X-chromosomal diversity may, or may not, reflect climate

Zarus Cenac
Independent Researcher
Email: z.cenac@alumni.ucl.ac.uk

Abstract

Previous research indicates that a climatic signal is present in mitochondrial diversity, but absent in other genetic diversities (e.g., X-chromosomal) and cranial diversity. Such research has included analysis which adjusted diversity for distance from Africa (i.e., to account for the global expansion), and seeing whether this adjusted diversity is associated with climate. Detection of a relationship between adjusted diversity and climate may be affected by the choice of African location from which distance is measured. To bypass this potential effect of location, some analyses in the present research only featured populations who are located outside of Africa. The present research (Studies 1 and 2) used various diversities which were sourced from previous research. Autosomal, X-chromosomal, and cranial diversities were adjusted for distance from Africa. This adjustment was not used regarding Y-chromosomal diversity; a preprint suggested that Y-chromosomal diversity does not signify expansion from Africa. In Study 1, adjusted X-chromosomal diversity increased with minimum temperature. Other adjusted diversities were not related to minimum temperature. In contrast to previous research, an association was indicated between the ratio of X- to Y-chromosomal diversity and climate in Study 2, which may suggest a relationship between sex-biased migration and climate—this relationship could drive the climatic signal in X-chromosomal diversity. Conscious of replicability, Study 2 examined whether a climatic signal is present in X-chromosomal diversity when an alternative dataset is used; a signal was not detected. Therefore, it seems unclear whether X-chromosomal diversity is associated with climate.

Introduction

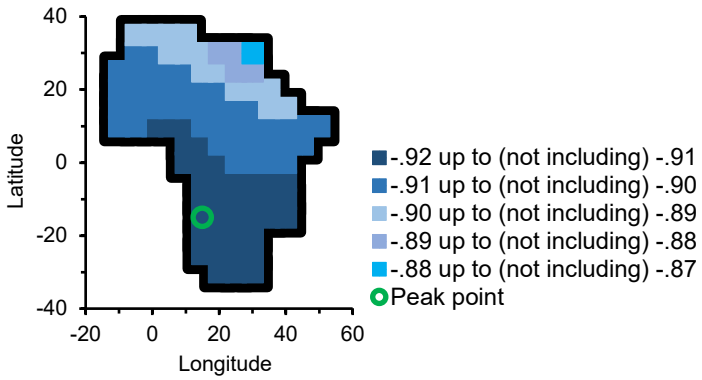
Climate and global expansion

Research has considered factors which may affect human biological diversity, e.g., by examining climate and expansion from Africa (e.g., Balloux et al., 2009; Betti et al., 2012). In modern humans, population-level diversity falls as geographical distance from Africa rises (Betti et al., 2012; Prugnolle et al., 2005; von Cramon-Taubadel & Lycett, 2008), indicating a worldwide expansion (through population bottlenecks) from Africa (Manica et al., 2007; Ramachandran et al., 2005; von Cramon-Taubadel & Lycett, 2008; see Figure 1). As for climate, mitochondrial DNA is known to be have been affected by climate (Balloux et al., 2009); mitochondria are acknowledged to play a role in climatic adaptation (Grover-Thomas et al., 2026). Indeed, in Balloux et al. (2009), climate (minimum temperature) was associated with mitochondrial diversity. This association persisted when

mitochondrial diversity was adjusted for linear distance from Africa—adjusted mitochondrial diversity rose with minimum temperature (Balloux et al., 2009; see Supplementary Note 1). No such association was apparent for adjusted autosomal, X-, or Y-chromosomal diversities (Balloux et al., 2009) (see Supplementary Note 2). When it comes to skeletal diversity, minimum temperature is related to the shape diversity of the femur and tibia (Betti et al., 2012), whilst a relationship with climate fails to be supported for cranial form diversity (Betti et al., 2009; see Supplementary Note 3) and pelvic shape diversity (Betti et al., 2012). Therefore, a climatic signal seems to be evident in some biological diversities (Balloux et al., 2009; Betti et al., 2012) yet not evident in others (Balloux et al., 2009; Betti et al., 2009, 2012).

Figure 1

X-Chromosomal Diversity and the African Origin of Global Expansion



Note. In order to indicate where expansion started, various locations can be used (one at a time) as if each was the origin, and it can be seen how well distance from that location is related to diversity (e.g., Manica et al., 2007). Support for a location being the origin of expansion is indicated by the (relative) extent to which diversity declines as distance from a location increases—a stronger decline (and therefore a stronger expansion signal) bodes well for a location being the origin (e.g., Ramachandran et al., 2005; von Cramon-Taubadel & Lycett, 2008). The location which results in the most negative decline can be called the *peak point* (e.g., Cenac, 2022). Like has been done with correlation coefficients for relationships between genetic diversity and distance from Africa (Luca et al., 2011; Ramachandran et al., 2005), Figure 1 presents correlation coefficients (*r*—blue areas) regarding X-chromosomal microsatellite heterozygosity and distance from Africa (coefficients are from Cenac, 2023a, who used diversities from Balloux et al., 2009, and distances in Cenac, 2022). The *peak point* is the location giving the lowermost correlation coefficient (Cenac, 2022) (green circle). Figure 1 adjusts Figure 4 in Cenac (2023b) which employed coordinates (in Africa) that are contained in a figure by Betti et al. (2013) (Figure 4 in Cenac, 2023b, did not feature correlation coefficients, and it had some basis in Cenac, 2022).

