Human Origins in Southern African Palaeo-wetlands? Strong Claims from Weak Evidence

Carina M Schlebusch[1,2,3,]#, Liisa Loog[4]*, Huw S Groucutt[5,6,7], Turi King[8], Adam Rutherford[9], Chiara Barbieri[10], Guido Barbujani[11], Lounes Chikhi[12,13], Chris Stringer[14], Mattias Jakobsson[1,2,3], Anders Eriksson[15,16], Andrea Manica[17], Sarah A Tishkoff[18], Eleanor ML Scerri[19,20.21], Aylwyn Scally[4], Chris Brierley[22], Mark G Thomas[9,]#*

*Equal contribution

#Correspondence: carina.schlebusch@ebc.uu.se and m.thomas@ucl.ac.uk

1 Human Evolution, Department of Organismal Biology, Uppsala University, Sweden
2 Palaeo-Research Institute, University of Johannesburg, South Africa
3 SciLifeLab Uppsala, Sweden
4 Department of Genetics, University of Cambridge, UK
5 Extreme Events Research Group, Max Planck Institutes for Chemical Ecology, the Science of Human History, and Biogeochemistry, Jena, Germany
6 Department of Archaeology, Max Planck Institute for the Science of Human History, Jena, Germany
7 Institute of Prehistoric Archaeology, University of Cologne, Cologne, Germany
8 Department of Genetics and Genome Biology & School of Archaeology and Ancient History, University of Leicester, UK
9 Department of Genetics, Evolution and Environment, University College London, UK
10 Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland
11 Department of Life Sciences and Biotechnology, University of Ferrara, Italy
12 Evolution & Diversité Biologique lab, CNRS, Univ. Toulouse 3, France
13 Instituto Gulbenkian de Ciência, Oeiras, Portugal
14 Centre for Human Evolution Research (CHER), Department of Earth Sciences, Natural History Museum, London, UK
15 Center for Genomics, Evolution and Medicine (cGEM), University of Tartu, Estonia
16 Department of Medical and Molecular Genetics, King's College London, UK
17 Department of Zoology, University of Cambridge, UK
18 Departments of Genetics and Biology, University of Pennsylvania, Philadelphia, PA USA
19 Pan-African Evolution Research Group, Max Planck Institute for the Science of Human History, Jena, Germany
20 Department of Classics and Archaeology, University of Malta, Msida, Malta
21 Institute of Prehistoric Archaeology, University of Cologne, Cologne, Germany
22 Department of Geography, University College London, UK

Abstract

Attempts to identify a ‘homeland’ for our species from genetic data are widespread in the academic literature. However, even when putting aside the question of whether a ‘homeland’ is a useful concept, there are a number of inferential pitfalls in attempting to identify the geographic origin of a species from contemporary patterns of genetic variation. These included making strong claims from weakly informative data, treating genetic lineages as representative of populations, assuming a high degree of regional population continuity over hundreds of thousands of years, and using circumstantial observations as corroborating evidence without considering alternative hypotheses on an equal footing, or formally evaluating any hypothesis. In this commentary we review the recent publication that claims to pinpoint the origins of ‘modern humans’ to a very specific region in Africa (Chan, et al., 2019), demonstrate how it fell into these inferential pitfalls, and discuss how this can be avoided.

Keywords: Human origins, mitochondrial DNA, gene trees, demographic inference, Africa

Introduction

A paper published by Chan and colleagues titled "Human origins in a southern African palaeo-wetland and first migrations" reported 198 novel whole mitochondrial DNA (mtDNA) genomes and inferred that 'anatomically modern humans' originated in the Makgadikgadi–Okavango palaeo-wetland of southern Africa around 200 thousand years ago (Chan, et al., 2019). This claim relies on weakly informative data. In addition to flawed logic and questionable assumptions, the authors surprisingly disregard recent data and debate on human origins in Africa (Scerri, et al., 2018, Scerri, et al., 2019, Henn, et al., 2018, Schlebusch and Jakobsson, 2018). As a result, the emphatic and far-reaching conclusions of the paper are unjustified.

