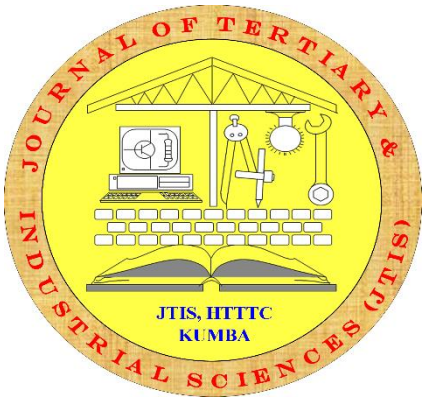
ISSN 2709-3409 (Online)

JOURNAL OF TERTIARY AND INDUSTRIAL SCIENCES

A MULTIDISCIPLINARY JOURNAL OF THE HIGHER TECHNICAL TEACHERS' TRAINING COLLEGE, KUMBA



VOLUME 3, NUMBER 1
MARCH, 2023

PUBLISHER: HIGHER TECHNICAL TEACHERS' TRAINING COLLEGE (HTTTC)
University of Buea
P.O Box: 249 Buea Road, Kumba
Tel: (+237) 33354691 – Fax: (+237) 33354692
Email: editorinchief@htttckumba.com
Website: http://www.htttckumba.com

ISSN 2709-3409 (Online)

EDITORIAL BOARD

ISSN 2709-3409 (Online)

JOURNAL OF TERTIARY AND INDUSTRIAL SCIENCES
VOLUME 3, NUMBER 1, MARCH 2023

TABLE OF CONTENTS

ISSN 2709-3409 (Online)

GENETIC DIVERSITY AND STAPLE CROPS CONSUMED IN CAMEROON: A GENETIC HISTORIAN'S APPRAISAL

By

FORKA LEYPEY MATHEW FOMINE[1]

Abstract

The primary objective of this article was to investigate what social and scientific factors differentiate north Cameroonian ethnic groups from those of the south. Cameroon has been described in the literature as Africa in miniature, taking into consideration the country's ethnic, linguistic, ecological, genetic, religious and social diversity. The findings of this paper reveal that food is just one of many socio-cultural historical features which vary and/or display continuity across time and space. Interestingly, there are many more differences between people in North and South Cameroon than might have been suggested by grains consumption in the north and tubers in the south. This research addressed the question of whether the choice of staples influenced the genetic diversity of the cultivating group either in the paternal line of inheritance, the maternal line of inheritance or both. This is evident because in historical genetic analysis paternal line inheritance is studied by critically analysing the non-recombining portion of the exclusively paternally inherited Y chromosome (NRY), while maternal line inheritance is on its part studied by analysing the exclusively maternally inherited mitochondrial (mtDNA). The paper further reveals that the high genetic diversity among northern Cameroonian ethnic groups emanated from the Fulani Jihads that led to an influx of a huge number of reproductive men and women. Similarly, the low genetic diversity among southern groups stemmed from the Trans-Atlantic Slave Trade that drained a significant fraction of reproductive men and women. Genetic information in this paper was collected through buccal swab samples from anonymous male donors over 18 years of age, unrelated at the grandparental level who gave their informed consent. Along with a buccal swab, each donor was questioned for information on their self-declared ethnicity, languages spoken and place of birth with similar information collected about their parents, paternal grandfather and maternal grandmother. Concerning staples, information was gathered from published books and articles. This paper concludes that a mastery of DNA study helps to illuminate when ethnic groups split up within a population.

Keywords: Genetic diversity, Y Chromosome, Mitochondrial DNA, Cameroon, Food.

[1] Faculty of Arts, University of Buea

ISSN 2709-3409 (Online)

1. Introduction

Cameroon lies at the intersection of Central and West Africa. It is geographically, ethnically, historically, culturally, linguistically, nutritionally and genetically a diverse country, proud of its nickname Africa in Miniature. While grains (millets and sorghums) are the chief staple crops consumed by the ethnic groups that inhabit the Northern part of Cameroon (Far North, North and Adamaoua Regions), tubers, including cassava and cocoyam are the main staples eaten by groups of the south (Centre, South, West, South-West, Littoral and North West regions). Food is just one of many socio-cultural historical features which vary and/or display continuity across time and space. There are many more differences between people of the North and the South of Cameroon that might have been suggested by grains consumption in the North and tubers in the South. This paper addresses the question of whether the choice of crop influences the genetic diversity of the cultivating group either in the paternal line of inheritance, the maternal line of inheritance or both. In genetic studies paternal line inheritance may be studied by analysing the non-recombining portion of the exclusively paternally inherited Y chromosome (NRY), while maternal line inheritance may be studied by analysing the exclusively maternally inherited mitochondrial genome (mtDNA). Since both the NRY and mtDNA can be characterised as haplotypes and sons have the same NRY haplotype (with the exception of rare mutations) as their biological fathers while both sons and daughters have the same mtDNA haplotype as their mothers, paternally mediated and maternally mediated genetic diversity may be assessed in terms of the diversities of NRY and mtDNA haplotypes respectively.

There are many social practices that prevent or reduce equal reproductive success of members of both sexes thereby reducing NRY and/or mtDNA haplotype diversity (hereafter referred to as hNRY and hmtDNA diversity respectively). While hNRY and hmtDNA may be expected to increase over time due to rare mutations hNRY will decrease if all males do not produce an equal number of sons and hmtDNA will decrease if all women do not produce an equal number of children. Two common examples will suffice to represent the large number of social practices that can affect hNRY and hmtDNA: a) a group in which male polygamy is practiced by a privileged class while males at the bottom of the economic structure are prevented from taking wives will reduce hNRY and b) groups practicing exogamy combined with patrilocality will experience higher hmtDNA than otherwise would be the case.