Diversity adjusted for distance

When studying whether diversity is related to climate, diversity can be adjusted for distance from Africa, and it can be seen if this adjusted diversity is related to climate (Balloux et al., 2009). However, when using populations pulled from across the world, the relationship between diversity

(adjusted for linear distance from Africa) and climate can be affected by the location in Africa which distances are measured from (Figure 2A and 2C). Therefore, a supposed lack of relationship with climate (e.g., for adjusted X-chromosomal diversity in Balloux et al., 2009) could be a product of analysis having used distances from a location which just so happens to lead to a weaker correlation coefficient.

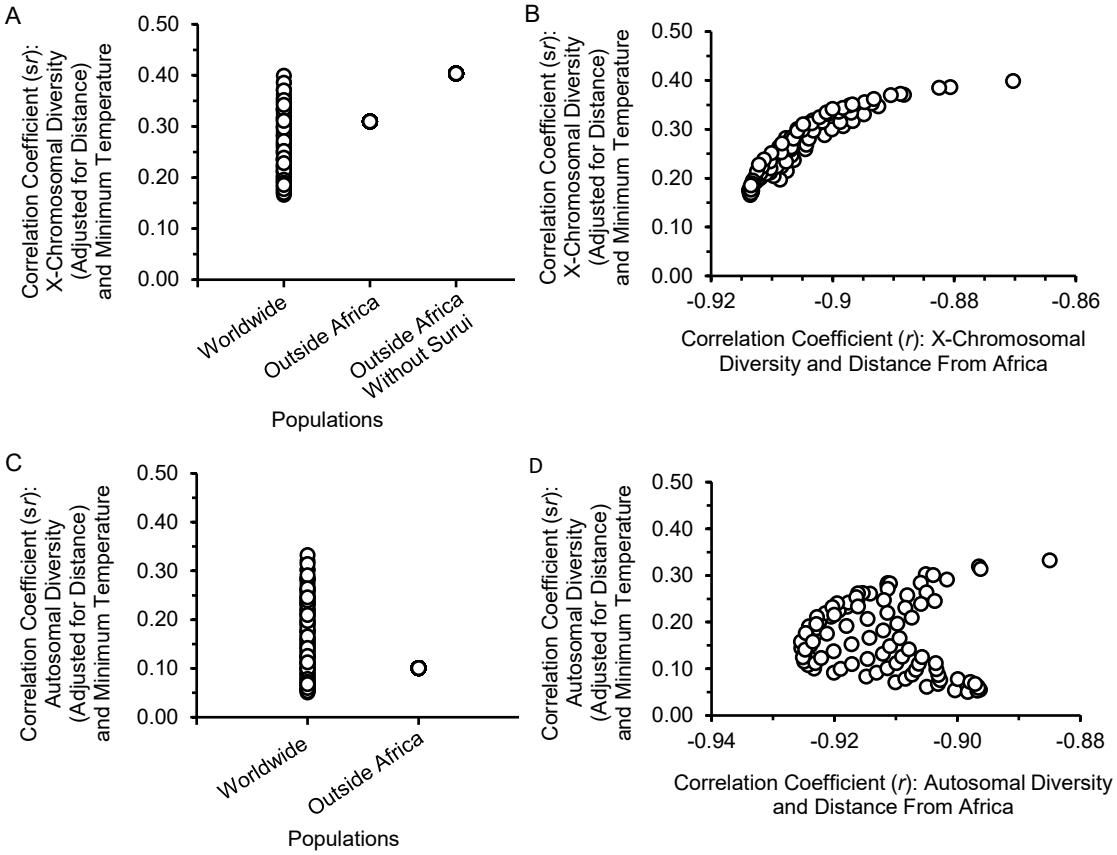
Distances from *which* location should be used to adjust diversity? The location whose distances deliver the strongest fall of diversity (e.g., Betti et al., 2013; Ramachandran et al., 2005) has been called the *peak point* (Cenac, 2022). The magnitude of the decline in diversity indicates the strength of the expansion signal (Ramachandran et al., 2005; von Cramon-Taubadel & Lycett, 2008). Therefore, to remove the influence of expansion on diversity, it would make sense to adjust diversity for distance from the peak point. However, different diversities can have different peak points (Cenac, 2022); did Balloux et al. (2009) use distance from diversity-specific peak points when they adjusted diversities? Balloux et al. (2009) used distance from a location (in Africa) from Manica et al. (2007). That location was arrived at through autosomal microsatellite heterozygosity and cranial form diversity (Balloux et al., 2009; Manica et al., 2007). The location would seem to be reasonably similar to the peak points for mitochondrial diversity and X-chromosomal diversity (in Cenac, 2022), less similar for the peak point (linear trend) for autosomal microsatellite heterozygosity (in Ramachandran et al., 2005), even less similar for autosomal SNP haplotype heterozygosity, but markedly different for Y-chromosomal microsatellite heterozygosity (in Cenac, 2022). Indeed, whilst peak points for autosomal, mitochondrial, and X-chromosomal diversity are in Africa (Cenac, 2022; Ramachandran et al., 2005), the peak point for Y-chromosomal microsatellite heterozygosity is not (Cenac, 2022).

Indeed, whereas Y-chromosomal microsatellite heterozygosity does decrease linearly as distance from Africa rises (Balloux et al., 2009), the strongest decline is not found when distance is from Africa; the peak point is located in Asia—Y-chromosomal microsatellite heterozygosity does not appear to portray expansion worldwide from Africa (Cenac, 2022). Therefore, when exploring if Y-chromosomal microsatellite heterozygosity is associated with climate, rather than adjusting diversity for distance from Africa (Balloux et al., 2009), it *may* make more sense to adjust diversity for distance from Asia. Nevertheless, global expansion from Asia seems unlikely (e.g., Li et al., 2008; Ramachandran et al., 2005). Therefore, all in all, Balloux et al. (2009) are likely to have greatly adjusted diversities for a global expansion signal except when it came to Y-chromosomal diversity.

