, 2016; Falkingham and Gatesy, 2017), a standard protocol for landmark identification can minimize the effect of subjectivity. By developing such a protocol, 2D methods can be used for the quantitative comparative assessment of track morphology.

Here we evaluate the Happisburgh tracks in a broader comparative context with other hominin tracks produced at similar speeds (but slightly different substrates) by applying a 2D geometric morphometric approach based on track photographs. This builds on the work of Berge et al. (2006), Bennett et al. (2009) and, more recently, Duveau et al. (2019), who also used 2D geometric morphometric approaches in comparative analyses of hominin footprints. This approach has previously been demonstrated to reflect differences in movement patterns and biometrics (e.g., body size) between species (e.g., Costa-Perez, 2019). Most importantly, this approach can assess the affinities of the Happisburgh tracks to other assemblages of footprints ascribed to the genus *Homo*. While this approach may not identify great disparity between taxa in the same way that we might find significant morphological disparity between the foot bones of the respective track makers from different sites, 2D geometric morphometric analyses of track shapes will enable between-site comparisons that are functionally and evolutionarily meaningful. For example, we would expect the Happisburgh prints to have a relatively adducted hallux in comparison to geologically older prints (i.e., the Laetoli, Tanzania prints), because an adducted hallux in the foot is one of the defining characteristics of efficient bipedality in *Homo* (e.g., Harcourt Smith et al., 2004).

In the present study, we aim to: 1) compare the 2D morphologies of the Happisburgh tracks with Pliocene, Pleistocene and Holocene tracks; and 2) evaluate the results of comparative analyses in functional and evolutionary contexts.

2. Materials and methods

2.1 Data acquisition

To compare the morphologies of the Happisburgh tracks with those of other hominin tracks, 2D data were collected from sites ranging from the Pliocene to the late Holocene (Table 1; Supplementary Online Material (SOM) S1). A number of footprints were excluded from this

dataset. Reasons for exclusion included: camera parallax issues during data capture, walking speed, poor outline definition, and/or substrate typology as discussed below.

Orthogonal photographs were collected from published or archival records, or taken directly by the authors (Fig. 2). Images were inspected for viewing angle, to ensure that the print was centred in the image and that camera distance was sufficient to avoid parallax distortion (i.e., the full print with the displacement rim had to be visible in each photograph alongside a small – not measured – border of surrounding substrate. If the photograph did not meet these criteria, the photograph was excluded). This precaution may not be necessary since Mullin and Taylor (2002) have shown that slight distortions in images are not always a problem in most geometric morphometric analyses. Despite this, we took a conservative view and excluded images that were not orthogonal or potentially suffered from parallax. In the case of tracks for which 3D data are available, an orthogonal image was created of the track and exported as a 2D image in MeshLab (Gignoni et al., 2008). Belvedere et al (2016) have demonstrated that only 3% measurement disparity exists between dimensions extracted from 2D images and those from 3D models. It is a reasonable assumption that non-significant variability will exist when extracting 2D data from 3D models.

Variation in speed of locomotion can induce changes to track shape that introduce confounding errors when comparing tracks belonging to different species (Costa-Perez et al., 2019). For tracks where speed estimates were possible, only those created at "walking speed" (classed as speeds below 1.5 m/s) were included in this study to minimise the effects of speed as a potential confounding factor. Qualitative categorization was based upon the gait classifications of Jordan and Newell (2008), whereby any speed above ~1.6m/s in humans is classed as a fast-paced walk and speeds above ~1.9m/s are classed as running. Speed was calculated using the method developed by Dingwall and colleagues (2013). Published stride and foot length values were used to calculate speeds for the Walvis Bay tracks (Morse et al., 2013). Stride and foot lengths were measured from the Formby Point footprints by AW in 2016/17. Speed was not calculated for the Happisburgh tracks as associating singular tracks into trackways was confounded by a mix of superimposed tracks in the sediment bed (Ashton et al., 2014). Published speed estimates were used for Laetoli Site G and Site S tracks (Masao et al., 2016) and for the Ileret sample (Dingwall et al., 2013). All tracks used in the current study are listed in SOM S1. Although it is acknowledged that based on the principal of dynamic similarity (Alexander and Jayes, 1983), step frequency will be higher in shorter hominins (e.g., *Australopithecus afarensis*) walking at similar speeds to taller hominins (e.g., *Homo sapiens*), the application of

the speed cut-off criterion for all fossil tracks is justifiable because tracks belonging to all shorter individuals from Laetoli and AMH sites (i.e., juvenile tracks) were travelling between 0.44-1.1m/s. We can expect that shorter individuals with a higher step frequency would transition to a running speed earlier than taller individuals, but such a low speed of 0.44-1.1m/s in shorter individuals does indeed represent walking behavior.

Only tracks which had clearly defined outlines were collected for this study, following the convention set by Marchetti et al. (2019) who described such tracks as those where it is still possible to discern distal toepad impressions, medial foot impressions etc. Tracks lacking clear outlines were excluded. In most cases, this involved omitting particularly deep tracks. Track morphology has been demonstrated to be influenced by substrate typology, and wide variation in track depths typically signals the existence of such substrate effects (Bates et al., 2013b; Morse et al., 2013). Bates et al. (2013b) noted that substrate effects were noticeably larger for tracks >20 mm deep. This threshold was therefore applied in the current study. We believe that omitting deep tracks (>20 mm deep) will help to constrain intra-group substrate-based variability and amplify the power of cross-site comparisons. Published print depths were used as the cut-off criterion for inclusion (Raichlen et al., 2010; Hatala et al., 2016). Depth was measured directly for the Walvis Bay and Formby Point footprints by fitting a plane to 3D models in CloudCompare and measuring the absolute depth of each print. Only the G1 trackway from Laetoli Site G was used in our study. We excluded the G2/3 tracks as the overlay/trampling of these tracks would probably introduce noise error within the Laetoli sample. Finally, if homologous landmarks could not be identified on a given track (Section 2.3), that track was also excluded from comparative analyses.

Across all sites a total sample of 274 footprints was identified that provided well-preserved track outlines from which measurements and defined homologous geometric landmarks could be identified. Only a small group of tracks were usable from the geologically oldest sites: Laetoli, Ileret and Happisburgh. Most of the sample (n=218) belongs to AMHs. For each footprint in the sample (except for the Laetoli prints), track-maker age was estimated using modern growth curves of the foot derived from the World Health Organisation (de Onis, 2006) as employed by Ashton et al. (2014) and by Altamura et al. (2018). Three classifications were created: an adult track was determined if footprint length exceeded 19 cm (following World Health Organisation protocol; see: de Onis, 2006), while prints shorter than this threshold were assigned to juveniles. Tracks which had an age prediction of 17-19 years old were assigned to sub-adult.

2.2 Linear footprint metrics

To test whether track dimensions differ between samples, four linear measurements of each track were taken in TPSDig 2.0 (Rohlf, 2004): the most distal point of the hallux to the most proximal point of the heel (henceforth, track length); the distal tip of the second digit to the most proximal point of the heel (henceforth, long axis of the track); forefoot breadth at the widest breadth; and heel breadth at the widest breadth whilst passing through the centre point of the heel (Fig.3). Hallux length was calculated for each track as the distance from the most distal point of the hallux (i.e., the most concave point) to the ridge between the hallux impression and the forefoot impression. Ray II length was measured as the most concave, distal point of the second digit impression to the tip of the heel. Track length was used to predict stature using regression equations published by Dingwall et al. (2013). The angle of hallux abduction was also measured for each track, as the angle between the long axis of the track and an intersecting line crossing from the tip of the hallux impression through the center-most point of the hallux impression (Bennett et al., 2009).

Replicability tests were computed to test observer error via assessing the reliability of measuring linear measurements from tracks (SOM S2). Using eight randomly selected tracks from each fossil location, metrics were repeatedly measured over a 10-day period by the same individual. The mean standard error of all measurements was determined to be <1.92%, well below the standard 5% measurement error accepted in biological sciences.

A stepwise Discriminant Function Analysis (DFA) using a leave one out classification to control for uneven sample sizes (Huberty, 1994; Lance et al., 2000) was computed on all track dimensions and hallucal angles to establish the probability of classification of tracks into the assumed species attribution for each site. Only the adult specimens were incorporated in the DFA (to exclude ontogeny as a potential factor driving statistical variance), with groups corresponding to the assumed species.

To test if track proportions (i.e., presumed foot proportions) changed from the Pliocene to the Holocene, we calculated the total lengths of the impressions for the hallux and the length of the impression of the ray II in each track (Fig.3). The proportion of forefoot length (i.e., hallucal/ray II length; both ratios were calculated for all tracks) to total track length was also calculated for each track. This allowed us to estimate the internal proportions of the foot that produced each track. Because some samples within our dataset included juvenile tracks (Happisburgh, Formby Point, Walvis Bay) and it is known that foot proportions change during

ontogeny (e.g., Davenport, 1932), tracks attributed to juveniles were excluded from these analyses.

2.3 Geometric morphometric analyses

We tested for changes in outline shape between groups (Pliocene, Pleistocene and Holocene adult and juvenile samples) by applying geometric morphometric methods (Bookstein, 1991; Slice, 2005). All tracks within a trackway belonging to a single individual were included in these analyses. Because there are potential or actual multiple tracks per individual, some individuals will be more heavily weighted in statistical assessments than others (i.e., variability can be assumed to increase within groups if the same individual is represented by more than one print). Consequently, all statistical analyses have incorporated 'trackway' (i.e., all tracks pertaining to a singular trackway) as a random effect to address this issue directly.