It is established in this paper that different practices related to food production are associated with different structures of social organisation and different roles in society for female and male members of groups. Differences in such roles affect reproductive success. Consequently it is possible to hypothesise an association between selection of staple crop and either hNRY, hmtDNA or both. However it is also necessary to consider what other influences may contribute to variation in hNRY and hmtDNA. In this paper statistical associations between hNRY and hmtDNA were also assessed for: group population size, geographic location (latitude) since crop choice is associated

with environmental conditions which, in Cameroon, vary considerably along a north-south axis, religion, political structure and descent system. Thirty groups were included in the study and data analysed to establish whether there were statistically significant associations between hNRY and hmtDNA respectively and, respectively, latitude, population size, most common religion practiced within the group, political structure, descent system and dominant staple crop.

The groups are divided into a northern set that grows seed crops. The most southerly of these-groups is the Nso' (822Km south from Lake Chad). The Bafut are the most northerly of the leaf crop (835Km). Bamun (875Km) are therefore included in the southern set because they are to the south of Kumbo (the reference point for the Nso') even though they grow maize as their primary staple crop. The northern group has 14 seed growers and no leaf growers. The southern set excluding Pygmies, has 1 seed grower and 11 leaf growers (Fisher exact test of difference <0.0001). The variables analysed in both the northern and southern groups are: distance south from Lake Chad, mtDNA diversity (h), NRY diversity (h), principal staple crop, total population, political system, descent system and principal religion.

2. Literature review

Although much has been written on human genetics in general and the historical importance of the Y chromosome in particular, less emphasis has been laid on the relationship between nutrition and genetics. Daniel L. Hartl and Andrew G. Clark (1980), writing in *Principles of Population Genetics*, reveal that the ultimate source of genetic variation is mutation, which leads to a heritable change in genetic material (Hart and Clark, 1977. According to them, mutation therefore includes change in the nucleotide sequence of a single gene as well as the formation of a chromosome rearrangement, such as an invasion or a translocation. They further discuss how migration enables mutations to spread among sub-populations. They also share the opinion that although mutations are generally rare, the alleles of large populations are at risk. Mutations introduce new alleles into population, and random genetic drift determines whether a neutral allele will ultimately be fixed or lost.

Cavalli-Sforza et al. (1996) have intimated the importance of DNA in general and that of mtDNA in particular. They reveal that the first polymorphisms examined in humans for evolutionary purposes were from mitochondrial DNA. They further explain mitochondria as being self-reproducing units contained in all cells of higher organisms (eukaryotes, from fungi to mammals), usually in many copies per cell (up to 10,000 or more). They further intimate that mtDNA are transmitted only by the mother but are present in both sexes. This is one of the issues that has been investigated and illuminated in this study. The scope of Cavalli-Sforza, Menozzi, and Piazza study is different from that of this thesis in that it encompasses the genetic history of the entire world while this is focused solely on the genetic history of selected Cameroonian populations.

Bradman and Mark (1998) unravel that the Y Chromosome is paternally inherited while the mtDNA is maternally inherited. Through the Y Chromosome analysis, they

are of the opinion that half or more of the Lemba have a Semitic origin although it is uncertain whether they were inherited from Jewish or Arab traders.

Turning to the social history of food, there are written works though limited in scope and focus. For instance, Edwin Ardener (1956) tackles the problem of nutrition among the Bakweri. He mentions the availability of staples in the region such as cocoyam, colocasia, yams and plantains. The only staple whose origin he traces is the cocoyam. He indicates that cocoyam were introduced in Bakweri land from Fernando Po (Equatorial Guinea) by the Baptist Missionaries.

In a joint article, Bah and Ghomsi (1986) have handled the cultivation of certain crops in the Sudan and the Lake Chad Basin. According to the two scholars, the Arabs introduced wheat in the Lake Chad Basin. Other crops that these scholars mention their cultivation in the area from time immemorial are rice and sugar cane.

Difference between causation and correlation

One of the most common errors made in social sciences is the confusion between *correlation* and *causation.* In theory, these are easy to distinguish — an action or occurrence can *cause* another (such as smoking causes lung cancer), or it can *correlate* with another (such as smoking is correlated with alcoholism). If one variable associates with another, then they are most certainly correlated. But just because two variables correlated does not mean that one caused the other (Daun, 2001).

Furthermore, eating breakfast has long been correlated with excellent performance in elementary school children. It would be erroneous to conclude that eating breakfast *causes excellent performance in elementary school children*. It is possible however, that those who do not eat breakfast are also more likely to be frequently absent in school — and it is of course absenteeism that plays a significant role in their poor academic performance. Independent of other factors, breakfast only helps undernourished children perform better.

The most important thing to remember is that correlation and causation are not the same thing. Correlation is a statistic representing how closely two variables co-vary. In other words, when one variable is studied next to another, an effect is likely to happen. Causation is when two variables directly affect each other. For example, if a dog stays the night outside and it gets sick, and this happens many times, it's likely that the dog got sick because it stayed the night outside. It might not be the cause, however. A virus or bacteria might be the cause that results in the dog's ill health.

It is very important in the social sciences to understand the difference between correlation and causation. People often make the mistake of assuming that if two things are correlated, there must be some causation involved in the relationship, which is not necessarily the case. Correlation refers to the strength and direction of a linear relationship between two variables. For example, if short students received higher scores on a class exam than tall students, there would be a correlation between height and exam scores. However, this does not imply a causal link between height, and test scores, as short people are not necessarily smarter than tall people. The actual cause of the correlation may be caused by a third factor, like the short students always doing their assignments.