Weakly informative data

Chan et al. (2019) sampled 198 Southern African individuals, sequenced the full mtDNA genome and merged their data with ~1000 other published mtDNA genomes, mostly from southern Africa. They then proceeded by inferring the phylogenetic structure of the mtDNA genealogy, estimating the ages of various branches on that tree (with a particular focus on branches found in southern Africa), and considering their modern geographic distribution. Mitochondrial DNA can be useful for addressing certain population history questions (e.g. Karmin, et al., 2015), and it is good to focus on under-sampled populations, but the phylogenetic tree at any single non-recombining locus (such as mtDNA) is one outcome of a strongly stochastic genealogical process that is only weakly constrained by population history. In practice, this means it contains useful but limited information about that population history. In particular, information content reduces rapidly as we look further back in time, no matter how many present-day samples are included. Near the time when all mtDNA lineages coalesce to their most recent common ancestor – the part of the tree where Chan and colleagues focus their analysis – mtDNA lineages represent only a small fraction of the thousands of ancestors that present-day humans have at that time, and therefore provides little or no information to distinguish between plausible models of human population history. If we consider another well-resolved single locus phylogenetic tree – that of the Y chromosome – a different phylogeographic structure is apparent (Mendez, et al., 2013).

Modern population genetic approaches take advantage of hundreds of thousands of independent gene trees (obtained from whole-genome sequencing and genome-wide SNP genotyping), gaining several orders of magnitude in statistical power (Kelleher, et al., 2019). While each tree is only weakly constrained by demographic history, by combining many trees with probabilistic approaches it is possible to quantify the likelihood of different demographic scenarios. However, in the absence of informative ancient DNA data from the periods in question, precise inferences on the location of origin of *Homo sapiens* still require strong and likely unjustified assumptions about human mobility in the intervening period, and so must be treated with caution. Genome-wide studies of contemporary and Holocene African populations indicate complex demography, including deep structure and varying migration rates over time (Schlebusch and Jakobsson, 2018, Tishkoff, et al., 2009, Schlebusch, et al., 2012, Schlebusch, et al., 2017, Pickrell, et al., 2012, Skoglund, et al., 2017, Mazet, et al., 2016, Bergström, et al., 2019, Henn, et al., 2011, Fan, et al., 2019). However, despite the far greater information content, no genome-wide study has been able to justify such geographic and temporal precision for modern human origins as that of Chan and colleagues.

Lineages are not populations

While the methods used to estimate the mtDNA tree and its node dates are relatively uncontroversial, it is important to note that there is no natural level at which branches should be assigned haplogroup status, and no special meaning for the age of any arbitrarily defined haplogroup. Chan and colleagues cherry-pick lineages and treat them as representing real populations of prehistoric humans. Lineages in a gene tree correspond to lines of parent-offspring descent, and each node in the tree corresponds to the most recent common ancestor (MRCA) of two lineages. Such lineages demonstrably do not correspond to populations today, and there is little reason to assume that they did in the past. Instead, multiple distinct lineages are usually shared across many different populations, and that degree of sharing is an outcome of random processes that are shaped by population demographic history. There is also little reason to assume that branch split dates in the mtDNA tree correspond directly to population-level splitting events. Instead, lineage split dates will usually predate a population split, often by a substantial margin which is itself shaped by the population size around the time of the split, and subsequent migration rates.

Chan and colleagues also applied Bayesian Skyline analysis to their arbitrarily selected lineages to infer ancestral population size change through time. This imposes a sampling bias that breaks a key assumption of Bayesian Skyline analysis (i.e. the data are from a lineage sample rather than from the model-assumed unstructured population sample), invalidating their reconstructions of population sizes through time (Heller, et al., 2013). Chan et al justify their use of this approach by citing two previous publications (Gandini, et al., 2016, Soares, et al., 2012). However, the latter of these publications (Gandini, et al., 2016) simply cites the earlier one (Soares, et al., 2012) on this issue, both publications merely state that Bayesian Skyline analysis "might nevertheless signal demographic processes" and neither provide any theoretical validation of this claim. Consequently, Chan et al simply repeat this error without credible justification.