Reliability tests of landmark placement were conducted to ensure that landmarks could be consistently identified within and across samples. Landmarks were placed over a period of ten days by the same researcher on three randomly selected tracks: one track each from Laetoli, Happisburgh and Formby Point. Landmark reliability tests consisted of a Generalised Procrustes Analysis (GPA) computed in R (R Core Team, 2017) to test for consistency in landmark digitization (Slice, 2005). The resulting Procrustes distances between each landmark consensus with the mean landmark configuration were calculated and then divided by the number of repeats (Slice, 2005; Zelditch et al., 2012). This process provided the error estimate (Type I error rate of 5%) for landmark placement within a 95% confidence interval. Mean values (Procrustes distances) over 0.05 specified that the distance between a landmark and the overall consensus was high and that the landmark is non-replicable (Profico et al., 2017). All mean values lower than 0.05 indicated good repeatability in landmark placement. Landmarks were placed at locations that were selected according to feasibility and likely repeatability of placement. While we acknowledge that some landmarks were less clearly defined in some prints because of differences in preservation, we do stress that all landmarks were found to be homologous between each repeat, permitting the following assessments to be conducted.

The mean Procrustes distance from the consensus was 0.03±0.01. This signifies that intra-observer error in repeatability of landmark placement was low, and that the landmark configuration is suitable for the subsequent analyses. This process resulted in the selection of 16 type II landmarks that all had a Procrustes distance <0.05. These landmarks were digitized on 270 prints (excluding Terra Amata, Vartop Cave and Langebaan tracks due to small sample

sizes) using TPSDig 2.0 (Rohlf, 2004) (Fig. 4). To circumvent the issue of asymmetry, all left landmark configurations were mirrored (Dryden and Mardia, 1998; Mardia et al., *20*00).

Landmark configurations were superimposed using a GPA (Gower, 1975). Shape variation was assessed using a between-groups Principal Components Analysis (bgPCA). This methodology allows the number of variables to be higher than the number of observations (Mitteroecker and Bookstein, 2011), which was particularly relevant for comparative analyses of the Laetoli, Ileret and Happisburgh samples. A nested MANOVA with mixed effects was computed on the resulting shape scores using trackway number (SOM S1) as a random effect, and age and fossil location as fixed effects to determine the statistical significance of morphological variation among fossil localities and across time. Analyses were computed in the geomorph (Adams and Otárola-Castillo, 2013) and morpho (Schlager, 2017) R packages (R Core Team, 2017).

3. Results

3.1 Linear footprint measurements

To evaluate changes in foot/track size from the Pliocene to the Holocene, four linear length and breadth measurements were computed and compared (SOM S3). Results from the one-way ANOVA and Games-Howell post-hoc tests demonstrated that track lengths and lengths of the long axes significantly increased from the Pliocene to the Pleistocene, despite high variation within the Laetoli sample as was revealed in other recent analyses (Masao et al., 2016). Broad similarity in track lengths was established between *Homo* species (SOM S4), consistent with previous comparative assessments (Kim et al., 2008). Forefoot and heel breadth were found to remain static across hominin prints from the Pliocene to the Holocene, except for variability in heel breadth dimensions between Holocene populations. Because track lengths increased between the Pliocene to the early Pleistocene samples, so did stature predictions (SOM S3). Comparisons of hallux abduction angles revealed a trend for a significant reduction in hallucal abduction ($P<=0.001$, $F=275.563$ between all groups) from the Pliocene to the Holocene (Table 2; Fig. 5).

Using all track dimensions, the range of presumed species assignment (see: Table 1) obtained from a DFA was between 20.0% (*H. antecessor* from Happisburgh) and 98.1% (*H. sapiens* from Formby and Walvis Bay) (Fig. 6). Tracks assigned to *Homo antecessor* were mostly (60.0%) classified as belonging to *H. erectus*, signifying that these two groups closely overlap in track dimensions. 20.0% of tracks from Happisburgh were incorrectly classified as *H. sapiens.* Only 20.0% were classified as *H. antecessor.* The first function was highly correlated

with hallucal angle (R^2=0.971) (the measurement which achieved the greatest discrimination between assumed species); the second function was correlated with track length (R^2=0.909) and forefoot (R^2=0.392) width; and the third function was driven by heel breadth (R^2=0.999) (Table 3). 83.3% of *A. afarensis* were correctly classified and 75.0% of *H. erectus* were correctly classified

To explore comparative foot proportions between tracks, digit lengths (henceforth referred to as hallux length) were calculated for each track as the distance from the most distal point of the hallux to the ridge between the toe impressions and the forefoot (Fig. 3) and then the ratio of distal track to total track length was calculated as a means of estimating the relative length of the load arm used for toe-off. Results indicate a 30.15% mean reduction in relative length of the hallux between the Laetoli and Ileret hominins (Table 4). There was a 4.4% reduction in hallux length established between the Ileret and Happisburgh individuals. Hallux length changed by -4.7-2.6%between the Happisburgh individuals and AMHs.

Synchronous with a reduction in the length of the distal foot, it was determined that the ratio of toe lengths (second digit) to total track length decreased from the Pliocene to the early Pleistocene (Table 5; Fig. 7). The second digit to total track length ratio was found to reduce as much as 26.2%. The ratio of toe length to total track length experienced very little variability thereafter, with miniscule changes being the probable result of the interactions of the foot with the underlying substrate, rather than reflecting changes to the foot's lever mechanics. The mean percentage of digit length to track length is found to be within modern human ranges (Keith, 1929) from the early Pleistocene, resulting in modern human-like foot proportions from the first appearance of trackways attributable to the genus *Homo*. The Laetoli toe lengths (44.4 ± 10.3 mm) were found to be within published skeletal estimates for the most frequently inferred track-maker, *A. afarensis* (49.4 mm) (Rolian et al., 2009).

3.2 Geometric morphometric results

To test the prediction that track shape varies between fossil localities, GM methods were applied on landmark configurations that synthesise the outline shapes of fossil prints. A PCA was performed using Procrustes-fitted landmarks across all samples of hominin tracks (Fig. 8). All categorical variables were treated as independent observations (e.g., different inferred species and the inclusion of several substrates) to identify which factor(s) explains the majority of shape change.

Variation along PC1 was characterized by a separation of negative PC scores for the Laetoli tracks and positive PC scores for the Ileret tracks. Positive and negative scores exist for all other hominin track samples. Variation between fossil localities explains 11.74% of the total variance in track outline shapes (P<=0.001, F=8.255), as determined by a MANOVA. Multiple other factors could explain variation in PC scores. For example, each site includes a different mixture of tracks produced by juvenile and adult individuals. However, relative age (e.g., juvenile or adult) of the track-maker explained just 1.78% of total shape variability (P=0.002, F=2.503) (Table 6). Further confounding variables are discussed below in section 3.3.

Variation along PC1 was visualised as shape deformation graphs within the morphospace (Bookstein, 1989). Shape change while moving positively along PC1 can be explained by three variables: increasing adduction of the hallux, the anteroposterior displacement of the medial longitudinal arch (MLA) and a reduction in heel width (Fig. 8). On the other hand, variation along PC2 seems related to the prominence of the MLA impression.

The axis of PC3 appears to highlight the morphological disparity between AMHs (majority distributed as PC3+ scores) and all other hominins (PC3- scores) (SOM Fig. S1). Shape change along PC3 can be explained by the prominence of the MLA impression, with PC4 explaining once more the change in the MLA but also hallucal adduction. Evidently, changes in the midfoot region accounts for much of the shape variance present within this sample (PC1 to PC11; 87.24%).

3.3 Confounding variables

Evolutionary differences are likely to be subtle and therefore potentially swamped by other variables which determine footprint outline, namely differences in the age of the track-maker, differences in walking speeds and substrate properties. In order to understand the contribution of such variables a number of additional analyses were performed.

<u>The effect of speed on track outline</u>

Dingwall et al. (2013) and McClymont et al. (2016) have both indicated that track topology is influenced by speed. Although tracks above a walking speed of 1.5 m/s were excluded from the original sample, speed remains a potential variable. A sub-sample of 137 tracks was used for this analysis with data from Laetoli, Ileret, Formby Point and Namibia being included. To determine if track morphology was affected by walking speed (m/s) across the sample, speed

was introduced as a covariate and a MANOVA that accounted for 100% of shape variance was computed to establish the relative influence of speed alongside fossil locality and track-maker age. In this analysis, the effect of speed was statistically significant (P=0.010; F=8.191) (Table 7). The effect of speed on outline shape explained 17.50% of total shape variance within this sample, whereas the "locality effect" explained only 15.21% of the total shape variance. Speed thus had a greater effect on track outline shapes than the inferred anatomical differences between the track-makers.

Potential ontogenetic effects on track shape variation

To determine if track shape variation between fossil samples could be affected by ontogenetic variation, track size (log-centroid size; henceforth log-CS) was introduced as a variable for the samples from which juvenile tracks were available (Pleistocene and Holocene samples). A MANOVA was computed between groups using all PC scores (describing 100% of shape variance) as response variables and with log-CS, track-maker age, and fossil locality as explanatory variables. Differences between juvenile and adult tracks within each fossil locality were found to be statistically significant (P=0.002; z=6.238 between the Formby Point juvenile and adult tracks. P=0.002; z=2.859 between the Walvis Bay juvenile and adult tracks. P=0.032; z=2.368 between the Happisburgh juvenile and adult tracks) (SOM S5). The contrasts in the z values reported here (grouped: P<=0.001; z>=2) have demonstrated that the greatest morphological disparities revealed by the GM analyses separate the juvenile tracks from all adult specimens (Holocene and Pleistocene samples).