In a nutshell, correlation does not imply causation. This statement is one of the fundamental dictums shared by the behavioural and natural sciences, and it is also one of the most important ideas scholars need to master. This statement serves to remind scholars that relationships in the world are not always what they seem to be. Unless we actively intervene into a situation by manipulating variables and measuring their effect on one another, we cannot assume that we know the causal order involved. Moreover, if correlation does not imply causation, then should such analyses be avoided? Not at all. Correlational research, the examination of associations among variables, is extremely valuable and very often revealing, even enlightening. More seriously, however, correlations are a start. They can be highly suggestive, pointing researchers in the right direction so that experiments identifying the causal ordering among known variables such as higher genetic diversity in North Cameroon than in the South and the unknown variables can be designed and subsequently executed. As a critical consumer of behavioural science research, however, one must avoid confusing correlated variables with causal variables.

Y Chromosome and Mitchondrial DNA

All human cells (other than mature red blood cells) possess a nucleus, which contains the genetic material (DNA) arranged into 46 chromosomes, themselves grouped into 23 pairs. In 22 pairs, both members are essentially identical, one deriving from the individual's mother, the other from the father; such pairs are known as autosomes. The 23rd pair is different: while in females this pair has two like chromosomes called "X," in males it comprises one "X" and one "Y," two extremely dissimilar chromosomes. It is primarily these chromosome differences, which determine sex. During the production of sperm and eggs (gametes), the paired chromosomes separate so that each gamete ends up with only one member of each chromosome pair. However, before separation occurs, the paired autosomes swap pieces of their DNA with each other. In women, this exchange process also takes place between the two X chromosomes but, in men, unmatched X and Y chromosomes do not exchange except at the ends of two chromosomes referred to as the pseudoautosomal regions (Bradman and Thomas, 1998).

All eggs (produced by females) contain an X chromosome while sperm (produced by males) contains an X and Y. Fertilization restores the chromosomes to their normal paired condition (Mange and Mange, 1999). Thus, a Y sperm fertilizing the X (all eggs carry only the X chromosome because females have got no Y) produces an XY zygote (cell produced by union of two gametes), which develops as male; fertilization by an X sperm definitely gives rise to a female XX zygote. Since every male must possess a Y (or male) chromosome, which can be exclusively inherited only from his biological father, a man's Y chromosome represents a unique record of his paternal inheritance.

As mentioned above, the chromosomes are mainly composed of DNA, that remarkable substance which constitutes "the book of life," the genetic instruction written in a four-letter alphabet: A (adenine), T (thymine), C (cytosine) and G (guanine). It is these letters that are read to uncover the history recorded in the 60 million letters of the Y chromosome. Since the Y chromosome, unlike the autosomes in the other 22 pairs, does not exchange letters with a partner (except for the comparatively small pseudoautosomal regions) those mistakes are the only changes that are passed on to the next generation (Bradman and Thomas, 1998).

Geneticists consider that in a particular island, over several generations, the population numbers remain constant. Some men have sons, some do not: those with sons pass on their Y chromosomes while those of others are lost to history. It is easy to show that eventually the descendants of only one of the original Y chromosomes (in many copies) survive in the population because, once a particular line dies out, it never reappears. In the Y chromosome passage through generations, changes occur randomly in its junk DNA and so Y chromosomes of the contemporary population retain a record of their passage through time: they can reveal paternal genealogy of their owners and the relationships between different groups of individuals. For the genetic historians (and other geneticists), that island is the whole world and we are the current generation whose genetic history is open for study.

At this point, some changes that occur in chromosomes and are of paramount importance to a genetic historian are worth mentioning. Changes that do occur in chromosomes from generation to generation are of four types: notably, indels – insertions into or deletions of the DNA at particular locations on the chromosome. One insertion particularly useful in population studies is the YAP, "Y chromosome alu polymorphism". Alu is a sequence of approximately 300 letters (based pairs) which inserted itself into a particular region of the DNA. There have been some half a million alu insertions into human DNA; YAP is one of the most recent.

Snips – are "single nucleotide polymorphisms" in which a particular nucleotide (an A, for example is changed (perhaps into a G). Stable indels and snips are relatively rare and, in the case of the latter, so infrequent that it is reasonable to assume they have occurred at any particular position in the genome only once in the course of human evolution. Snips and stable alus have been termed "unique event polymorphisms" (UEPs).

In using polymorphisms to study changes over time, we are fortunate in having markers, which change in different rates. Perhaps we can think of the UEPs as the hour hand, the microsatellite polymorphisms as the minute hand and the minisatellites as a sweep second hand of the evolutionary clock. Because most of the Y chromosome does not change DNA with a partner, a further benefit of using it to study evolution is that all the markers are joined one to another along its entire length. Such linkage of markers means that a haplotype constructed from a number of different markers records the evolutionary history of the particular Y chromosome on which they are all located. Due to the biological and historical importance of the Y chromosome, it currently occupies a conspicuous and fascinating position in the genetic historian's toolbox.

Mitochondria are structures within cells that convert the energy from food into a form that cells can use. Although most DNA is packaged in chromosomes within the nucleus, mitochondria also have a small amount of their own DNA. This genetic material is known as mitochondrial DNA (mtDNA). The first DNA polymorphisms examined in humans for evolutionary purposes were from mitochondrial DNA. Mitochondria are self-reproducing units contained in all cells of higher organisms (eukaryotes, from fungi to mammals), usually in many copies per cell (up to 10,000 or more) (Cavalli-Sforza et al., 1996).

Mitochondria are normally transmitted exclusively by the mother but are present in both sexes (recall that all mitochondria in a fertilized egg come from the egg only; that

mitochondria that are present in the tail section of the sperm are usually destroyed by the egg cell during the formation of the zygote definitely the genes in mtDNA follow an exclusively maternal inheritance pattern, generation after generation, without any reassortment or crossing over). So, any trait associated with a mitochondrial gene must be transmitted by the mother to all of her children, both male and female. In the other direction, this lineage should extend from any individual through the mother, maternal grandmother, the maternal great-grandmother, and so on, as far back as humans existed (Mange and Mange, 1999).