Populations are not static

A key issue with inferring geographic origins from contemporary genetic data is the extent to which population history is a 'palimpsest' of overlaid demographic processes (migration, population fission, fusion and size change, etc.) throughout the intervening period. Chan and colleagues' claims rest on the implicit assumption that the locations of present-day individuals represent those of past populations. This assumption of static populations over a period of more than 100,000 years is problematic in the light of well described (by both archaeological and ancient DNA studies) long and short range migrations, population contractions, expansions and replacements - not only in Eurasia, the Americas and Oceania, but also in Africa (e.g. Schlebusch, et al., 2017, Pickrell, et al., 2012, Skoglund, et al., 2017, Bergström, et al., 2019). The assumption of static populations would need to be supported by archaeological, fossil, and ideally ancient DNA evidence; to date no such evidence has been presented for such a deep time period anywhere in Africa. On the contrary, studies of raw material transport demonstrate high levels of mobility among Pleistocene hunter-gatherers that are at odds with the notion of isolated populations over a period of more than 100,000 years (e.g. Scerri, 2017, Brooks, et al., 2018). This issue remains regardless of whether single locus (e.g. mtDNA) data or more strongly informative multi-locus genomic data are considered.

Palaeoanthropology

Chan and colleagues conclude that 'anatomically modern humans' evolved in the Makgadikgadi–Okavango palaeo-wetland of southern Africa around 200,000 years ago. However, fossil hominin data from across Africa indicate the presence of morphology characteristic of *H. sapiens* more than 300 thousand years ago at the opposite end of the continent (Hublin, et al., 2017). The evolution of *H. sapiens* morphology is characterised by a mosaic of different derived and primitive features across time and space, and these data do not indicate a single origin point (Scerri, et al., 2018, Gunz, et al., 2009, Stringer, 2016). Chan and colleagues use archaeological evidence for 'modern human behaviour' from southern Africa between 60 and 100 ka to corroborate their claims of human origins in that region more than 100 ka earlier (Henshilwood, 2012). However, evidence of complex culture is found in other African regions around this period (e.g. (Scerri, 2017, Brooks, et al., 2018)). Together, current palaeoanthropological data do not support claims that one single area of Africa was the 'homeland' of contemporary *H. sapiens.*

Climate reconstructions

Using climate reconstructions to contextualise inferences based on genetic data is a commendable aim. However, while Chan and colleagues argue that wetlands were a stable habitat, they give no evidence that wetlands provide ecologies that are well-suited to ancient humans, and do not consider palaeo-wetlands elsewhere in Africa (Wrangham, et al., 2009). Without addressing these issues, it is difficult to see how any type of climate reconstruction can provide a test of their origins model. Other ecologies, such as grasslands, savannahs and Mediterranean biomes were also inhabited by *H. sapiens* for long periods of time, and were therefore clearly suitable for long term *H. sapiens* occupation (Scerri, et al., 2018, Hublin, et al., 2017). An equally plausible alternative hypothesis (Potts, 1998) is that semi-desert conditions would be a suitable habitat for humans, because of their strong variations in climate on seasonal/annual timescales. The climate model that Chan and colleagues employ would not capture those features well. General overviews of African palaeoclimatic and palaeoanthropological data for the relevant time period exists (e.g. Blome, et al., 2012). Relationships between climate and human demography are best explored through spatially explicit modelling (e.g. Eriksson and Manica, 2012) rather than *post-hoc* explanations which ignore other relevant data from across Africa.

Concluding remarks

Although admittedly common in human evolutionary studies, a particularly troubling aspect of Chan and colleagues' study is that the authors make no attempt to quantify how well their data can be explained either by simpler models of population history than the one they claim, or by equally complex but dramatically different models, such as those involving an origin of *H. sapiens* elsewhere in Africa, or indeed, in multiple locations throughout Africa.