Pairwise comparisons of log-CS to shape (PC scores) were computed using only the adult tracks from the Pliocene, Pleistocene and Holocene. Results indicated that there are no significant differences for this comparison between the adult tracks from the Pliocene, Pleistocene or Holocene (P>=0.05; z>=1 between all groups, within a 95% confidence interval). This suggests that the relationship between track size and shape remained similar between hominin adult groups, despite eco-geographical and temporal differences, and variability in substrate typologies. Alongside these differences, there was a wide range of variations in the anatomies of australopith and *Homo* feet (e.g., Aiello and Dean, 2002; De Silva et al., 2018), so it is quite surprising to find such similarity between the tracks. Alternatively, the lack of apparent differences in track morphologies could be due to the stark contrast in sample sizes (Cohen, 1988; Collyer et al., 2015), as geologically older samples (e.g.,

in the Ileret and Happisburgh samples) are much smaller than Holocene samples (Walvis Bay and Formby Point).

Because shape variance was dominated by the presence of juvenile prints in the dataset, an additional PCA and MANOVA using only the adult specimens (now characterised as dependent observations) were computed, so as to reduce the number of confounding variables (Table 8). The results of the PCA indicate that there was broad similarity between all tracks. Speed explained 17.11% of the total variance (P=0.001) in outline shape. Fossil locality (and therefore their eco-geographical and temporal properties) explained 16.12% of the total variance in the adult tracks, although an overlay of Procrustes scores makes it difficult to clearly distinguish shape differences between different inferred species. The "locality effect" was higher here, indicating that between-site variations are more apparent in the adult-only sample.

The effect of substrate on track shapes

Although particularly deep tracks were excluded from these analyses, we took a conservative approach and examined the extent to which substrate may influence the variations observed in the outline shapes of tracks. A PCA and a MANOVA were computed on the two Holocene track samples from Formby Point and Walvis Bay which were produced on different substrates with speed introduced as a random effect. The PCA results demonstrate a mixture of Holocene PC- and PC+ scores (R^2=0.016; F=3.121; P=0.005), indicating that differences in substrate material properties only accounted for 1.61% of morphological variation. Rather, other factors, such as biometric variation, are likely to have greater influence on the variance of track outline shapes.

To test the effect of substrate on fossil track shapes composed in a larger variety of sediments (e.g., natrocarbonatite ash and sandy deposits), a final PCA and MANOVA were computed using track samples which represent the deeper and shallower ends of the spectrum (Section 2.1; Fig. 9). Results were found to be similar to the PCA inclusive of all track data (Fig. 8): the geologically oldest tracks show little intra-group variability along PC1, represented by strong negative characterisation along PC1 in both the deep tracks (R^2=0.123; F=4.836; P<=0.001) and the shallow tracks (R^2=0.108; F=8.396; P<=0.001). The Holocene tracks have a mix of PC scores, with a broad overlap with the Happisburgh scores. Differences in locality (inferred species) account for 70.27% of the total variance in the deep tracks and 76.34% for shallow tracks. This signifies that most of the shape variation is influenced by the track-maker and not by track depth. Some consideration should still be given to substrate as despite depth being

non-influential, this study sampled seven different substrate typologies which will likely introduce some error into these analyses (i.e., between Holocene samples, the influence of substrate was 1.61%).

4. Discussion

4.1 Disparities and affinities in hominin track shapes

Differences in track shapes were identified between the geologically oldest Pliocene tracks (Laetoli) and Pleistocene tracks ascribed to *Homo* species, indicating that there may be differences in form and function between genera. Although some morphological differences were established between tracks assumed to have been created by australopithecines and *Homo* species as indicated weakly by the PCA, no shape differences were identifiable between *Homo* groups. Given the anatomical differences in *Homo* feet (Aiello and Dean, 2002; De Silva et al., 2018) this result is perhaps surprising, but could reflect a lack of functional differences despite subtle skeletal anatomical differences.

In contrast, a DFA was able to correctly classify the majority of track ascribed to *H. erectus* (from Ileret) (75%) and *H. sapiens* (from Formby and Walvis Bay) (98.1%), suggesting that track dimensions may still be useful for ichno-taxonomy within the genus *Homo*. Importantly, the Happisburgh prints were similar to *H. erectus,* with some sharing closer affinities with *H. sapiens.* This aligns with previous assumptions that the Happisburgh tracks belong to the genus *Homo* and are in a sense morphologically intermediate between prints assigned to *H. erectus* and *H. sapiens.* This is consistent with their age and inferred attribution to *H. antecessor* (Ashton et al., 2014).

4.2 Trends in track morphology inferred from comparative analyses

As a result of previous studies (e.g., Meldrum et al., 2011; Crompton et al., 2013; Hatala et al., 2016; Bennett et al., 2016) one might suspect that midfoot impressions should vary in tracks from the Pliocene to the Holocene as the medial longitudinal arch was more prominent in certain later hominins. This morphological change is hypothesized to have occurred in conjunction with a more adducted hallux. This study suggests that from the Pliocene to the late Holocene, hallux adduction increased (Fig. 5) while the MLA became more prominent (inferred from TPS grids – Fig. 8), coinciding with the hallux becoming more adducted. This coincides with bony configurations from the ascribed *Homo* species (e.g., Harcourt-Smith and Aiello, 2004) suggesting that foot proportions were within modern human ranges in *Homo,* and

outside those of australopiths. Assuming these anatomical specifications reflect functional capabilities (Harcourt-Smith and Aiello, 2004; Sellers et al., 2005; Bates et al., 2013a; Holowka et al., 2017), these results hint at possible functional differences between the feet of the Laetoli track-makers and *Homo* track-makers, as have been proposed elsewhere (e.g., Bennett et al., 2009; Hatala et al., 2016a). However, it should be noted that the extent to which tracks reflect longitudinal arch morphology might be highly dependent on substrate properties (e.g., Meldrum, 2004; Bennett et al., 2016a; Hatala et al., 2018), that this region of the foot can also deform during locomotion (Bates et al., 2013a; Pataky et al., 2013; McClymont et al., 2016), and that the longitudinal arch deforms differently across different substrates (Hatala et al., 2018).

Differences in hallucal abduction were more readily apparent across fossil track samples. The DFA actually identified that the majority of species classification is driven by hallucal abduction. This finding fits well with previous assumptions that the hallux became more adducted in the genus *Homo* and strongly differs between genera (e.g., Aiello and Dean, 2002; Proctor et al., 2008; Bennett et al., 2009).

Additionally, one might expect that foot proportions will vary between hominin track samples of different geological ages, which may imply different patterns in foot function across the taxa. In the modern human foot, the distal foot constitutes ~18% of the total foot length, whereas in chimpanzees (a habitual quadruped) the distal foot accounts for ~35% of total foot length (Keith, 1929; Aiello and Dean, 2002). By having a smaller ratio of phalanx to foot length, humans shorten the load arm at the metatarsophalangeal joints and, therefore, decrease energy expenditure during locomotion whilst increasing the mechanical efficiency of foot propulsion (i.e., toe-off in later stance) during bipedality relative to chimpanzees. The current work shows that relative toe lengths were found to be within modern human ranges for all Pleistocene and Holocene tracks corresponding to bone lengths in the ascribed species (e.g., see: Aiello and Dean, 2002). We can therefore assume that foot proportions and bony configurations as inferred from the footprints in the ascribed *Homo* track-makers were similar to modern human foot anatomy. The Laetoli tracks, on the other hand, are characterised by relatively longer toe impressions. Changes in foot proportions suggest that the Pleistocene and Holocene hominins sampled here may have been better suited for bipedal locomotor efficiency than the Laetoli hominins, at least during running (Rolian et al., 2009).

4.3 Morphological affinity of the Happisburgh tracks

The Happisburgh tracks were found to share closer affinities to Pleistocene and Holocene groups than to Pliocene tracks. No tracks from Happisburgh were incorrectly classified as belonging to the presumed australopith grouping, suggesting that linear dimensions may be more suitable for inferring genus disparity between footprints than a landmark-based approach. The classification scores as determined within this study indicate that the prints from Happisburgh are most similar to *H. erectus* prints. The results are consistent with predictions that early Pleistocene *Homo* species share anatomical affinities.

To this end, it is possible that this result in some way reflects that locomotor activity has probably remained relatively consistent within the genus *Homo* since the Pleistocene. However, inference on biomechanical affinity/disparity between-groups should be cautious, as extracting biomechanical data from track morphology has previously been demonstrated to be complicated (D'Août et al., 2010; Bates et al., 2013b; Hatala et al., *20*13; Pataky et al., 2013). Further exploration into the complex relationships between foot motion and substrate mechanics is necessary before drawing comprehensive functional conclusions about fossil tracks (Hatala et al., 2018). We can draw confident conclusions about similarities or differences in track morphologies between fossil sites, but linking these comparisons to biomechanical conclusions will require further research.