The maternally inherited mitochondrial DNA (Smith, 1989) is being widely used (in population genetics, genetic history and in other genetic studies) to infer aspects of human female population histories (Wilson et al., 2000). In contrast to the nuclear DNA, mtDNA plays no part in the development of the phenotype, but functions independently during reproduction of energy within the cell. In some ways mtDNA is best thought of as a benign and essentially within each cell. Apart from its different function, mtDNA has a number of distinctive characteristics. First, compared with nuclear DNA there is not much of it in technical terms, it is only 16, 451 bases long, as opposed to the 30, 000, 000 (million) in nuclear genes. This makes it relatively easy to study and compare between individuals, populations and species.

Second, and most important, it is inherited solely through the female line – that is, males do not pass their mtDNA at all, and both males and females inherit their mtDNA from their mothers. Third, parts of the mtDNA change relatively rapidly, making it an appropriate tool for investigating recent evolutionary history. Sequences of the mtDNA have been employed to construct the genetic history of human species. Wilson and Stoneking, among others, have pioneered this type of analysis, and it is from their work that the idea of the African Eve has emerged. Essentially, what they did was to look at the pattern of variation in human mtDNA across the world, region by region, and population by population. What they found was twofold: first, it appeared that there was more variation in mtDNA among Africans than anywhere else in the world, and furthermore, that any "tree" of relationships among human populations showed a major split between Africans and other populations.

Second, using the estimates of the length of time it takes for change to occur within the mtDNA, they calculated that the last common ancestor of all the types of mtDNA found today lived approximately 200, 000 years ago. As mtDNA is inherited maternally, this individual must have been a woman. As there is more variation in Africa than anywhere else, this implies that Eve's descendants have been in Africa longer than anywhere else in the world, and also that she must have lived in Africa. This is the basis for the African Eve (Foley, 1998).

Since Mitochondrial Eve is believed to have lived in Africa, she is sometimes referred to as African Eve, an ancestor who has been hypothesized on the grounds of fossil as well as DNA evidence. According to the most recent interpretation of the mitochondrial DNA data, the title belongs to the same hypothetical woman. Family tree (or "phylogenies") constructed on the basis of mitochondrial DNA comparisons show that the living humans whose mitochondrial lineages branched earliest from the tree are indigenous Africans, whereas the lineages of

indigenous peoples on other continents all branch off from African lines. Researchers therefore reason that all living humans descend from Africans, some of whom migrated out of Africa and populated the rest of the world. If the mitochondrial analysis is correct, then because mitochondrial Eve represents the root of the mitochondrial family tree, she most have predated the exodus and lived in Africa. Therefore many researchers take the mitochondrial evidence as support for the "single-origin" or Out-of-Africa model/theory.

The concept of "mitochondrial Eve" is widely misunderstood. It does not mean that she was the only woman of her time who was ancestral to the modern humans. In other words, mitochondrial Eve was not a Biblical Eve. However the Biblical Eve, if she had existed, might well be mitochondrial Eve (though not necessarily: it would be one of her female descendants).

The name Eve, in retrospect, is perhaps the worst possible name to give to the entity in question. It is believed that this is probably the cause of so much confusion in understanding what this title is. Some scholars think that this title has some theological or religious consequences. Nothing of that sort. The concept of Mitochondrial Eve has not validated the bible (or the book of genesis). The Mitochondrial Eve of some 200, 000 years ago is not our common ancestor, or even common genetic ancestor. She is the most-recent common ancestor of all humans alive on earth today with respect to matrilineal descent. That might seem like a mouthful, but without even a single one of those qualifying phrases, any description or discussion of the Mitochondrial Eve reduces to a lot of nonsense.

The term mitochondrial Eve itself is a title given retroactively to a woman. Often the conferring of the title occurs many hundreds of thousands of years after the death of the woman in question. Mitochondrial Eve lived with many other humans (men and women); she was certainly not alone. When she was alive, she was most certainly not the Mitochondrial Eve of the time. The title at that time was held by a different ancestor of hers.

The existence of the Mitochondrial Eve is not a theory; it is a fact, unless something like a multiple-origins theory of human evolution that is, the human species arose independently in different geographically separated populations, and that the present-day ease of interbreeding is the result of a remarkable convergent evolution, is true. Few people subscribe to the multiple-origins theory, and the Mitochondrial Eve observation is a refutation of multiple origins. Mitchondrial Eve is also known as African Eve an apparent indication that the origin of humanity is Africa.

3. Methodology

Buccal swab samples were collected from anonymous male donors over 18 years of age, unrelated at the grandparental level who gave informed consent. Towns and villages, as well as rural areas which were known to contain large numbers of a particular ethnic group, were targeted for sample collection. Samples were collected on a first come first taken basis, provided that donors were not related at the paternal grandparental level to any previous sample donor. Along with a buccal swab, each donor was questioned for information on their self-declared ethnicity, languages spoken and place of birth with similar information collected about their parents, paternal grandfather and maternal grandmother.

ISSN 2709-3409 (Online)

Donors were instructed on the procedure for buccal swab DNA collection after informed consent was given. The cotton swab is rubbed on the inside of both cheeks and the inside of the lips of the donor for at least 20 seconds. After swabbing was complete, the swab was immediately placed inside a sample collection tube containing 1ml solution of 0.05M Ethylenediaminetetraacetic acid (EDTA) and 0.5% Sodium Dodecyl Sulfate (SDS), that immersed the cotton end of the swab, and retarded breakdown of the DNA sample during transit to the laboratory in England. Each sample collection tube was then securely sealed with PVC tape, and kept in as cool and dark place as possible until it could be transported to the laboratory.