One hundred years ago, in the first study using biomarkers, Hirschfeld and Hirschfeld (1919) typed different populations from around the world at a single genetic locus (ABO blood groups). While the study was revolutionary, the authors concluded that there were two ancestral human 'races', corresponding to the A and B blood groups, with two different geographical origins that subsequently mixed globally. Later developments in population genetics, genomics, archaeology and palaeoanthropology have revealed a very different and more complex picture. But the dangers of basing conclusions on limited data and problematic assumptions remain just as relevant today. Presentation of such studies, with predictably high-profile media and broader scientific community attention, and simple yet questionable results, does not serve the science of human origins, nor its broader dissemination.

Competing Interest Statement

The authors declare no competing interests

Author Contributions

CMS, LL and MGT drafted and wrote the manuscript with contributions from other authors. All authors read and approved the manuscript

References

Chan, E.K.F., Timmermann, A., Baldi, B.F., Moore, A.E., Lyons, R.J., Lee, S.S., Kalsbeek, A.M.F., Petersen, D.C., Rautenbach, H., Fortsch, H.E.A., Bornman, M.S.R., Hayes, V.M., 2019. Human origins in a southern African palaeo-wetland and first migrations, Nature.

Scerri, E.M.L., Thomas, M.G., Manica, A., Gunz, P., Stock, J.T., Stringer, C., Grove, M., Groucutt, H.S., Timmermann, A., Rightmire, G.P., d'Errico, F., Tryon, C.A., Drake, N.A., Brooks, A.S., Dennell, R.W., Durbin, R., Henn, B.M., Lee-Thorp, J., deMenocal, P., Petraglia, M.D., Thompson, J.C., Scally, A., Chikhi, L., 2018. Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter?, Trends Ecol Evol 33, 582-594.

Scerri, E.M.L., Chikhi, L., Thomas, M.G., 2019. Beyond multiregional and simple out-of-Africa models of human evolution, Nat Ecol Evol 3, 1370-1372.

Henn, B.M., Steele, T.E., Weaver, T.D., 2018. Clarifying distinct models of modern human origins in Africa, Curr Opin Genet Dev 53, 148-156.

Schlebusch, C.M., Jakobsson, M., 2018. Tales of Human Migration, Admixture, and Selection in Africa, Annu Rev Genomics Hum Genet 19, 405-428.

Karmin, M., Saag, L., Vicente, M., Wilson Sayres, M.A., Jarve, M., Talas, U.G., Rootsi, S., Ilumae, A.M., Magi, R., Mitt, M., Pagani, L., Puurand, T., Faltyskova, Z., Clemente, F., Cardona, A., Metspalu, E., Sahakyan, H., Yunusbayev, B., Hudjashov, G., DeGiorgio, M., Loogvali, E.L., Eichstaedt, C., Eelmets, M., Chaubey, G., Tambets, K., Litvinov, S., Mormina, M., Xue, Y., Ayub, Q., Zoraqi, G., Korneliussen, T.S., Akhatova, F., Lachance, J., Tishkoff, S., Momynaliev, K., Ricaut, F.X., Kusuma, P., Razafindrazaka, H., Pierron, D., Cox, M.P., Sultana, G.N., Willerslev, R., Muller, C., Westaway, M., Lambert, D., Skaro, V., Kovacevic, L., Turdikulova, S., Dalimova, D., Khusainova, R., Trofimova, N., Akhmetova, V., Khidiyatova, I., Lichman, D.V., Isakova, J., Pocheshkhova, E., Sabitov, Z., Barashkov, N.A., Nymadawa, P., Mihailov, E., Seng, J.W., Evseeva, I., Migliano, A.B., Abdullah, S., Andriadze, G., Primorac, D., Atramentova, L., Utevska, O., Yepiskoposyan, L., Marjanovic, D., Kushniarevich, A., Behar, D.M., Gilissen, C., Vissers, L., Veltman, J.A., Balanovska, E., Derenko, M., Malyarchuk, B., Metspalu, A., Fedorova, S., Eriksson, A., Manica, A., Mendez, F.L., Karafet, T.M., Veeramah, K.R., Bradman, N., Hammer, M.F., Osipova, L.P., Balanovsky, O., Khusnutdinova, E.K., Johnsen, K., Remm, M., Thomas, M.G., Tyler-Smith, C., Underhill, P.A., Willerslev, E., Nielsen, R., Metspalu, M., Villems, R., Kivisild, T., 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture, Genome research 25, 459-466.