4.4 Limitations of substrate

The results presented here should be interpreted with some caution since the dataset comprises fossil tracks generated on an array of different substrates. These substrates range from fluvial-lacustrine at Ileret to natrocarbonatite ash at Laetoli and, consequently, vary in their material properties including their lithology and heterogeneity. Variability in material properties impacts the mechanics of substrate deformation when a foot strikes the ground and, subsequently, the morphology of the print that is left behind (Morse et al., 2013; Bennett and Morse, 2014; Hatala et al., 2018; Costa-Perez et al., 2019). Most sites incorporated in this study (with the exception of the Laetoli trackways) were created in similarly soft substrates, based on qualitative between-site comparisons of trackway depths and topographies. Deeply deformed tracks associated with soft substrates were also excluded from the sample. Deformation primarily impacts of 3D topology of tracks rather than on their 2D outlines (Bates et al., 2013b; Morse et al., 2013), suggesting that our cautious approach could mitigate the impacts of substrate on our results. However, if 3D analyses of the Happisburgh tracks had

been possible, we would have been afforded more analytical power to assess the potential influences of substrate variation on the 2D comparisons made here.

5. Conclusion

The dataset used within the current study includes hominin trackways that have been attributed to six distinct hominin species within two genera, spanning from the Pliocene to the Holocene. Even across such a broad sample of time and space, some aspects of track morphology are found to be remarkably consistent. However, between-sample differences were identified in three morphological aspects of the tracks. These differences are related to the prominence and position of the medial midfoot impression, the abduction angle of the hallux impression, and the length of the forefoot relative to the rest of the track. Generally, comparing sites across time from the Pliocene to the Holocene, the MLA is more prominent, the hallux is less abducted (this variable achieved the greatest discrimination between assumed species), and the forefoot is relatively shorter in more recent track samples. The linear dimensions classified the potential *H. antecessor* tracks from Happisburgh as being most similar to the *H. erectus* prints from Ileret, suggesting the dimensions and shape of Pleistocene tracks were likely similar.

Importantly, this is the first study to specifically examine the morphology of the Happisburgh tracks within such a broad comparative context. The Happisburgh tracks are found to be morphologically similar to other early Pleistocene and Holocene hominin tracks consistent with the geological age of the site, yet distinct from the Pliocene tracks from Laetoli.

Acknowledgements

This research was funded by Liverpool John Moores University, the Calleva Foundation and Human Origins Research Fund. We would like to thank Professor Bogdan Onac for access to the Vartop footprint, and Masao and colleagues (2016) for making the Site S footprints available online for further research. We would also like to thank the National Trust at Formby for permission to conduct research, the National Museum of Kenya, the Iziko Museums of South Africa, and the Terra Amata Museum, Nice. Finally, we would like to thank one identified and three anonymous reviewers for their comments on a previous version of this paper. MRB acknowledges Natural Environment Research Council grants (NE/H004246/1 and NE/M021459/1).

REFERENCES

Ashton, N. M, Lewis, S. G., De Groote, I., Duffy, S. M., Bates, M., Bates, R., Hoare, P., Lewis, M., Parfitt, S. A., Peglar, S., Williams, C. & Stringer, C. 2014. Hominin footprints from early Pleistocene deposits at Happisburgh, UK. PLoS One 9, e88329.

Bates, K. T., Collins, D., Savage, R., McClymont, J., Webster, E., Pataky, T. C., D'Aout, K., Sellers, W. I., Bennett, M. R. & Crompton, R. H. 2013a. The evolution of compliance in the human lateral mid-foot. The Royal Society Publishing: Proccedings B. 280(1769): 20131818.

Bates, K. T., Savage, R., Pataky, T. C., Morse, S. A., Webster, E., Falkingham, P. L., Ren, L., Qian, Z., Collins, D., Bennett, M. R., McClymont, J. & Crompton R. H. 2013b. Does footprint depth correlate with foot motion and pressure? Journal of The Royal Society Interface 10, 20130009.

Belvedere, M., Ernst, S. & Marty, D. The precision of outline drawings: a comparison based on dinosaur tracks. Inchnia 2016: abstract books. UNESCO Geopark Naturtejo/International Ichnological Association, Castelo Branco.

Bennett, M. R., Harris, J. W., Richmond, B. G., Braun, D. R., Mbua, E., Kiura, P., Olago, D., Kibunjia, M., Omuombo, C., Behrensmeyer, A. K., Huddart, D. & Gonzalez, S. 2009. Early hominin foot morphology based on 1.5-million-year-old footprints from Ileret, Kenya. Science 323, 1197-201.

Bennett, M. R. & Morse, S. A. 2014. Inferences from Human Tracks. Human Footprints: Fossilised Locomotion? Springer, London.

Bennett, M. R., Morse, S. A., Liutkus-Pierce, C., Mcclymont, J., Evans, M., Crompton, R. H. & Francis Thackeray, J. 2014. Exceptional preservation of juvenileren's footprints from a Holocene footprint site in Namibia. Journal of African Earth Sciences 97, 331-341.

Bennett, M. R., Reynolds, S. C., Morse, S. A. & Budka, M. 2016a. Footprints and human evolution: Homeostasis in foot function? Palaeogeography, Palaeoclimatology, Palaeoecology 461, 214-223.

Bennett, M. R., Reynolds, S. C., Morse, S. A. & Budka, M. 2016b. Laetoli's lost tracks: 3D generated mean shape and missing footprints. Science Reports 6, 21916.

Berge, C., Penin, X. & Pellé, É. 2006. New interpretation of Laetoli footprints using an experimental approach and Procrustes analysis: Preliminary results. Comptes Rendus Palevol 5, 561-569.

569 Carbonell, E., Bermudez De Castro, J. M., Pares, J. M., Perez-Gonzalez, A., Cuenca-Bescos,
570 G., Olle, A., Mosquera, M., Huguet, R., Van Der Made, J., Rosas, A., Sala, R.,
571 Vallverdu, J., Garcia, N., Granger, D. E., Martinon-Torres, M., Rodriguez, X. P.,
572 Stock, G. M., Verges, J. M., Allue, E., Burjachs, F., Caceres, I., Canals, A., Benito,
573 A., Diez, C., Lozano, M., Mateos, A., Navazo, M., Rodriguez, J., Rosell, J. &
574 Arsuaga, J. L. 2008. The first hominin of Europe. Nature 452, 465-9.
575 Cohen, J,. 1988. Statistical Power Analysis for the Behavioural Sciences. New Jersey:
576 Erlbaum.
577 Collyer, M. L., Sekora, D. J. & Adams, D. C. 2015. A method for analysis of
578 phenotypic change for phenotypes described by high-dimensional data. Heredity
579 115:357–365.
580 Costa-Perez, M., Moratalla, J. J. & Marugan-Lobon, J. 2019. Studying bipedal dinosaur
581 trackways using geometric morphometrics. Palaeontologia Electronica 22.3, pvc-3, 1-
582 13.
583 Crompton, R. H., Pataky, T. C., Savage, R., D'aout, K., Bennett, M. R., Day, M. H., Bates,
584 K., Morse, S. & Sellers, W. I. 2012. Human-like external function of the foot, and
585 fully upright gait, confirmed in the 3.66 million year old Laetoli hominin footprints by
586 topographic statistics, experimental footprint-formation and computer simulation.
587 Journal of The Royal Society Interface 9, 707-19.
588 D'aout, K., Meert, L., Van Gheluwe, B., De Clercq, D. & Aerts, P. 2010. Experimentally
589 generated footprints in sand: Analysis and consequences for the interpretation of
590 fossil and forensic footprints. American Journal of Physical Anthropology 141, 515-
591 25.
592 Davenport, C. B. 1932. The growth of the human foot. American Journal of Physical
593 Anthropology 17(2), 167-211.
594 de Onis, M. 2006. Assessment of differences in linear growth among populations in the
595 WHO Multicentre Growth Reference Study. Acta Paediatr 95, 56-65.
596 Dingwall, H. L., Hatala, K. G., Wunderlich, R. E. & Richmond, B. G. 2013. Hominin stature,
597 body mass, and walking speed estimates based on 1.5 million-year-old fossil
598 footprints at Ileret, Kenya. Journal of Human Evolution 64, 556-68.
599 Domjanic, J., Seidler, H. & Mitteroecker, P. 2015. A combined morphometric analysis of
600 foot form and its association with sex, stature, and body mass. American Journal of
601 Physical Anthropology 157, 582-91.
602 Dryden, I. L. & Mardia, K. V. 1998. Statistical Shape Analysis. Wiley, Chichester.