Turning to the collection of ethnographic information on sample donors, it is worth mentioning the point that, for each sample donor, a questionnaire was completed in English on behalf of the sample donor. Each anonymous donor was allocated an alphanumeric code that was written on the sample collection tubes of the donor's DNA sample and recorded on the donor's questionnaire. For each donor, the following information was recorded:

-The donor's age

-Self-declared cultural identity (or ethnicity)

- First and second language spoken

-The donor's religion

-Place of birth

-The cultural identity (or ethnicity) of the donor's father, mother, paternal grandfather and maternal grandmother

-First and second language spoken by the donor's father, mother, paternal grandfather and maternal grandmother. The place of birth and current residence of the donor's father, mother, paternal grandfather and maternal grandmother was also recorded.

In relation to laboratory methods, DNA from buccal swab samples was extracted using a phenol/chloroform method. To each buccal swab sample collection tube containing the swab and 1ml of 0.05M EDTA and 0.5% SDS, 0.8ml of 0.02 $mgml^{-1}$ of proteinase K solution was added, and then incubated at 56°C for two hours. Furthermore, 0.8ml of the digested buccal swab sample solution was then added to microfuge tubes containing a 1:1 mix of phenol/chloroform. The sample solution and phenol/chloroform was mixed, and then centrifuged for 10 minutes at over 2000g. The aqueous upper layer was then transferred to a 2ml screw-top microfuge tube (used for long-term DNA storage) containing 0.7ml isopropanol. The sample and isopropanol solution was mixed, chilled at -20°C for two hours, and then centrifuged for 13 minutes at over 2000g. The supernatant was decanted off carefully to avoid dislodging the precipitated DNA pellet, and the microfuge tube allowed to dry for 1 minute while inverted at a 45° angle. 0.8ml of 70% ethanol was then added to the microfuge tube, and then centrifuged for a final 10 minutes at over 2000g.

All samples with a minimum sequence covering nucleotides 16019-16400 were compared to the Cambridge Reference Sequence (Anderson et al., 1981) in order to identify the polymorphic nucleotide positions thus defining the mtDNA haplotype consisting of the nucleotide positions where substitutions, insertions or deletions occurred plus details of the base change.

Genetic data

The cheek cells analysed in this paper were collected by the author (see plate 01) during intermittent field trips in selected Cameroonian communities. Two hundred mouth swaps (from which cheek cells were extracted) were collected from each of the following locations (where the staple crops consumed were also studied) – Buea, Mbo, Foumban, Banyo, Poli and Guider. Six hundred samples were also collected from Kousseri and its environs. To inform the analyses, 400 extra samples were collected in Kribi and Lolodorf. A total of 2,200 consenting self-identified males above the ages of 18 and unrelated at the paternal grandfather level were therefore anonymously sampled. For the purpose of this analysis, cheek cells were extracted from 1952 (NRY)/1957 (mtDNA) individuals belonging to the thirty groups listed in Table 01 above, DNA was extracted and NRY and mtDNA haplotypes analysed at The Centre for Genetic Anthropology (TCGA), University College London in accordance with their standard procedures.

Plate 1

Photograph of the author collecting DNA sample during fieldwork

Source: Snapped with the help of author's camera during fieldwork, 10 August 2015.

Statistical methods

hNRY and hmtDNA were calculated as Gene Diversity, (*h*, the probability of randomly sampling two different alleles from a population), estimated using the unbiased formulae of Nei 1987 that is $h = n(1 - \underline{\Sigma}\chi_i^2)/(n-1)$, (Nei, 1987) where n is the sample size and χ_i^2 is the squared frequency of the *i*th allele. Calculations were performed in Excel. Means were compared applying the Mann-Whitney, unpaired t (with and without the Welch correction) and Kruskal-Wallis (non-parametric

ANOVA) tests and correlations assessed by Pearson and Spearman Rank tests using the Graph Pad InStat 3 program.

4. Results

These results contain the names, gene diversity values (hNRY and hmtDNA), geographic distance south of Lake Chad, sample numbers analysed for each group, principal staple crop, total estimated number of individuals, traditional political system, descent system and principal religion of each of the 27 Cameroonian groups for which NRY and mtDNA were typed for 15 or more samples. Distance south of Lake Chad ranged from 108Km to 1,234KM (average 965KM), sample numbers ranged from 15 to 366 (mtDNA, average 71) and 16-359 (NRY, average 71), gene diversity h ranged from 0.958-1.000 mtDNA, average 0.986 and NRY 0.703-0.983, average 0.902, estimated population sizes ranged from 28,000 to 20 million average 1.777million. There were five separate crops identified as the principal staple (cassava, cocoyam, colocasia, maize and millet), three principal religions (Islam, Christianity and Traditional African Religion), three traditional political systems (centralised, decentralised and mixed) and three descent systems (paternal, maternal and both maternal and paternal). Map 1 shows the geographic location of all the 27 groups.

In the complete dataset (see table 02) hNRY and hmtDNA are weakly, but significantly, correlated (Pearson) r^2=0.1437 p=0.0467 but there was no correlation between sample size and gene diversity (hNRY: Pearson p = 0.3531, Spearman p = 0.9731; hmtDNA: Pearson p = 0.7159, Spearman p = 0.4633). hmtDNA was strongly associated with the principal staple crop (Kruskal-Wallis Test p=0.0067) while NRY was not quite significant p=0.0773. The order of diversity was average hmtDNA: colocasia 0.995, millet 0.991, maize 0.990, cocoyam 0.983 and cassava 0.969. The equivalent order for hNRY was maize 0.935, millet 0.922, colocasia 0.907, cocoyam 0.886 and cassava 0.846.

Neither hNRY nor hmtDNA were correlated with population size (NRY: Pearson r^2 = 0.087, p=0.1363; Spearman r = 0.3711, p=0.0586. mtDNA: Pearson r^2 = 0.085, p=0.140; Spearman r = 0.2651, p=0.181).