Mendez, F.L., Krahn, T., Schrack, B., Krahn, A.M., Veeramah, K.R., Woerner, A.E., Fomine, F.L., Bradman, N., Thomas, M.G., Karafet, T.M., Hammer, M.F., 2013. An African American paternal lineage adds an extremely ancient root to the human Y chromosome phylogenetic tree, Am J Hum Genet 92, 454-459.

Kelleher, J., Wong, Y., Wohns, A.W., Fadil, C., Albers, P.K., McVean, G., 2019. Inferring whole-genome histories in large population datasets, Nat Genet 51, 1330-1338.

Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.M., Doumbo, O., Ibrahim, M., Juma, A.T., Kotze, M.J., Lema, G., Moore, J.H., Mortensen, H., Nyambo, T.B., Omar, S.A., Powell, K., Pretorius, G.S., Smith, M.W., Thera, M.A., Wambebe, C., Weber, J.L., Williams, S.M., 2009. The genetic structure and history of Africans and African Americans, Science 324, 1035-1044.

Schlebusch, C.M., Skoglund, P., Sjodin, P., Gattepaille, L.M., Hernandez, D., Jay, F., Li, S., De Jongh, M., Singleton, A., Blum, M.G., Soodyall, H., Jakobsson, M., 2012. Genomic Variation in Seven Khoe-San Groups Reveals Adaptation and Complex African History, Science 338, 374-379.

Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R., Steyn, M., Soodyall, H., Lombard, M., Jakobsson, M., 2017. Ancient genomes from southern Africa pushes modern human divergence beyond 260,000 years ago, bioRxiv.

Pickrell, J.K., Patterson, N., Barbieri, C., Berthold, F., Gerlach, L., Guldemann, T., Kure, B., Mpoloka, S.W., Nakagawa, H., Naumann, C., Lipson, M., Loh, P.R., Lachance, J., Mountain, J., Bustamante, C.D., Berger, B., Tishkoff, S.A., Henn, B.M., Stoneking, M., Reich, D., Pakendorf, B., 2012. The genetic prehistory of southern Africa, Nat Commun 3, 1143.

Skoglund, P., Thompson, J.C., Prendergast, M.E., Mittnik, A., Sirak, K., Hajdinjak, M., Salie, T., Rohland, N., Mallick, S., Peltzer, A., Heinze, A., Olalde, I., Ferry, M., Harney, E., Michel, M., Stewardson, K., Cerezo-Roman, J.I., Chiumia, C., Crowther, A., Gomani-Chindebvu, E., Gidna, A.O., Grillo, K.M., Helenius, I.T., Hellenthal, G., Helm, R., Horton, M., Lopez, S., Mabulla, A.Z.P., Parkington, J., Shipton, C., Thomas, M.G., Tibesasa, R., Welling, M., Hayes, V.M., Kennett, D.J., Ramesar, R., Meyer, M., Paabo, S., Patterson, N., Morris, A.G., Boivin, N., Pinhasi, R., Krause, J., Reich, D., 2017. Reconstructing Prehistoric African Population Structure, Cell 171, 59-71 e21.

Mazet, O., Rodriguez, W., Grusea, S., Boitard, S., Chikhi, L., 2016. On the importance of being structured: instantaneous coalescence rates and human evolution--lessons for ancestral population size inference?, Heredity (Edinb) 116, 362-371.