Nonetheless, regarding populations who are located across the world, peak points (when they are in Africa) appear to be greatly determined by African populations (Cenac, 2022). This is problematic because few African populations *may* feature in data which are used in the calculation of a peak point (e.g., seven African populations regarding X-chromosomal and autosomal haplotype diversity) (Cenac, 2022). And so, it seems unclear if it is wise to simply use the peak point when diversity is adjusted for distance.

Figure 2

Diversity, Climate, and Distance from Africa



Note. Figure 2 used distances from 99 locations in Africa (Cenac, 2022), X-chromosomal microsatellite heterozygosities, autosomal single nucleotide polymorphism (SNP) haplotype heterozygosities, and minimum temperatures (Balloux et al., 2009). Figure 2A and 2B are with respect to X-chromosomal diversity, whilst Figure 2C and 2D are regarding autosomal diversity. In some instances, like in Balloux et al. (2009), diversity was adjusted for linear distance from Africa (see the *y*-axes of 2A, 2B, 2C, and 2D). With respect to minimum temperature (Balloux et al., 2009) and adjusted diversity, semi-partial Pearson correlation coefficients were calculated (*y*-axes). This procedure was done with and without populations in Africa. For X-chromosomal diversity, in the absence of the populations in Africa (Figure 2A), there was an atypical datapoint. The Surui population in the HGDP-CEPH data (Balloux et al., 2009) was this datapoint. Regarding populations from across the world, Balloux et al. used distance from a particular location in Africa—the present research observed quite similar R^2s (compared to Balloux et al.) when using distances in the vicinity of that particular location—the corresponding s*r*s are generally amongst the lowest when populations are from across the world in Figure 2A and 2C. For Figure 2B and 2D, populations from across the world were used—a significance test was not run (there likely would be nonindependence—see Judd et al., 2009, regarding nonindependence), although patterns do seem to be visually indicated. For some diversities, a weaker correlation coefficient may indicate that the signal of expansion from Africa has been controlled for better—when it comes to X-chromosomal diversity, it appears that correlation coefficients may become weaker as more of the expansion signal is removed (Figure 2B). However, this is likely not the case for all diversities as it seems to not be the case for autosomal SNP haplotype heterozygosity (Figure 2D).

Therefore, when adjusting diversity, there is ambiguity concerning which location the distances should be from. An alternative approach could be to only use populations outside of Africa—when only those populations are used, the correlation coefficient does not vary according to the location in Africa that is used as the origin (Figure 2A and 2C). However, to comprehensively understand if there is a climatic signal in diversity, it would not be desirable to exclude an entire continent.

Cranial form diversity

Betti et al. (2009) explored climate and diversity in a different way than Balloux et al. (2009) did. Betti et al. used coordinates across the Earth like each was an origin. For each origin, they started with a model for predicting cranial form diversity, using distance (to populations) terms and climatic terms. For each origin, Betti et al. (2009) found the most effective model for predicting diversity. Out of all the origins, the origin which had a model that best predicted diversity was an origin in Africa, and climate was not included in that model (Betti et al., 2009). This finding happened with respect to male crania, and female crania too (Betti et al., 2009). However, as mentioned above, the peak point may be particularly influenced by African populations (Cenac, 2022), which could be worth considering in the event that models using other *origins* included climatic terms in Betti et al. (2009).

The present research

In two studies, it was examined whether diversities appear to have a climatic signal when the signal of expansion is controlled for (see Supplementary Note 4). Prior research has used minimum temperature as a climatic variable (Balloux et al., 2009; Betti et al., 2009, 2012; Manica et al., 2007; see Supplementary Note 5), and minimum temperature was used as such in the present research. Using populations from across the world, earlier research adjusted genetic diversities for distance from Africa (distance from the same location in Africa for each diversity), and examined if these adjusted diversities are related to minimum temperature (Balloux et al., 2009); this approach was broadly taken in the present research, but with some differences. To counter a potential influence of location in Africa on whether relationships with climate are apparent (e.g., Figure 2), several analyses solely featured populations outside of Africa. Those analyses concerned autosomal microsatellite heterozygosity, autosomal SNP haplotype heterozygosity, X-chromosomal microsatellite heterozygosity, and cranial form diversity. Certain analyses (global/worldwide analyses) featured African populations. Distances from the peak point specific to the relevant diversity (Cenac, 2022, 2023a) were used in the global analyses. These analyses featured the aforementioned autosomal and X-chromosomal diversities, as well as (for completeness) Y-chromosomal microsatellite heterozygosity. In Study 1, a climatic signal was indicated in X-chromosomal diversity.

Previous research found a climatic signal in mitochondrial in *two* datasets (Balloux et al., 2009); exploring replicability in science is important (e.g., Nosek & Errington, 2020) and, in the present research, a second study (Study 2) was undertaken as a follow-up to Study 1 in order to test

whether the signal in X-chromosomal diversity is indicated in an alternative dataset (see Balloux et al., 2009; Mallick et al., 2016).

Research has occurred with respect to sex-biased migration in human history (e.g., Keinan & Reich, 2010; Rasteiro et al., 2012). As a way of indicating whether a relationship between sex-biased migration and climate may underlie the climatic signal in mitochondrial diversity, prior research examined whether the ratio in diversity between sex chromosomes is related to minimum temperature (Balloux et al., 2009); Study 2 examined whether the ratio (after adjusting X-chromosomal diversity for distance from Africa) is related to minimum temperature, and this was done for the purpose of exploring whether sex-biased migration and climate are related because their relationship could provide a reason for why, in Study 1, X-chromosomal diversity was related to climate.