603 Duveau, J., Berillon, G., Verna, G., Laisné, G. & Cliquet., D. 2019. The composition of a
604 Neandertal social group revealed by the hominin footprints at Le Rozel (Normandy,
605 France). Proceedings of the National Academy of Sciences 116 (39), 19409-19414.
606 Falkingham, P. L. 2012. Acquisition of high resolution 3D models using free, open-source,
607 photogrammetric software. Palaeontologia Electronica 15(1), 1-15.
608 Ganley, K. J. & Powers, C. M. 2005. Gait kinematics and kinetics of 7-year-old juvenileren: a
609 comparison to adults using age-specific anthropometric data. Gait Posture 21, 141-5.
610 Gierliński, G. D., Niedźwiedzki, G., Lockley, M. G., Athanassiou, A., Fassoulas, C.,
611 Dubicka, Z., Boczarowski, A., Bennett, M. R. & Ahlberg, P. E. 2017. Possible
612 hominin footprints from the late Miocene (c. 5.7 Ma) of Crete? Proceedings of the
613 Geologists' Association 128, 697-710.
614 Hatala, K. G., Demes, B. & Richmond, B. G. 2016a. Laetoli footprints reveal bipedal gait
615 biomechanics different from those of modern humans and chimpanzees. Proceedings
616 of Royal Soceity B: Biological Sciences 283: 20160235.
617 Hatala, K. G., Dingwall, H. L., Wunderlich, R. E. & Richmond, B. G. 2013. The relationship
618 between plantar pressure and footprint shape. Journal of Human Evolution 65**,** 21-8.
619 Hatala, K. G., Perry, D. A. & Gatesy, S. M. 2018. A biplanar X-ray approach for studying the
620 3D dynamics of human track formation. Journal of Human Evolution 121, 104-118.
621 Hatala, K. G., Roach, N. T., Ostrofsky, K. R., Wunderlich, R. E., Dingwall, H. L., Villmoare,
622 B. A., Green, D. J., Braun, D. R., Harris, J. W. K., Behrensmeyer, A. K., Richmond,
623 B. G. 2017. Hominin track assemblages from Okote Member deposits near Ileret,
624 Kenya, and their implications for understanding fossil hominin paleobiology at 1.5
625 Ma. Journal of Human Evolution 112, 93-104.
626 Hatala, K. G., Roach, N. T., Ostrofsky, K. R., Wunderlich, R. E., Dingwall, H. L.,
627 Villmoare, B. A., Green, D. J., Harris, J. W. K., Braun, D. R., Richmond, B. G.
628 2016b. Footprints preserve direct evidence of group behavior and locomotion in
629 *Homo erectus*. Scientific Reports 6, 28766.
630 Hatala, K. G., Wunderlich, R. E., Dingwall, H. L. & Richmond, B. G. 2016c. Interpreting
631 locomotor biomechanics from the morphology of human footprints. Journal of Human
632 Evolution 90**,** 38-48.
633 Hitchcock, E. 1858. Ichnology of New England: A Report on the Sandstone of the
634 Connecticut Valley, Especially its Fossil Footmarks, Made to the Government of the
635 Commonwealth of Massachusetts Boston, Massachusetts: W. White xii + 220 pp.

636 Jordan, K. & Newll, K. M. 2008. The structure and variability in human walking and running
637 is speed dependent. Exercise and Sport Science Reviews 36(4), 200-2004.
638 Kanchan, T., Krishan, K., Prusty, D. & Machado, M. 2014. Heel–Ball index: An analysis of
639 footprint dimensions for determination of sex. Egyptian Journal of Forensic Sciences
640 4, 29-33.
641 Keith, A. 1929. The history of the human foot and its bearing on orthopaedic practise. The
642 Journal of Bone and Joint Surgery 11, 10-32.
643 Kim, J. L., Kim, K. S., Lockley, M. G. & Matthews, N. 2008. Hominid Ichnotaxonomy: An
644 Exploration of a Neglected Discipline. Ichnos 15(3-4), 126-139.
645 Krishan, K. 2007. Individualizing characteristics of footprints in Gujjars of North India--
646 forensic aspects. Forensic Science International 169, 137-44.
647 Krishan, K. 2008. Establishing correlation of footprints with body weight--forensic
648 aspects.Forensic Science International 179, 63-9.
649 Lockley, M. G., Meldrum, D. J. & Kim, J. Y. 2016. Major events in hominin evolution. In:
650 M. G. Mángano, L.A. Buatois (Eds.), The Trace-Fossil Record of Major Evolutionary
651 Events in Geobiology. 40, 411-448, Springer, Dordrecht.
652 Marchetti, L., Belvedere, M., Voigt, S., Klein, H., Castanera, D., Diaz-Martinez, I., Matry,
653 D., Xing, L., Feola, S., Melchor, R. N. & Farlow, J. O. 2019. Defining the
654 morphological quality of fossil footprints. Problems and principles of preservation in
655 tetrapod ichnology with examples from the Palaeozoic to the present. Earth-Science
656 Reviews 193, 109-145.
657 Mardia, K.V., Bookstein, K. V. & Moreton, I. J. 2000. Statistical assessment of bilateral
658 symmetry of shapes. Biometrika 87:285-300.
659 Masao, F. T., Ichumbaki, E. B., Cherin, M., Barili, A., Boschian, G., Iurino, D. A.,
660 Menconero, S., Moggi-Cecchi, J. & Manzi, G. 2016. New footprints from Laetoli
661 (Tanzania) provide evidence for marked body size variation in early hominins. Elife
662 5:e19568.
663 McClymont, J., Pataky, T. C., Crompton, R. H., Savage, R. & Bates, K. T. 2016. The nature
664 of functional variability in plantar pressure during a range of controlled walking
665 speeds. Royal Society Open Science 3, 160369.
666 Meldrum, D. J. 2007. Renewed perspective on the Laetoli trackways: The earliest hominid
667 footprints. New Mexico Museum of Natural History and Science Bulletin 42, 233–
668 239.

669 Meldrum, D.J., Lockley, M. G., Lucas, S. G. & Musiba, C. M. 2011. Ichnotaxonomy of the
670 Laetoli Trackways: The earliest hominin footprints. Journal of African Earth Science
671 60, 1-12.
672 Mitteroecker, P. & Bookstein, F. 2011. Linear discrimination, ordination, and the
673 visualization of selection gradients in modern morphometrics. Evolutionary Biology
674 38, 100-114.
675 Morse, S. A., Bennett, M. R., Gonzalez, S. & Huddart, D. 2010. Techniques for verifying
676 human footprints: reappraisal of pre-Clovis footprints in Central Mexico. Quaternary
677 Science Reviews 29, 2571-2578.
678 Morse, S. A., Bennett, M. R., Liutkus-Pierce, C., Thackeray, F., Mcclymont, J., Savage, R. &
679 Crompton, R. H. 2013. Holocene footprints in Namibia: the influence of substrate on
680 footprint variability. American Journal of Physical Anthropology 151, 265-79.
681 Mullin, S. K. & Taylor, P. J. 2002. The effects of parallax on geometric morphometric data.
682 Computers in Biology and Medicine 32, 455-464.
683 Pablos, A., Lorenzo, C., Martinez, I., Bermudez De Castro, J. M., Martinon-Torres, M.,
684 Carbonell, E. & Arsuaga, J. L. 2012. New foot remains from the Gran Dolina-TD6
685 Early Pleistocene site (Sierra de Atapuerca, Burgos, Spain). Journal of Human
686 Evolution 63, 610-23.
687 Pataky, T. C., Savage, R., Bates, K. T., Sellers, W. I. & Crompton, R. H. 2013. Short-term
688 step-to-step correlation in plantar pressure distributions during treadmill walking, and
689 implications for footprint trail analysis. Gait & Posture 38(4), 1054-1057.
690 Raichlen, D. A. & Gordon, A. D. 2017. Interpretation of footprints from Site S confirms
691 human-like bipedal biomechanics in Laetoli hominins. Journal of Human Evolution
692 107, 134-138.
693 Raichlen, D. A., Pontzer, H. & Sockol, M. D. 2008. The Laetoli footprints and early hominin
694 locomotor kinematics. Journal of Human Evolution 54, 112-7.
695 Rohlf, F. J. 2003. Bias and error in estimates of mean shape in geometric morphometrics.
696 Journal of Human Evolution 44, 665-683.
697 Stavlas, P., Grivas, T. B., Michas, C., Vasiliadis, E. & Polyzois, V. 2005. The evolution of
698 foot morphology in children between 6 and 17 years of age: a cross-sectional study
699 based on footprints in a Mediterranean population. Journal of Foot amd Ankle
700 Surgery 44, 424-8.
701 Tuttle, R. H. 2008. Footprint clues in hominid evolution and forensics: Lessons and
702 limitations. Ichnos 15, 158-165.

703 Vaughan, C. L. & Blaszczyk, M. B. 2008. Dynamic similarity predicts gait parameters for
704 *Homo floresiensis* and the Laetoli hominins. American Journal of Human Biology 20,
705 312-6.
706 Webb, S. 2007. Further research of the Willandra Lakes fossil footprint site, southeastern
707 Australia. Journal of Human Evolution 52, 711-5.
708 Webb, S., Cupper, M. L. & Robins, R. 2006. Pleistocene human footprints from the
709 Willandra Lakes, Southeastern Australia. Journal of Human Evolution 50, 405-13.
710 Wiseman, A. L .A. & De Groote, I. 2018. A three-dimensional geometric morphometric
711 study of the effects of erosion on the morphologies of modern and prehistoric
712 footprints. Journal of Archaeological Science: Reports 17, 93-102.
713 Wood, B. & Richmond, B. G. 2000. Human evolution: taxonomy and paleobiology. Journal
714 of Anatomy 196, 19-60.
715

716 **Supporting Information**

717 Table S1 listing tracks used in the study. Where initials are used, data was collected by
718 authors. Formby Point tracks provided by AW were excavated/recorded in 2016 and 2017 by
719 AW. For access to these tracks please contact AW.