Both hNRY and hmtDNA were significantly associated with political system (NRY: unpaired t test 0.0047 centralised mean = 0.943, decentralised mean = 0.870. mtDNA: Mann-Whitney test 0.0138 centralised mean = 0.992, decentralised mean = 0.980) as they were with the principal religion NRY: Kruskal-Wallis Test p = 0.0073; Islam mean 0.946, African Traditional 0.882, Christianity 0.872. mtDNA: Kruskal-Wallis Test p = 0.0217; Islam mean 0.992, African Traditional 0.988, Christianity 0.978). However there was no significant statistical association between either hNRY or hmtDNA and descent system (NRY p =0.0931, mtDNA p = 0.2479 both unpaired t test). Similarly population size was not associated with any of: political system p = 0.1788, principal staple crop p = 0.3733, principal religion p = 0.9039 (Kruskal-Wallis Test), with descent system p = 0.4142 (Mann-Whitney Test) nor with distance south (Pearson p = 0.1171, Spearman p = 0.0784). There was evidence of Islam being associated with a traditional centralised political system and Christianity with a non-centralised system (Fisher

exact test p = 0.0237). Groups practicing Traditional African Religion divided 4 centralised and 4 not centralised. There was significant association between distance south from Lake Chad and both hNRY and hmtDNA (all $p < 0.05$).

Given the considerable differences in environment and religious affiliation with distance south from Lake Chad the data were also analysed in two separate sets: less than 1,000Km south of Lake Chad (20 groups) and more than 1,000Km south of Lake Chad (7 groups). In addition given the possible considerable difference in the demographic history of the Pygmies the data were analysed both including and excluding them in the more than 1,000Km set.

With or without including Pygmies population size in both the under and over 1,000Km sets was not associated with distance south ($p > 0.1$). Comparing the northern and southern groups and excluding Pygmies there were significant differences between them in mean hNRY and hmtDNA (NRY: Mann Whitney Test p = 0.0091; northern group 0.991, southern group 0.982. mtDNA unpaired t test with Welch correction p = 0.0377, northern group 0.925, southern group 0.878). When Pygmies were included the differences were still significant (NRY: Unpaired t Test with Welch correction p = 0.0007; mtDNA unpaired t test p = 0.0203).

There were also significant differences between the two regions in descent system p = 0.012 Fisher Exact Test (northern group paternal 14, maternal and paternal 0; southern group paternal 8, maternal and paternal 5) and principal religion p = 0.0119 Fisher Exact Test (northern group Islam 9, African Traditional Religion 3, Christian 2; southern group Christian 6). There were no significant differences in population size (p = 0.271 Mann-Whitney Test) or political system (p = 0.391 χ^2) between the two regions.

Analysing NRY diversity in the northern set there was a significant association between distance south of Lake Chad with diversity decreasing (Spearman r = -4622 p = 0.0402) while there was no association between hNRY and, respectively, population size, religion and political system ($p > 0.13$). There were too few data points to assess association with descent system. The equivalent analysis of mtDNA diversity found no association with distance south of Lake Chad, population size, religion or political system ($p > 0.25$).

In the southern set the most notable observation was of uniformity of religion (all are Christian except Pygmies and all have a traditional decentralised political system. In addition all groups grew either colocasia, cocoyam or cassava as their principal crop. For NRY diversity there was no statistically significant association between distance south from Lake Chad, population size or descent system ($p > 0.15$). With hmtDNA while there was no statistically significant association with distance south when Pygmies were included in the analysis ($p > 0.08$) there was when they were excluded (Pearson $r^2 = 0.7607$ p = 0.0235). There was no significant association with population size ($p > 0.05$) or descent system (unpaired t test with Welch correction p = 0.878).

The data were next analysed defining a Northern group comprising 14 groups the most southern of which was the Nso and a Southern group that excluded Pygmies comprised 12 groups the most northern of which was the Bafut. The Northern group may be characterised as follows: hNRY range: 0.703-0.983 average 0.925; hmtDNA range: 0.981-1.000 average 0.991; principal crop one group grows maize and the remainder millet; estimated population size ranged from 28,000 to 20 million with a mean of 2.97 million; all groups were patrilineal while 7 were classified as centralised, 5 decentralised and 2 semi centralised; 9 primarily practised Islam, 3 Traditional African Religion and 2 Christianity. The Southern group (excluding Pygmies) are characterised as follows: hNRY range: 0.823-0.934 average 0.878; hmtDNA range: 0.958-1.000 average 0.981; principal crop one group grows maize, four cocoyam, four cassava and three colocasia; estimated population size ranged from 358,000 to approximately 2.5 million with a mean of approximately half a million; 7 groups were patrilineal and 5 patrilineal and matrilineal; 4 were classified as centralised, 7 decentralised and 1 semi centralised; only 1 primarily practised Islam, 3 Traditional African Religion and 8 Christianity.

Both hNRY and hmtDNA were significantly different between the Northern and Southern sets whether or not Pygmies were included in the analysis (unpaired t test all < 0.025) within all cases the Southern set being less diverse. However in the Northern set there was no statistically significant association between either hNRY or hmtDNA and, respectively, distance south of Lake Chad, estimated population size, religion or traditional political structure (NRY all $p > 0.07$; hmtDNA all >0.21). In contrast for the Southern set mtDNA was strongly associated with distance south from Lake Chad (both Pearson $r^2 = 0.9245$ and Spearman $r = -0.9507$ both $p < 0.0001$) but not with population size ($p > 0.29$). There was no similar association with either religion or descent system ($p > 0.19$). There was, however, a just significant association with the traditional political system ($p = 0.0424$ Mann Whitney Test with centralised societies being more diverse $h = 0.992$ than decentralised societies 0.974). The pattern for NRY was broadly similar with the association of hNRY and distance south being significant or almost significant (Pearson $r^2 = 0.4882$ $p = 0.0115$ and Spearman $r = -0.5754$ $p = 0.5754$). There was a significant association between religion and hNRY (unpaired t test $p = 0.0003$; Christian 0.855 African Traditional Religion 0.930) but not between hNRY and either the descent system or traditional political structure.