Method

Study 1: Genetic and cranial diversities

Data

Study 1 used genetic diversities in Balloux et al. (2009). These diversities were heterozygosities (51 populations from the HGDP-CEPH) regarding autosomal microsatellites, autosomal SNP haplotypes, X-chromosomal microsatellites, and Y-chromosomal microsatellites (Balloux et al., 2009). Balloux et al. acquired the haplotype diversities from Li et al. (2008). Additionally, the form diversity of male crania (90 populations) and female crania (34 populations) outside of Africa (Betti et al., 2009) was used (regarding cranial data, Betti et al., 2009, cite Manica et al., 2007).

Geographical distances were from prior research (Betti et al., 2009; Cenac, 2022), some of which (in Cenac, 2022) included distances from i) 99 locations in Africa from Betti et al. (2013), and ii) 32 global locations found in von Cramon-Taubadel and Lycett (2008). Geographical distances (Betti et al., 2009; Cenac, 2022) from peak points (Betti et al., 2009; Cenac, 2022, 2023a) were used to adjust diversities. These distances (Betti et al., 2009; Cenac, 2022) were from peak points which are i) in Africa for cranial form diversity (Betti et al., 2009), X-chromosomal microsatellite heterozygosity, autosomal SNP haplotype heterozygosity (Cenac, 2022), and autosomal microsatellite heterozygosity (Cenac, 2023a), and (for completeness) ii) in Asia for Y-chromosomal microsatellite heterozygosity (Cenac, 2022). Because peak points can vary between diversities (Cenac, 2022), the distances (from peak points) used for adjusting diversities were the ones relevant/specific to the type of diversity. Distances from peak points were utilised even when analyses only consisted of populations outside Africa (despite reasoning in the *Introduction*). For peak point coordinates, see Table S1 (or references given in Table S1).

Regarding genetic data, it was known from beforehand which continent populations are located within (Cenac, 2022); it was clear which populations are inside Africa, and which are not. Labelling employed in Betti et al. (2009) was used to identify which populations featured in their study are outside of Africa.

Minimum temperatures shown in previous research (Balloux et al., 2009; Betti et al., 2009) were used (both Balloux et al., 2009, and Betti et al., 2009, sourced minimum temperatures from WORLDCLIM, and they cited Hijmans et al., 2005, with respect to WORLDCLIM).

Analysis

Analysis took place in R Version 4.0.5 (R Core Team, 2021) and Microsoft Excel. ppcor (Kim, 2015) was used for running semi-partial correlation tests, except for the semi-partial Pearson correlation tests used regarding autosomal microsatellite heterozygosity, and the semi-partial Spearman correlation test employed regarding the diversity of male crania.

Autosomal microsatellite heterozygosity may *appear* to fall linearly when distance from Africa increases (Prugnolle et al., 2005), but this type of diversity aligns more strongly with a non-linear trend (quadratic) than a linear decline (Cenac, 2023a). Therefore, autosomal microsatellite heterozygosity was adjusted for a quadratic (rather than a linear) relationship with distance from Africa. It was then seen if the adjusted autosomal diversity is related to minimum temperature. Semi-partial Pearson correlation coefficients concerning adjusted autosomal diversity and minimum temperature were calculated (for populations globally, and only outside Africa). The correlation coefficients were converted to *t*-statistic values using a formula in Kim (2015). A *p*-value can be found for a *t*-statistic value (e.g., Kim, 2015), and *p*-values were indeed calculated for the converted *t*-values.

In Betti et al. (2009), there was a non-linear association between distance from Africa and cranial form diversity for male crania (105 populations from across the globe). The model used in Betti et al. (2009) for predicting diversity from distance from the peak point had an intercept, a linear term, and a cubic term (see Supplementary Note 6); a variant of their model was used to calculate residuals in the present study (the residuals being diversity adjusted for distance from Africa)—the same distances were used as in Betti et al., an intercept was also utilised, as was a linear term, and also a cubic term when trying to predict diversity from distance, but populations were solely outside of Africa. The semi-partial Spearman correlation coefficient for adjusted cranial diversity and minimum temperature was determined. The formula in Kim (2015) was used for converting the correlation coefficient to a *t*-value, for which (like with autosomal diversity) a *p*-value was calculated.

Linear declines (with increasing distance from Africa) are shown in autosomal SNP haplotype heterozygosity and X-chromosomal diversity (Balloux et al., 2009), with a quadratic relationship attaining no support (over a linear relationship) (Cenac, 2023a). And so, adjusting those diversities for linear distance from Africa (e.g., Balloux et al., 2009) would seem to be acceptable. Therefore, analysis in the present study set out to adjust those diversities for linear distances.

Y-chromosomal microsatellite heterozygosity has a peak point in Asia (and indicates an origin of expansion which possibly is exclusive to Asia), with a linear decline from there (Cenac, 2022), and an absence of support for a non-linear relationship (Cenac, 2023a). Therefore, it was planned, for the

present study, to adjust Y-chromosomal diversity for linear distance from Asia (whereas Balloux et al., 2009, adjusted for linear distance from Africa).

There are various ways of assessing whether datapoints are atypical (e.g., Field, 2013; Judd et al., 2009). In the present study, datapoints which have *z*-scores of residuals over |3.29| were (conservatively) noted as being atypical (e.g., Field, 2013). The Surui population (Balloux et al., 2009) had an atypical standardised residual in some analyses (see *Results*). In the absence of atypical datapoints, it was seen whether heteroscedasticity was indicated. It was also seen if the presence of heteroscedasticity was resolved in the absence of any one population. To visually assess whether heteroscedasticity was present (in parametric analysis), adjusted diversity (*y*-axis) was placed against minimum temperature (*x*-axis) in graphs. Several graphs of this sort are presented in Balloux et al. (2009).