720

721 3D data availability:

722 For access for Site G footprints see: http://footprints.bournemouth.ac.uk/archive/Laetoli/

723 For access for Site S footprints see:
724 https://www.morphosource.org/Search/Index?search=laetoli

725 For access for Namibian footprints see:
726 http://footprints.bournemouth.ac.uk/archive/Namibian%20Footprints/

727

728 Table S1 overleaf.

FOSSIL LOCALITY	SITE	TRACK NUMBER	TRACK ID	DATA PROVIDED BY	DATA ONLINE?
LAETOLI	Site G		G1-35	MRB	Yes
	Site G		G1-36	MRB	Yes
	Site G		G1-37	MRB	Yes
	Site G		G1-38	MRB	Yes
	Site G		G1-39	MRB	Yes
	Site G		G1-25	MRB	Yes
	Site G		G1-26	MRB	Yes
	Site G		G1-27	MRB	Yes
	Site G		G1-28	MRB	Yes
	Site G		G1-30	MRB	Yes
	Site G		G1-31	MRB	Yes
	Site G		G1-34	MRB	Yes
	Site S		L8S11	Masao *et al.* 2016	Yes
	Site S		L8S12	Masao *et al.* 2016	Yes
	Site S		L8S13	Masao *et al.* 2016	Yes
	Site S		L8S14	Masao *et al.* 2016	Yes
	Site S		M9S12	Masao *et al.* 2016	Yes
	Site S		M9S13	Masao *et al.* 2016	Yes
	Site S		TP2S2111	Masao *et al.* 2016	Yes
ILERET	FwJj14E Upper Footprint Layer		FU-A	KGH	No
	FwJj14E Upper Footprint Layer		FU-H	KGH	No
	FwJj14E Upper Footprint Layer	1	FUT1-6	KGH	No
	FwJj14E Upper Footprint Layer	1	FUT1-7A	KGH	No
	FwJj14E Upper Footprint Layer		FUT1-7B	KGH	No
	FwJj14E Upper Footprint Layer	1	FUT1-12	KGH	No
	FwJj14E Upper Footprint Layer	1	FUT1-13	KGH	No
	FwJj14E Upper Footprint Layer	1	FUT1-16	KGH	No
	FwJj14E Upper Footprint Layer	2	FUT2-1	KGH	No
	FwJj14E Upper Footprint Layer	2	FUT2-2	KGH	No
	FwJj14E Upper Footprint Layer	2	FUT2-4	KGH	No
	FwJj14E Upper Footprint Layer		FUT3-1	KGH	No
HAPPISBURGH	-	-	Print 33	IDG&CS&NA&SD	No
	-	-	Print 39	IDG&CS&NA&SD	No
	-	-	Print 40	IDG&CS&NA&SD	No
	-	-	Print 49	IDG&CS&NA&SD	No
	-	-	Print 3	IDG&CS&NA&SD	No
	-	-	Print 4	IDG&CS&NA&SD	No
	-	-	Print 5	IDG&CS&NA&SD	No
	-	-	Print 6	IDG&CS&NA&SD	No
	-	-	Print 8	IDG&CS&NA&SD	No
	-	-	Print 9	IDG&CS&NA&SD	No
	-	-	Print 11	IDG&CS&NA&SD	No
	-	-	Print 12	IDG&CS&NA&SD	No
	-	-	Print 14	IDG&CS&NA&SD	No

	-	-	Print 18	IDG&CS&NA&SD	No
LANGEBAAN	-	-	-	Iziko Museum	No
	-	-	-	Iziko Museum	No
TERRA AMATA	-	-	-	Terra Amata Museum	No
VARTOP CAVE	-	-	-	Prof. Bogdan Onac	No
WALVIS BAY	Site One	-	PATCH 7.1b	MRB	Yes
	Site One	-	PATCH 41 4	MRB	Yes
	Site One	-	PATCH 74 4	MRB	Yes
	Site One	-	H302	MRB	Yes
	Site One	-	H305	MRB	Yes
	Site One	-	H308	MRB	Yes
	Site One	-	H309	MRB	Yes
	Site One	-	H310	MRB	Yes
	Site One	-	H311	MRB	Yes
	Site One	-	H312	MRB	Yes
	Site One	-	H313	MRB	Yes
	Site One	-	H314	MRB	Yes
	Site One	-	H315	MRB	Yes
	Site One	-	H316	MRB	Yes
	Site One	-	H141	MRB	Yes
	Site One	-	H142	MRB	Yes
	Site One	-	H143	MRB	Yes
	Site One	-	H144	MRB	Yes
	Site One	-	H145	MRB	Yes
	Site One	-	H149	MRB	Yes
	Site One	-	H155	MRB	Yes
	Site One	-	H158	MRB	Yes
	Site One	-	H159	MRB	Yes
	Site One	-	H162	MRB	Yes
	Site One	-	H318	MRB	Yes
	Site One	-	H319	MRB	Yes
	Site One	-	H321	MRB	Yes
	Site One	-	H322	MRB	Yes
	Site One	-	H323	MRB	Yes
	Site One	-	H324	MRB	Yes
	Site One	-	H59	MRB	Yes
	Site One	-	H68	MRB	Yes
	Site One	-	H65	MRB	Yes
	Site One	-	H60	MRB	Yes
	Site One	-	H66	MRB	Yes
	Site One	-	H51	MRB	Yes
	Site One	-	H55	MRB	Yes
	Site One	-	H56	MRB	Yes
	Site One	-	H57	MRB	Yes
	Site One	-	H42	MRB	Yes
	Site One	-	H45	MRB	Yes
	Site One	-	H05	MRB	Yes
	Site One	-	H06	MRB	Yes
	Site One	-	H07	MRB	Yes
	Site One	-	H13	MRB	Yes
	Site One	-	H14	MRB	Yes
	Site One	-	H70	MRB	Yes
	Site One	-	H070A	MRB	Yes
	Site One	-	H72	MRB	Yes
	Site One	-	H74	MRB	Yes
	Site One	-	H075	MRB	Yes
	Site One	-	H14	MRB	Yes
	Site One	-	H21	MRB	Yes
	Site One	-	H42	MRB	Yes

	Site One	-	H43	MRB	Yes
	Site One	-	H45	MRB	Yes
	Site One	-	H49	MRB	Yes
	Site One	-	H47	MRB	Yes
	Site One	-	301	MRB	Yes
	Site One	-	317	MRB	Yes
	Site One	-	H078	MRB	Yes
	Site One	-	H079	MRB	Yes
	Site One	-	H082	MRB	Yes
	Site One	-	H086	MRB	Yes
	Site One	-	H086	MRB	Yes
	Site One	-	H087	MRB	Yes
	Site One	-	H091	MRB	Yes
	Site One	-	H097	MRB	Yes
	Site One	-	H098	MRB	Yes
	Site One	Trail One	H10	MRB	Yes
	Site One	Trail One	H11	MRB	Yes
	Site One	Trail One	H1	MRB	Yes
	Site One	Trail One	H2	MRB	Yes
	Site One	Trail One	H3	MRB	Yes
	Site One	Trail One	H4	MRB	Yes
	Site One	Trail One	H5	MRB	Yes
	Site One	Trail One	H6	MRB	Yes
	Site One	Trail One	H7	MRB	Yes
	Site One	Trail One	H8	MRB	Yes
	Site One	Trail One	H9	MRB	Yes
	Site One	Trail One	H10	MRB	Yes
	Site One	Trail One	H11	MRB	Yes
	Site One	Trail One	H12	MRB	Yes
	Site One	Trail One	H16	MRB	Yes
	Site One	Trail One	H17	MRB	Yes
	Site One	Trail One	H18	MRB	Yes
	Site One	Trail One	H19	MRB	Yes
	Site One	Trail One	H20	MRB	Yes
	Site One	Trail One		MRB	Yes
	Site One	Trail One	H22	MRB	Yes
	Site One	Trail One	H23	MRB	Yes
	Site One	Trail One	H24	MRB	Yes
	Site One	Trail One	h26	MRB	Yes
	Site One	Trail One	H27	MRB	Yes
	Site One	Trail One	H29	MRB	Yes
	Site One	Trail One	H30	MRB	Yes
	Site One	Trail One	H31	MRB	Yes
	Site One	Trail One	H32	MRB	Yes
	Site One	Trail One	H33	MRB	Yes
	Site One	Trail One	H33	MRB	Yes
	Site One	Trail One	H34	MRB	Yes
	Site One	Trail One	H35	MRB	Yes
	Site One	Trail One	H36	MRB	Yes
	Site One	Trail One	H37	MRB	Yes
	Site One	Trail One	H38	MRB	Yes
	Site One	Trail One	H39	MRB	Yes
	Site One	Trail One	H37	MRB	Yes
	Site One	Trail One	H44	MRB	Yes
	Site One	Trail One	H46	MRB	Yes
	Site One	Trail One	H48	MRB	Yes
	Site One	Trail One	H61	MRB	Yes
	Site One	Trail One	H62	MRB	Yes
	Site One	Trail One	H64	MRB	Yes
	Site One	Trail One	H69	MRB	Yes