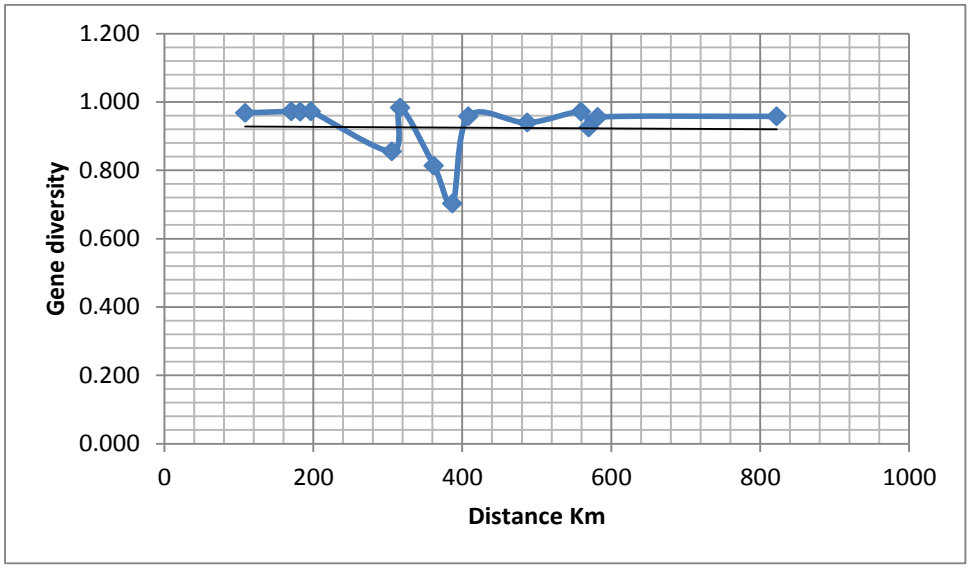
Discussion

It is clear from the collected data and analyses undertaken in this chapter that the peoples of the northern part of Cameroon and the region to the south are very different both in their economic activities and social structures but not in the average sizes of the ethnic groups living in those areas. The north represented by peoples living north of latitude 5.96 is characterised by patrilineal societies, a majority of ethnic groups practicing Islam and the prevalence of seed crops. The southern groups however are mostly decentralised, have a majority of predominantly Christian affiliation, and a substantial minority of decentralised traditional systems. In all but one case they grow root crops with the single exception growing maize being the mainly Muslim Bamun.

ISSN 2709-3409 (Online)

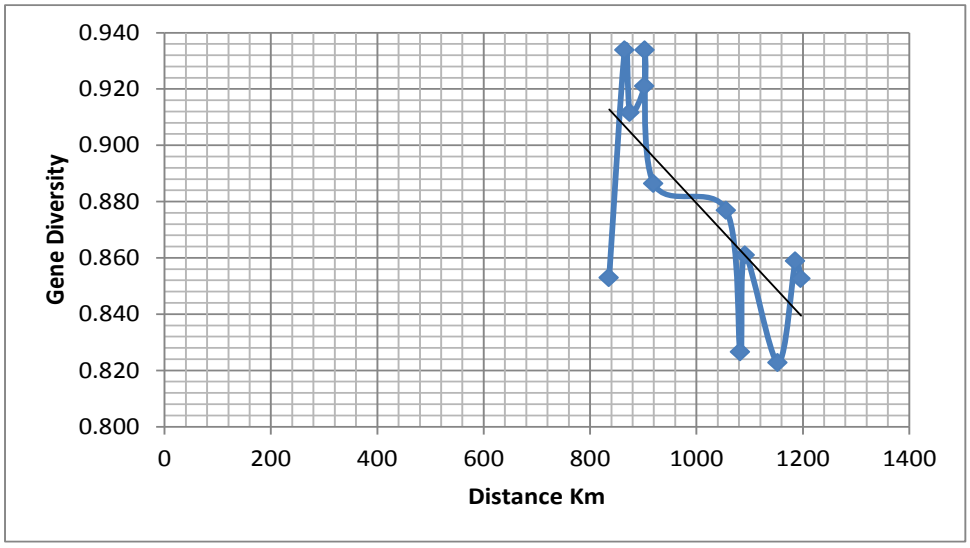
The statistical difference in gene diversity between the northern group and the southern group is striking (hNRY p = 0.0091 hmtDNA 0.0377) with both hNRY and hmtDNA, as recorded above, being more diverse in the north. Further there is a marked difference between the slope of decline in diversity with distance south observed in both hNRY and hmtDNA with the south much steeper in both cases (see figure A-D).