To see whether positive spatial autocorrelation was present amongst residuals, the three-step method of Chen (2016) was employed. This was done utilising a spreadsheet which is available in Chen (2016). Spatial Durbin-Watson values (spatial *DW*s) are produced in the three-step method, and Durbin-Watson bounds are applicable to the spatial *DW* (Chen, 2016). Consequently, 5% Durbin-Watson bounds (Savin & White, 1977) were referred to in order to assess if positive spatial autocorrelation was at hand. However, spatial autocorrelation would enhance the Type I error rate (e.g., Deblauwe et al., 2012), therefore, when significant relationships were not evident (e.g., Figure 3), positive spatial autocorrelation would not *need* to be tested for.

The Bonferroni correction (reviewed in Abdi, 2010, alongside other methods of adjusting *p*-values) was used on *p*-values in the current study; *p*-values were multiplied by five (as there were five types of diversity of interest in this study) regarding the genetic diversities, but multiplied by 10 concerning cranial diversity (multiplied by five given the five types of diversity, and multiplied by two because there were separate analyses of males and females).

Study 2: X-chromosomal diversity and the sex chromosome diversity ratio

Data

Study 2 explored whether there is a climatic signal in X-chromosomal diversity. The study used X-chromosomal expected heterozygosity (calculated from base pair dissimilarities) of persons whose sex is female (77 populations outside of Africa) from the Simons Genome Diversity Project (SGDP) (Mallick et al., 2016). The procedure of Williams (2011) was used for the purpose of calculating population geographical distances (from Africa). To calculate those distances, population coordinates (Mallick et al., 2016), waypoints (Cenac, 2022; von Cramon-Taubadel & Lycett, 2008), and an origin (von Cramon-Taubadel & Lycett, 2008) were used. Minimum temperatures (each month from 1981 to 2025) were obtained from NASA POWER (2026) using the population coordinates. The downloaded data (NASA POWER, 2026) included a column featuring the lowest minimum temperature (from the 12 months) for each year. These lowest minimum temperatures were averaged (across the years 1981

to 2025) in order to arrive at the *minimum temperature variable* which was used in correlation tests in the current study.

Additionally, using HGDP-CEPH data (X- and Y-chromosomal diversities) and minimum temperatures (Balloux et al., 2009), the study expanded upon Balloux et al. (2009) who looked into whether the ratio between the diversities of sex chromosomes is related to minimum temperature; the present study adjusted X-chromosomal diversity for distance from Africa before dividing the result by Y-chromosomal diversity—it was then seen if the resulting ratio correlated with minimum temperature.

Analysis

Analysis utilised Spatial Analysis in Macroecology (SAM) Version 4.0 (Rangel et al., 2010) (available from https://www.ecoevol.ufg.br/sam/), R Version 4.0.5 (R Core Team, 2021), and Microsoft Excel. In order to run the semi-partial correlation test between adjusted X-chromosomal diversity (i.e., adjusted for distance from Africa) and minimum temperature, ppcor (Kim, 2015) was used initially.

As with Study 1, in parametric analysis, datapoints were highlighted as atypical when *z*-scored residuals surpassed |3.29| (e.g., Field, 2013). In instances where X-chromosomal diversity (SGDP data) was only available for one person (in a population), the diversity of that person was simply used to represent their population. In instances where the diversity of more than one person was available for a population, an average was calculated to represent their population. As in Study 1, data were examined for the presence of positive spatial autocorrelation (Chen, 2016; Savin & White, 1977). However, positive spatial autocorrelation was not tested for when examining whether X-chromosomal diversity is correlated with distance from Africa. Instead, SAM Version 4.0, which adjusts the degrees of freedom (Rangel et al., 2010), was cautiously used in the event of spatial autocorrelation being present. A Bonferroni correction (see Abdi, 2010) was applied.

Results

Study 1: Genetic and cranial diversities

Outside of Africa

For populations outside of Africa, X-chromosomal microsatellite heterozygosity (adjusted for distance from Africa) positively correlated with minimum temperature, $sr(40) = .40$, $p = .040$, spatial $DW = 1.91$. This result was found following Surui not being included due to their atypical standardised residual ($z = -3.70$) (as in part of Figure 2A). Adjusted autosomal SNP haplotype heterozygosity had no correlation with minimum temperature, $sr(41) = .10$, $p = 1.00$, spatial $DW = 1.71$. The same can be said for adjusted autosomal microsatellite heterozygosity, $sr(39) = .25$, $p = .57$, spatial $DW = 1.92$ (in the absence of Surui due to their standardised residual, $z = -3.41$). As for adjusted cranial form diversity, it was not found to have a correlation with minimum temperature for females, $sr_s(31) = -.13$, $p = 1.00$, or for males, $sr_s(86) = -.02$, $p = 1.00$.