	Site One	Trail One	H71	MRB	Yes
	Site One	Trail One	H73	MRB	Yes
	Site One	Trail Two	H77	MRB	Yes
	Site One	Trail Two	HR20	MRB	Yes
	Site One	Trail Two	HR21	MRB	Yes
	Site One	Trail Two	HR29	MRB	Yes
	Site One	Trail Two	HR31	MRB	Yes
	Site One	Trail Two	HR36	MRB	Yes
	Site One	Trail Two	HR44	MRB	Yes
	Site One	Trail Two	HR51	MRB	Yes
	Site One	Trail Two	HR67	MRB	Yes
	Site One	Trail Two	HR86	MRB	Yes
	Site One	Trail Two	HR89	MRB	Yes
	Site One	Trail Two	HR91	MRB	Yes
	Site One	Trail Two	HR116	MRB	Yes
	Site One	Trail Two	HR130	MRB	Yes
	Site One	Trail Two	HARIETTE4	MRB	Yes
	Site One	Trail Two	HARRIETTE 12	MRB	Yes
	Site One	Trail Two	HARRIETTE 17	MRB	Yes
	Site One	Trail Two	HARRIETTE 18	MRB	Yes
	Site One	Trail Two	HARRIETTE 20	MRB	Yes
	Site One	Trail Two	HARRIETTE 21	MRB	Yes
	Site One	Trail Two	HARRIETTE 29	MRB	Yes
	Site One	Trail Two	HARRIETTE 46	MRB	Yes
	Site One	Trail Two	HARRIETTE 51	MRB	Yes
	Site One	Trail Two	HARRIETTE 54	MRB	Yes
	Site One	Trail Two	HARRIETTE 20	MRB	Yes
	Site One	Trail Two	HARRIETTE 21	MRB	Yes
	Site One	Trail Two	HARRIETTE 29	MRB	Yes
	Site One	Trail Two	HARRIETTE 46	MRB	Yes
	Site One	Trail Two	HARRIETTE 51	MRB	Yes
	Site One	Trail Two	HARRIETTE 54	MRB	Yes
	Site One	Trail Two	HAR18239	MRB	Yes
FORMBY POINT	Sefton Coast	-	PRINT A	MRB	Yes
	Sefton Coast	-	PRINT AA	MRB	Yes
	Sefton Coast	-	PRINT B	MRB	Yes
	Sefton Coast	-	PRINT F	MRB	Yes
	Sefton Coast	-	PRINT I	MRB	Yes
	Sefton Coast	-	PRINT J	MRB	Yes
	Sefton Coast	-	PRINT K	MRB	Yes
	Sefton Coast	-	PRINT L	MRB	Yes
	Sefton Coast	-	PRINT M	MRB	Yes
	Sefton Coast	-	PRINT N	MRB	Yes
	Sefton Coast	-	PRINT O	MRB	Yes

	Sefton Coast	-	PRINT P	MRB	Yes
	Sefton Coast	-	PRINT Q	MRB	Yes
	Sefton Coast	-	PRINT R	MRB	Yes
	Sefton Coast	-	PRINT S	MRB	Yes
	Sefton Coast	-	PRINT TT	MRB	Yes
	Sefton Coast	-	PRINT W	MRB	Yes
	Sefton Coast	-	PRINT X	MRB	Yes
	Sefton Coast	-	PRINT ZZ	MRB	Yes
	Sefton Coast	-	PRINT T5	MRB	Yes
	Cornerstone N	Track 13	285	AW	No
	Cornerstone N	Track 13	286	AW	No
	Cornerstone N	Track 13	289	AW	No
	Cornerstone N	Track 13	292	AW	No
	Cornerstone N	Track 11	231	AW	No
	Cornerstone N	Track 11	225	AW	No
	Cornerstone N	Track 11	219	AW	No
	Cornerstone N	Track 11	220	AW	No
	Cornerstone N	Track 7	202	AW	No
	Cornerstone N	Track 8	204	AW	No
	Cornerstone N	Track 8	205	AW	No
	Cornerstone N	Track 8	210	AW	No
	Cornerstone N	Track 9	212	AW	No
	Cornerstone N	Track 9	213	AW	No
	Cornerstone N	Track 9	215	AW	No
	Cornerstone N	Track 9	214	AW	No
	Cornerstone N	Track 10	216	AW	No
	Cornerstone N	Track 10	216-a	AW	No
	Cornerstone N	Track 10	216-b	AW	No
	Cornerstone N	Track 10	233	AW	No
	Cornerstone N	Track 10	223	AW	No
	Cornerstone N	Track 10	229	AW	No
	Cornerstone N	-	202	AW	No
	Blundell Path	1	1295	AW	No
	Blundell Path	1	1296	AW	No
	Blundell Path	1	1297	AW	No
	Blundell Path	1	1298	AW	No
	Blundell Path	1	1299	AW	No
	Blundell Path	1	1300	AW	No
	Blundell Path	1	1301	AW	No
	Blundell Path	2	1303	AW	No
	Blundell Path	2	1304	AW	No
	Blundell Path	2	1305	AW	No
	Blundell Path	3	1272	AW	No
	Blundell Path	3	1365	AW	No
	Blundell Path	3	1366	AW	No
	Blundell Path	3	UNI	AW	No
	Blundell Path	3	dpL	AW	No
	Blundell Path	3	dpR	AW	No
	Blundell Path	-	350	AW	No
	Blundell Path	-	280	AW	No
	Blundell Path	-	273	AW	No
	Blundell Path	Track 5	309	AW	No
	Blundell Path	Track 5	310	AW	No
	Blundell Path	Track 18	348	AW	No
	Blundell Path C	Track 18	349	AW	No
	Blundell Path C		220	AW	No
	Blundell Path C	-	316	AW	No
	Blundell Path C	Track 4	261	AW	No
	Blundell Path C	Track 4	262	AW	No
	Blundell Path C	Track 4	263	AW	No

	Blundell Path C	Track 4	264	AW	No
	Blundell Path C	Track 4	265	AW	No
	Gypsy Path	Track 15	print f101	AW	No
	Gypsy Path	Track 15	print f110	AW	No
	Gypsy Path	Track 15	**print f8**	AW	No
	Gypsy Path	Track 15	print f31	AW	No
	Gypsy Path	Track 15	print f113	AW	No
	Gypsy Path	Track 15	**print f40**	AW	No

729

Supplementary Information: Table S2. Results of the replicability tests for all track measurements from each fossil location.

	Track Length		**Long Axis**		**Hallux Length**	
	MSE	VARIANCE	MSE	VARIANCE	MSE	VARIANCE
Laetoli	1.107%	0.069% ± 1.215	0.877%	0.126% ± 1.359	0.627%	1.716% ± 0.971
Ileret	0.442%	0.416% ± 1.091	0.678%	0.416% ± 1.051	0.899%	0.422% ± 1.392
Happisburgh	0.041%	0.513% ± 0.294	0.038%	0.653% ± 1.601	0.926%	0.237% ± 1.434
Formby Point	0.455%	0.179% ± 0.705	0.230%	0.185% ± 0.356	0.067%	0.186% ± 0.104
Walvis Bay	0.489%	0.494% ± 0.757	0.587%	0.281% ± 0.909	0.411%	0.691% ± 0.637

	Forefoot width		**Heel width**	
	MSE	VARIANCE	MSE	VARIANCE
Laetoli	0.028%	0.830% ± 0.647	0.931%	1.920% ± 0.196
Ileret	0.020%	0.304% ± 0.517	0.014%	0.457% ± 0.351
Happisburgh	0.251%	0.046% ± 0.389	0.941%	0.420% ± 1.458
Formby Point	0.612%	0.471% ± 0.942	0.243%	0.911% ± 0.367
Walvis Bay	0.842%	0.596% ± 1.307	0.524%	1.612% ± 0.789

Supplementary Information: Table S3. Mean measurements (mm) and mean predicted stature (mm) of each individual. As determining which track belongs to a certain individual in the Happisburgh hominins is subjective, the group means are reported for inferred juvenile and adult prints. Individual tracks not belonging to a trackway from Ileret, Formby Point and Walvis Bay are not reported here. Group means provided from Group One and Group Two from Walvis Bay are provided (Bennett et al. 2014).

	Track ID	Length	Stature	2nd digit to heel	Heel breadth	Forefoot breadth
Laetoli	M9-S1	256.71	1711.67	247.02	65.46	101.85
	L8-S1	261.02	1740.13	262.32	78.75	103.25
	G1	173.93	1159.54	165.26	46.08	73.44
	TP2-S1	271.01	1806.73	272.11	82.00	99.45
Ileret	FUT1A	261.06	1740.39	259.45	48.08	82.79
	FUT2	283.98	1893.17	274.02	57.20	93.96
Happisburgh	Juvenile mean	150.02	1000.11	150.40	31.34	63.15
	Adult mean	217.72	1451.49	208.90	49.09	77.34
Terra Amata	Single print	242.66	1617.75	250.13	53.78	83.34
Vârtop Cave	Single Print	222.25	1481.17	210.68	77.20	113.72
Langebaan	1	220.00	1466.67	/	62.96	89.42
	2	220.50	1470.00	/	/	/
Formby Point	1	113.87	759.12	106.11	41.89	76.20
	2	250.78	1671.85	241.21	58.79	88.73
	3	204.67	1364.47	198.72	50.97	72.25
	4	274.86	1832.41	263.94	46.45	88.50
	5	230.15	1534.33	210.35	45.34	82.55
	6	207.03	1380.17	192.46	40.96	64.97
	7	259.54	1730.26	230.36	51.33	87.34
	8	235.52	1570.11	217.77	34.27	76.39
	9	260.74	1738.26	251.92	47.72	86.53
	10	278.96	1859.73	255.96	47.79	102.99
Walvis Bay	Group One	172.89	1490.85	158.92	42.16	61.12
	Group Two	204.94	1366.27	189.08	45.12	62.40
	Trail One	255.25	1678.58	238.11	62.33	88.75
	Trail Two	229.43	1529.56	212.67	52.44	75.14

741 **Supplementary Information: Table S4**. Results of the ANOVA and Games-Howell Test. Table displays the between-groups variability of linear
742 measurements of the track and stature. Both df1 (between-groups) and df2 (within-groups) are reported. Levels of significance are reported within
743 a 95% confidence level. Significant P values are in bold.
744

	One-way ANOVA				**Games-Howell Test**			
Measurement (mm)	df1	df2	f	P	*Between-groups variability*		Std. error (mm)	P
Foot length	4	220	18.4	**<0.001**	Laetoli	Ileret	13.26	**<0.001**
						Happisburgh	26.09	0.997
						Formby Point	9.21	**<0.001**
						Walvis Bay	7.78	**0.009**
					Ileret	Happisburgh	27.23	0.126
						Formby Point	12.06	0.169
						Walvis Bay	11.01	**0.005**
					Happisburgh	Formby Point	25.50	0.476
						Walvis Bay	25.02	0.900
					Formby Point	Walvis Bay	5.50	**0.002**
Stature	4	220	19.266	**<0.001**	Laetoli	Ileret	88.38	**<0.001**
						Happisburgh	173.93	0.997
						Formby Point	61.14	**<0.001**
						Walvis Bay	51.89	**0.009**

Table S4 cont. Results of the ANOVA and Games-Howell Test. Table displays the between-groups variability of linear measurements of the track and stature. Both df1 (between-groups) and df2 (within-groups) are reported. Levels of significance are reported within a 95% confidence level. * indicates statistically significant variability between-groups.