Bergström, A., McCarthy, S., Hui, R., Almarri, M.A., Ayub, Q., Danecek, P., Chen, Y., Felkel, S., Hallast, P., Kamm, J., Blanché, H., , e.a., 2019. Insights into human genetic variation and population history from 929 diverse genomes, bioRxiv Jan 1:674986.

Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodriguez-Botigue, L., Ramachandran, S., Hon, L., Brisbin, A., Lin, A.A., Underhill, P.A., Comas, D., Kidd, K.K., Norman, P.J., Parham, P., Bustamante, C.D., Mountain, J.L., Feldman, M.W., 2011. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans, Proc Natl Acad Sci U S A 108, 5154-5162.

Fan, S., Kelly, D.E., Beltrame, M.H., Hansen, M.E.B., Mallick, S., Ranciaro, A., Hirbo, J., Thompson, S., Beggs, W., Nyambo, T., Omar, S.A., Meskel, D.W., Belay, G., Froment, A., Patterson, N., Reich, D., Tishkoff, S.A., 2019. African evolutionary history inferred from whole genome sequence data of 44 indigenous African populations, Genome Biol 20, 82.

Heller, R., Chikhi, L., Siegismund, H.R., 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history, PLoS One 8, e62992.

Gandini, F., Achilli, A., Pala, M., Bodner, M., Brandini, S., Huber, G., Egyed, B., Ferretti, L., Gómez-Carballa, A., Salas, A., 2016. Mapping human dispersals into the Horn of Africa from Arabian Ice Age refugia using mitogenomes, Scientific reports 6, 25472.

Soares, P., Alshamali, F., Pereira, J.B., Fernandes, V., Silva, N.M., Afonso, C., Costa, M.D., Musilova, E., Macaulay, V., Richards, M.B., Cerny, V., Pereira, L., 2012. The Expansion of mtDNA Haplogroup L3 within and out of Africa, Mol Biol Evol 29, 915-927.

Scerri, E.M.L., 2017. The North African Middle Stone Age and its place in recent human evolution, Evol Anthropol 26, 119-135.

Brooks, A.S., Yellen, J.E., Potts, R., Behrensmeyer, A.K., Deino, A.L., Leslie, D.E., Ambrose, S.H., Ferguson, J.R., d'Errico, F., Zipkin, A.M., Whittaker, S., Post, J., Veatch, E.G., Foecke, K., Clark, J.B., 2018. Long-distance stone transport and pigment use in the earliest Middle Stone Age, Science 360, 90-94.

Hublin, J.J., Ben-Ncer, A., Bailey, S.E., Freidline, S.E., Neubauer, S., Skinner, M.M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., Gunz, P., 2017. New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens, Nature 546, 289-292.

Gunz, P., Bookstein, F.L., Mitteroecker, P., Stadlmayr, A., Seidler, H., Weber, G.W., 2009. Early modern human diversity suggests subdivided population structure and a complex out-of-Africa scenario, Proc Natl Acad Sci U S A 106, 6094-6098.

Stringer, C., 2016. The origin and evolution of Homo sapiens, Philos Trans R Soc Lond B Biol Sci 371.

Henshilwood, C.S., 2012. The Still Bay and Howiesons Poort: ‘Palaeolithic’ techno-traditions in southern Africa, J World Pre-hist 25 205–237.

Wrangham, R., Cheney, D., Seyfarth, R., Sarmiento, E., 2009. Shallow-water habitats as sources of fallback foods for hominins, Am J Phys Anthropol 140, 630-642.

Potts, R., 1998. Variability selection in hominid evolution, Evolutionary Anthropology 7, 81-96.

Eriksson, A., Manica, A., 2012. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins, Proc Natl Acad Sci U S A 109, 13956-13960.

Blome, M.W., Cohen, A.S., Tryon, C.A., Brooks, A.S., Russell, J., 2012. The environmental context for the origins of modern human diversity: a synthesis of regional variability in African climate 150,000-30,000 years ago, J Hum Evol 62, 563-592.

Hirschfeld, L., Hirschfeld, H., 1919. Serological Differences between the Blood of Different Races, The Lancet 194, 675-679.