Worldwide

X-chromosomal microsatellite heterozygosity (adjusted for distance from Africa) was found to have a positive association with minimum temperature, sr(46) = .39, p = .031, spatial DW = 1.92. However, there were caveats, with this correlation being calculated in the absence of two populations. When using 51 populations, Surui had an atypically low standardised residual (z = -3.52). Whether with or without Surui, heteroscedasticity was indicated. In the absence of Surui, out of the 50 remaining populations, any one of the populations was removed to see if any particular population may be driving the heteroscedasticity amongst the 50 populations. The absence of the Mbuti Pygmy population (Balloux et al., 2009) seemed to (visually) have the most impact. Indeed, without Mbuti Pygmy, heteroscedasticity appeared to be absent; amongst the 50 populations, it seemed like Mbuti Pygmy may have resulted in the heteroscedasticity. And so, without Surui and Mbuti Pygmy, there was a correlation between adjusted X-chromosomal diversity and minimum temperature (also see Supplementary Note 7). If a less stringent criterion (e.g., |1.96|) was used for detecting atypical datapoints (e.g., Field, 2013), then, when 51 populations were used, Mbuti Pygmy would have been atypical (z = -2.91).

Autosomal SNP haplotype heterozygosity (adjusted for distance from Africa) was not observed to be associated with minimum temperature, sr(48) = .16, p = 1.00, spatial DW = 1.81. Likewise, autosomal microsatellite heterozygosity (controlling for distance from Africa) did not yield a correlation with minimum temperature, sr(46) =.28, p =.28, spatial DW = 1.92 (without Surui because of their standardised residual, z = -3.69), and neither did Y-chromosomal microsatellite heterozygosity (adjusted for distance from Asia), sr_s(48) = -.11, p = 1.00. A semi-partial Spearman test was used regarding Y-chromosomal diversity because heteroscedasticity seemed apparent in parametric analysis, and it persisted when any one population was absent (see Supplementary Note 8).

Study 2: X-chromosomal diversity and the sex chromosome diversity ratio

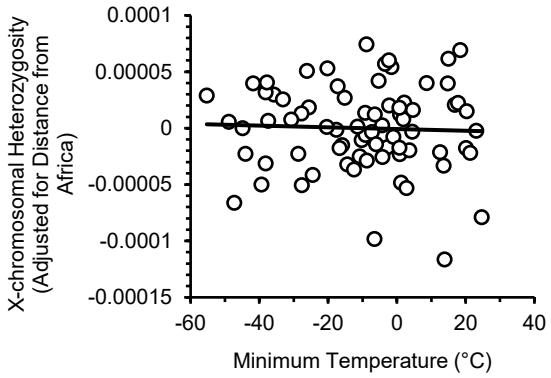
X-chromosomal diversity

When it came to the relationship between X-chromosomal diversity (SGDP data) and distance from Africa, two datapoints were indicated to be atypical. These datapoints were for the Mixtec population and the Palestinian population (Mallick et al., 2016); the former was atypically high (z = 3.62), and the latter was atypically low (z = -4.59). Therefore, those two datapoints were not used when i) testing whether X-chromosomal diversity is associated with distance from Africa, and when ii) X-chromosomal diversity was adjusted for distance from Africa. X-chromosomal heterozygosity (SGDP data) was found to decrease as distance from Africa advanced, r(12.30) = -.77, p = .003, therefore, the X-chromosomal diversity (of populations who are located outside of Africa) was adjusted for distance from Africa. When examining if there is a correlation between adjusted X-chromosomal diversity and minimum temperate (i.e., if there is a significant semi-partial Pearson correlation), the author was unable to generate a p-value through the ppcor package. Therefore, the p-value was calculated from a t-value, and the t-value was arrived at (from the correlation coefficient)

via a formula which is presented in Kim (2015). In contrast to results in Study 1, adjusted X-chromosomal diversity was not found to correlate with minimum temperature, sr(72) = -.04, p = 1.00 (Figure 3).

Figure 3

A Correlation was not Observed Between Adjusted X-Chromosomal Diversity and Minimum Temperature in Study 2



Note. The values of the datapoints are given in Table S2 as are the geographical distances (from Africa) which were used to adjust X-chromosomal diversity.

Sex chromosome diversity ratio

Research on diversity and climate has considered sex-biased migration (Balloux et al., 2009). It is unlikely that an association between mitochondrial diversity and climate is because of some relationship between climate and sex-biased migration—previous research did not find that the ratio of diversity between sex chromosomes is related to minimum temperature (Balloux et al., 2009). Therefore, it could seem improbable that the climatic signal in the X chromosome (Study 1) arose from an association between sex-biased migration and climate. However, Y-chromosomal diversity does not have a climatic signal (Balloux et al., 2009; Study 1), unlike X-chromosomal diversity in Study 1. Therefore, it would not be surprising if there actually is a link between sex-biased migration and climate. Although Balloux et al. (2009) examined if a link is indicated, they do not appear to have accounted for the expansion signal in X-chromosomal diversity; such an accounting could allow for a more powerful test of whether the sex chromosome diversity ratio is associated with climate. In Study 2, the ratio of adjusted X-chromosomal to Y-chromosomal diversity increased with minimum temperature (Mbuti Pygmy and Surui were not included in analysis—see Study 1 results), r(47) = .45, p = .004, Spatial DW = 1.91; as minimum temperature cools, the X chromosome becomes less diverse relative to the Y chromosome.

Discussion

Previous research has not found support for a climatic signal being present in autosomal diversity, Y-chromosomal diversity (Balloux et al., 2009), or cranial form diversity (Betti et al., 2009). This lack of support was matched in the present manuscript which addressed potential limitations of previous studies (see *Introduction*). On the other hand, Study 1 diverged from previous research (i.e., Balloux et al., 2009) when it came to X-chromosomal diversity; in Balloux et al. (2009), a climatic signal was not indicated in X-chromosomal diversity, unlike in Study 1.