	One-way ANOVA				**Games-Howell Test**			
Measurement (mm)	df1	df2	f	P	*Between-groups variability*		Std. error (mm)	P
					Ileret	Happisburgh	181.52	0.126
						Formby Point	80.22	0.203
						Walvis Bay	73.41	**0.005**
					Happisburgh	Formby Point	169.93	0.449
						Walvis Bay	166.82	0.900
					Formby Point	Walvis Bay	36.28	**0.001**
Long axis of foot	4	220	18.008	**<0.001**	Laetoli	Ileret	14.50	**<0.001**
						Happisburgh	27.46	0.993
						Formby Point	9.61	**<0.001**
						Walvis Bay	8.47	**0.033**
					Ileret	Happisburgh	28.80	0.105
						Formby Point	12.95	**0.026**
						Walvis Bay	12.13	**0.0028**

Table S4 cont. Results of the ANOVA and Games-Howell Test. Table displays the between-groups variability of linear measurements of the track and stature. Both df1 (between-groups) and df2 (within-groups) are reported. Levels of significance are reported within a 95% confidence level. * indicates statistically significant variability between-groups.

	One-way ANOVA				**Games-Howell Test**			
Measurement (mm)	df1	df2	f	P	*Between-groups variability*		Std. error (mm)	P
					Happisburgh	Formby Point	26.67	0.690
						Walvis Bay	26.28	0.970
					Formby Point	Walvis Bay	5.40	**0.006**
Forefoot breadth	4	220	2.489	**0.044**	Laetoli	Ileret	6.76	0.327
						Happisburgh	12.85	1.000
						Formby Point	4.12	0.323
						Walvis Bay	3.91	0.998
					Ileret	Happisburgh	13.61	0.863
						Formby Point	6.08	0.909
						Walvis Bay	5.95	0.309
					Happisburgh	Formby Point	12.51	0.964
						Walvis Bay	12.44	1.000
					Formby Point	Walvis Bay	2.57	0.065

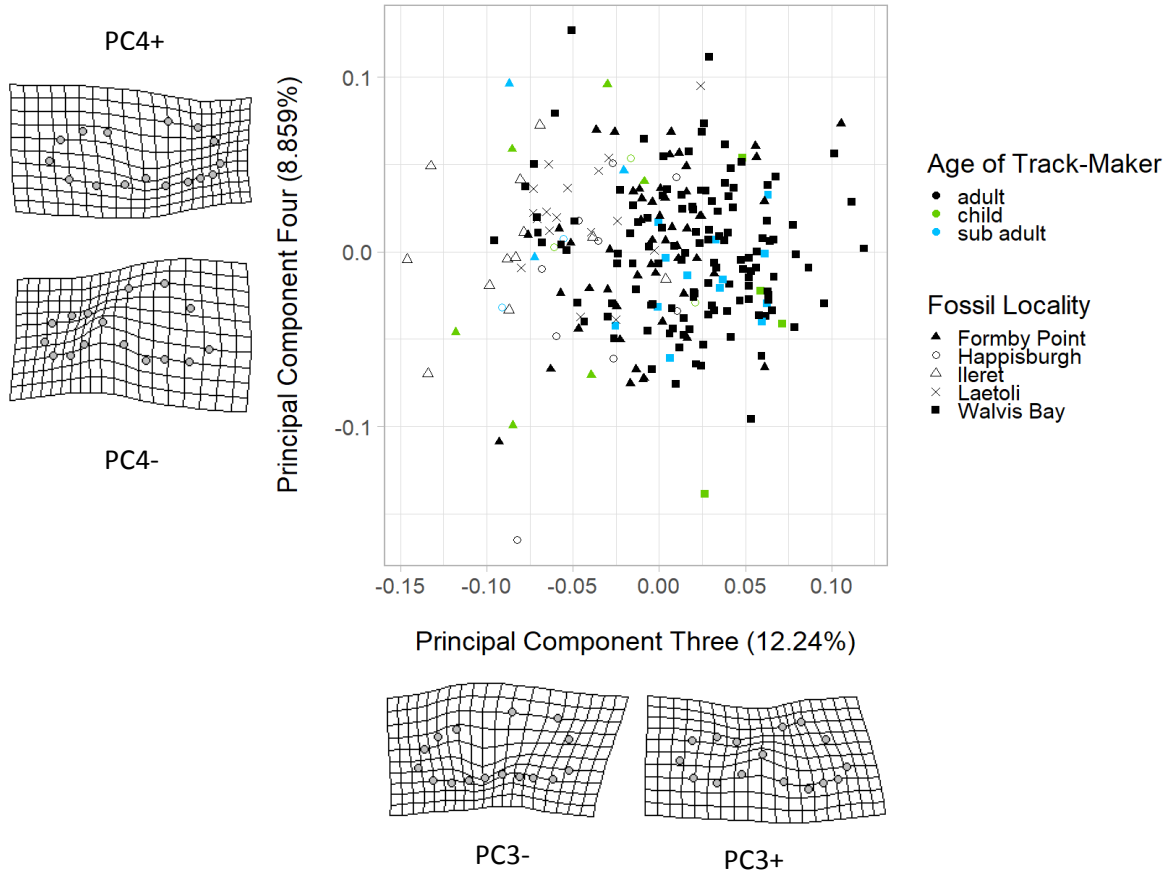
Table S4 cont. Results of the ANOVA and Games-Howell Test. Table displays the between-groups variability of linear measurements of the track and stature. Both df1 (between-groups) and df2 (within-groups) are reported. Levels of significance are reported within a 95% confidence level. * indicates statistically significant variability between-groups.

	One-way ANOVA				**Games-Howell Test**			
Measurement (mm)	df1	df2	f	P	*Between-groups variability*		Std. error (mm)	P
Heel breadth	4	220	3.82	**0.005**	Laetoli	Ileret	5.32	0.990
						Happisburgh	7.11	0.969
						Formby Point	2.93	0.715
						Walvis Bay	2.69	0.728
					Ileret	Happisburgh	8.15	0.915
						Formby Point	4.95	0.728
						Walvis Bay	4.81	1.000
					Happisburgh	Formby Point	6.83	1.000
						Walvis Bay	6.73	0.780
					Formby Point	Walvis Bay	1.85	**0.002**

745

Supplementary Information: Figure S1.

Graphical results of the PCA plotting PC3 against PC4 scores. The axis of PC3 appears to highlight the morphological disparity between AMHs (PC3+ scores) and Pleistocene and Pliocene tracks (PC3- scores).



Supplementary Information: Table S5. Effect sizes (*z*) (Cohen 1988) table displaying the significant shape variability between juvenile and adult fossil tracks, as produced from the MANOVA computed between-groups using the PC scores that represent 100% of shape variance and log-CS.

A shaded grey box indicates that the variability was non-significant between-groups (P≥0.05, within a 95% confidence interval). A shaded green box indicates that significant shape disparity was found between-groups (P<0.05, within a 95% confidence interval). Boxes with a thick black outline indicate within-groups variability (e.g., the juvenile tracks differ in shape from the adult tracks within modern humans at Formby Point).

Δ – Juvenile Track
▼ – Sub-adult Track (these tracks were identified to belong to individuals which were borderline adult; i.e., age predictions were 17-19 years old)
▲ – Adult Track

		Formby Point			Happisburgh			Ileret	Laetoli	Walvis Bay		
		▲	Δ	▼	▲	Δ	▼	▲	▲	▲	Δ	▼
Formby Point	▲	0										
	Δ	6.238	0									
	▼	0.867	-0.413	0								
Happisburgh	▲	-0.465	-1.067	-0.311	0							
	Δ	-0.909	-0.867	-0.489	2.368	0						
	▼	-0.006	-0.878	-0.585	-0.481	-0.553	0					
Ileret	▲	-0.104	-1.421	-1.049	0.094	-0.583	-2.319	0				
Laetoli	▲	-0.103	-0.418	-1.107	-1.238	0.708	-0.316	-0.362	0			
Walvis Bay	▲	-0.219	1.318	0.774	-0.538	-0.522	-0.211	-0.474	-0.639	0		
	Δ	1.552	2.221	1.939	0.578	1.130	1.638	2.739	0.081	2.859	0	
	▼	-0.330	0.873	0.905	1.070	-0.633	0.357	0.217	2.551	5.336	3.251	0