This incongruity between Balloux et al. (2009) and Study 1 is unlikely to have been caused by Balloux et al. and Study 1 not using the same location as the *origin*—similar origins were used regarding X-chromosomal diversity (see *Introduction*; Table S1). Instead, the incongruity likely arose because correlation tests in Study 1 omitted populations (Mbuti Pygmy and Surui) which Balloux et al. (2009) included in their analysis. In the present research, if the correlation coefficient is calculated when those two populations are included (global analysis), the coefficient is numerically smaller in magnitude (r = .18) than when the populations are absent. If the coefficient is calculated without Surui, but including Mbuti Pygmy, the correlation coefficient is also numerically smaller (r = .28). It is also worth stating that when only populations outside of Africa are used, the inclusion of Surui (who had an atypical standardised residual) leads to a numerically weaker correlation coefficient (r = .31) than when Surui are absent (Figure 2A).

In contrast to Study 1, Study 2 did not lend support to there being a climatic signal in X-chromosomal diversity. Therefore, the present research leaves a question mark when answering if there is a climatic signal. Regarding X-chromosomal diversity, more populations were used in Study 2 (SGDP data) than in Study 1 (HGDP-CEPH data) (Balloux et al., 2009; Mallick et al., 2016; see *Method*). However, per population, far fewer individuals were represented in the SGDP data (see Balloux et al., 2009; Mallick et al., 2016); heterozygosity from the HGDP-CEPH data may generally be more representative of populations. Therefore, it could be reasonable to *provisionally* favour results found with the HGDP-CEPH data (i.e., Study 1). However, further research clearly is required in order to say, with any reasonable confidence, if there is a climatic signal.

Mitochondria are recognised for being tied to climatic adaptation (Grover-Thomas et al., 2026), unlike the X chromosome (e.g., Lasne et al., 2019). Therefore, Study 1's indication of a climatic signal in the X chromosome may be surprising. Nevertheless, there are, theoretically, other routes through which genetic diversity could be related to climate, distinct of climatic adaptation, e.g., a link between sex-biased migration and climate (Balloux et al., 2009). However, previous research did not find a relationship between the sex chromosome diversity ratio and climate, and this absence of a relationship suggested that the climatic signal in mitochondrial diversity is not explainable by a relationship between sex-biased migration and climate (Balloux et al., 2009). On the other hand, Study 2 found a relationship between the ratio and climate; the present research may give weight to the idea that the climatic signal in mitochondrial diversity *and* X-chromosomal diversity could reflect an association between climate and sex-biased migration. Of course, given the apparent climatic

signal in X-chromosomal diversity when HGDP-CEPH data are used, perhaps it is reasonable to expect a relationship between the sex chromosome diversity ratio (calculated from HGDP-CEPH data) and climate in Study 2 (see *Method*). Consequently, because a climatic signal in the X chromosome was not evident in SGDP data (Study 2), there is clear reason to question the correlation between the ratio and climate. The SGDP data contains Y-chromosomal data (e.g., Willems et al., 2016), and X-chromosomal (e.g., Mallick et al., 2016) which were used in Study 2; SGDP data could be used to see if the relationship between the sex chromosome diversity ratio and climate is replicated.

When examining if adjusted diversity is related to minimum temperature, heteroscedasticity i) may have been present in some analyses in Balloux et al. (2009) (see Supplementary Note 8), and it ii) seems to have occurred in the present research when analysis used adjusted X-chromosomal diversity (50 populations) and adjusted Y-chromosomal diversity in Study 1. Heteroscedasticity is known to arise for a number of reasons, e.g., crucial variables not being included in analysis (e.g., Ohaegbulem & Iheaka, 2024); research on climate and diversity has not been limited to minimum temperature when it comes to climatic variables (Betti et al., 2009, 2012; Manica et al., 2007). Therefore, the actual/potential heteroscedasticity in Balloux et al. and Study 1 could possibly have arisen because certain climatic variables were not taken into account. Consequently, regarding climatic variables, it would seem to be prudent for future research to look beyond minimum temperature when addressing if there is a climatic signal in X-chromosomal diversity. Nevertheless, in previous research which found a relationship between climate and the diversities of the tibia and femur, backward stepwise regression resulted in minimum temperature being included as the only predictor/independent variable (not maximum temperature, precipitation, climatic interaction terms, or indeed distance from Africa) regarding the prediction of diversity (Betti et al., 2012).

Conclusion

In agreement with previous research (Balloux et al., 2009; Betti et al., 2009), a climatic signal did not seem to be evident in autosomal, Y-chromosomal, or cranial form diversities. In contrast to previous research (Balloux et al., 2009), X-chromosomal diversity seemed to have a climatic signal in Study 1. However, no such signal was apparent in a second study which used a different dataset. With there being conflicting findings over whether there is a climactic signal in X-chromosomal diversity (Studies 1 and 2), it seems rather questionable if there truly is a climatic signal in the X chromosome. Hopefully, the present research i) functions as a reminder of how important replicability is (see Nosek, & Errington, 2020) and, ii) encourages further replication research on whether X-chromosomal diversity has a climatic signal.

References

Abdi, H. (2010). Holm's sequential Bonferroni procedure. In N. J. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 573–577). Sage.

Balloux, F., Lawson Handley, L.-J., Jombart, T., Liu, H., & Manica, A. (2009). Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1672), 3447–3455. https://doi.org/10.1098/rspb.2009.0752

Betti, L., Balloux, F., Amos, W., Hanihara, T., & Manica, A. (2009). Distance from Africa, not climate, explains within-population phenotypic diversity in humans. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1658), 809–814. https://doi.org/10.1098/rspb.2008.1563

Betti, L., von Cramon-Taubadel, N., & Lycett, S. J. (2012). Human pelvis and long bones reveal differential preservation of ancient population history and migration out of Africa. *Human Biology*, *84*(2), 139–152. https://doi.org/10.3378/027.084.0203

Betti, L., von Cramon-Taubadel, N., Manica, A., & Lycett, S. J. (2013). Global geometric morphometric analyses of the human pelvis reveal substantial neutral population history effects, even across sexes. *PLoS ONE*, *8*(2), Article e55909. https://doi.org/10.1371/journal.pone.0055909

Cenac, Z. (2022). *Modern human expansion from (southern?) Africa: Origin, non-African ancestry, and linearity*. bioRxiv. https://doi.org/10.1101/2022.07.31.500977

Cenac, Z. (2023a). *Expansion of modern humans over the world: The origin when considering non-linearity*. bioRxiv. https://doi.org/10.1101/2023.03.11.532168

Cenac, Z. (2023b). *Y-chromosomal microsatellites and expansion from Africa*. bioRxiv. https://doi.org/10.1101/2023.04.29.538812

Chen, Y. (2016). Spatial autocorrelation approaches to testing residuals from least squares regression. *PLoS ONE*, *11*(1), Article e0146865. https://doi.org/10.1371/journal.pone.0146865

Deblauwe, V., Kennel P., Couteron, P. (2012). Testing pairwise association between spatially autocorrelated variables: A new approach using surrogate lattice data. *PLoS ONE*, *7*(11), Article e48766. https://doi.org/10.1371/journal.pone.0048766

Field, A. (2013). *Discovering statistics using IBM SPSS statistics (4th ed.)*. Sage.

Grover-Thomas, F., van Dorp, L., Balloux, F., Andrés, A. M., & Camus, M. F. (2026). Climate associated natural selection in the human mitochondrial genome. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msag044

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2009). *Data analysis: A model comparison approach*. Routledge.

Keinan, A., & Reich, D. (2010). Can a sex-biased human demography account for the reduced effective population size of chromosome X in non-Africans? *Molecular Biology and Evolution*, *27*(10), 2312–2321. https://doi.org/10.1093/molbev/msq117

Lasne, C., Van Heerwaarden, B., Sgrò, C. M., & Connallon, T. (2019). Quantifying the relative contributions of the X chromosome, autosomes, and mitochondrial genome to local adaptation. *Evolution*, *73*(2), 262–277. https://doi.org/10.1111/evo.13647

Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., Cann, H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L., & Myers, R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, *319*(5866), 1100–1104. https://doi.org/10.1126/science.1153717

Luca, F., Hudson, R. R., Witonsky, D. B., & Di Rienzo, A. (2011). A reduced representation approach to population genetic analyses and applications to human evolution. *Genome Research*, *21*(7), 1087–1098. https://doi.org/10.1101/gr.119792.110

Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, *22*(6), 665–674. https://doi.org/10.5351/CSAM.2015.22.6.665

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., Skoglund, P., Lazaridis, I., Sankararaman, S., Fu, Q., Rohland, N., Renaud, G., Erlich, Y., Willems, T., Gallo, C., … Reich, D. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, *538*, 201–206. https://doi.org/10.1038/nature18964

Manica, A., Amos, W., Balloux, F., & Hanihara, T. (2007). The effect of ancient population bottlenecks on human phenotypic variation. *Nature*, *448*(7151), 346–348. https://doi.org/10.1038/nature05951

NASA POWER (2026). *NASA POWER Data Access Viewer*. https://power.larc.nasa.gov/data-access-viewer/

Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLoS Biology*, *18*(3), Article e3000691. https://doi.org/10.1371/journal.pbio.3000691

Ohaegbulem, E. U., & Iheaka, V. C. (2024). On remedying the presence of heteroscedasticity in a multiple linear regression modelling. *African Journal of Mathematics and Statistics Studies*, *7*(2), 225–261. https://doi.org/10.52589/AJMSS-TJ9XI8HD

Prugnolle, F., Manica, A., & Balloux, F. (2005). Geography predicts neutral genetic diversity of human populations. *Current Biology*, *15*(5), R159–R160. https://doi.org/10.1016/j.cub.2005.02.038

R Core Team (2021). *R: A Language and Environment for Statistical Computing (*Version 4.0.5) [Computer software]. https://www.R-project.org

Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., & Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National*

Academy of Sciences of the United States of America, *102*(44), 15942–15947. https://doi.org/10.1073/pnas.0507611102

Rangel, T. F., Diniz-Filho, J. A. F., & Bini, L. M. (2010). SAM: A comprehensive application for Spatial Analysis in Macroecology. *Ecography*, *33*, 46–50. https://doi.org/10.1111/j.1600-0587.2009.06299.x

Rasteiro, R., Bouttier, P.-A., Sousa, V. C., & Chikhi, L. (2012). Investigating sex-biased migration during the Neolithic transition in Europe, using an explicit spatial simulation framework. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1737), 2409–2416. https://doi.org/10.1098/rspb.2011.2323

Savin, N. E., & White, K. J. (1977). The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. *Econometrica*, *45*(8), 1989–1996. https://doi.org/10.2307/1914122

von Cramon-Taubadel, N., & Lycett, S. J. (2008). Brief communication: Human cranial variation fits iterative founder effect model with African origin. *American Journal of Physical Anthropology*, *136*(1), 108–113. https://doi.org/10.1002/ajpa.20775

Willems, T., Gymrek, M., Poznik, G. D., Tyler-Smith, C., The 1000 Genomes Project Chromosome Y Group, & Erlich, Y. (2016). *Population-scale sequencing data enables precise estimates of Y-STR mutation rates*. bioRxiv. https://doi.org/10.1101/036590

Williams, E. (2011). *Aviation formulary V1.46*. http://edwilliams.org/avform.htm