Figure A: hNRY IN THE NORTHERN SET



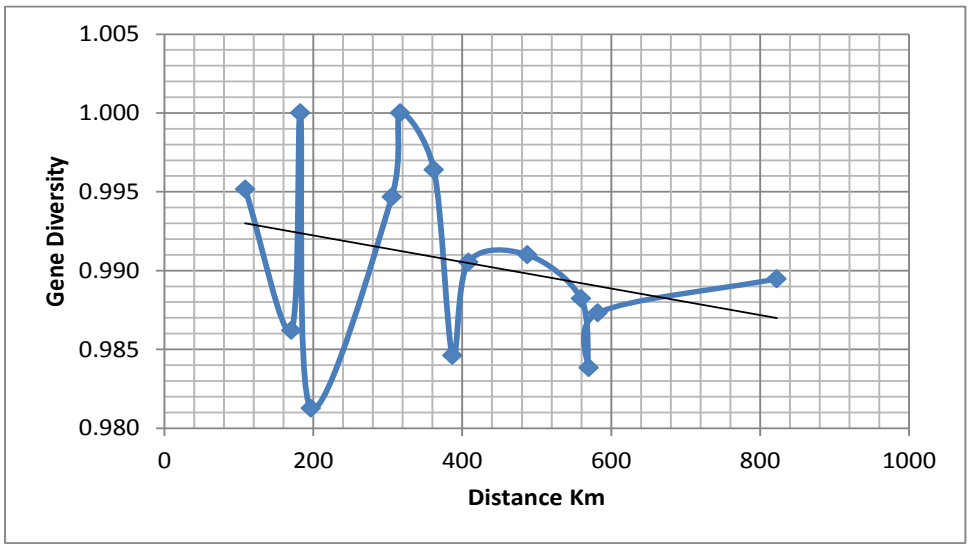
Source: Computed by the Author with the assistance of genetic historians at The Centre for Genetic Anthropology, University College London.

Figure B: hNRY IN THE SOUTHERN SET



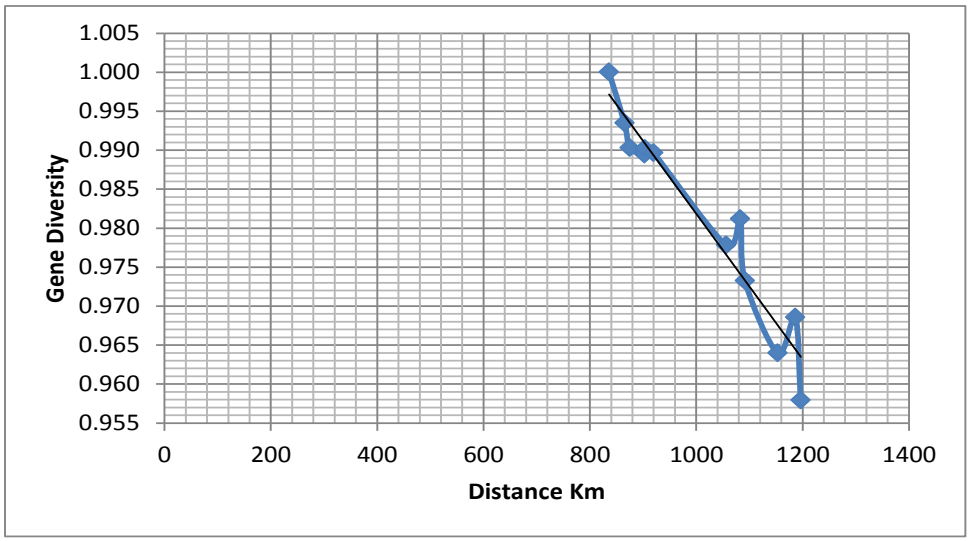
Source: Computed by the Author with the assistance of genetic historians at The Centre for Genetic Anthropology, University College London.

Figure C: hmtDNA IN THE NORTHERN SET



Source: Computed by the Author with the assistance of genetic historians at The Centre for Genetic Anthropology, University College London.

Figure D: hmtDNA IN THE SOUTHERN SET



Source: Computed by the Author with the assistance of genetic historians at The Centre for Genetic Anthropology, University College London.

5. Conclusion

On the bases of the above historical genetic analysis it can be deduced that there is a strong correlation between crops grown (seed as opposed to root) and gene diversity (hmtDNA unpaired t test p = 0.0074; means: root 0.098, seed 0.991; hNRY unpaired t test with Welch correction p = 0.0426; means: root 0.875, seed 0.924) with seed being associated with increased diversity in both genetic systems. However association and causation are not identical phenomena and it would be premature to conclude that in

this study one was causative of the other. Much larger data sets than it was possible to collect or analyse in this pioneering study will be required to separate the relative contributions to genetic diversity made by the different economic and social practices studied in this survey. This study is the most extensive yet undertaken on sex-specific diversity in Cameroon, an important region in the context of the peopling of sub-Saharan Africa. It has established clear differences in diversity of both paternally mediated and maternally mediated genetic diversity between the northern region and the area to the south, close to that from which, on linguistic evidence, it has been suggested that the expansion of the Bantu-speaking peoples began some 3,000 to 5,000 years before present (Vansina, 1995).

It has been hypothesized in this paper that the high genetic diversity among northern Cameroonian ethnic groups emanated from the Fulaini jihads of the early 19th century granted the fact that the jihads led to an unprecedented influx of huge bands of reproductive Fulani. Conversely, the low genetic diversity among southern Cameroonian ethnic groups can conveniently be attributed to the Bimbian Trans-Atlantic Slave Trade that drained and forcefully dispatched a very significant number of reproductive men and women to the New World.
An awareness of the genetic diversity of Cameroonian populations is important for several reasons: first it can help pharmaco-geneticists to understand why some drugs are more efficacious among certain populations than in others, second it can help to understand lactose tolerance and intolerance (the reason why some Cameroonian populations can digest uncooked cow milk and others cannot), third, a knowledge of Cameroonian genetic diversity can enable the police to understand why some Cameroonian populations are more prone to criminality. Lastly, a mastery of DNA study helps to illuminate when ethnic groups split up within a population.

6. References

Books

Ardener, Edwin. (1956), Coastal Bantu of the Cameroons, London, International African Institute.

Bah, T. M., Ghomsi, E. (février 1986) ''Problématique des transmissions des techniques à travers le Sahara du 8e au 16e siecle,'' in AFRIKA ZAMANI no. 16 and 17, pp.22-35.

Cavalli-Sforza, L. L., Menozzi, P. and Piazza, A. (1996). The History and Geography of Human Genes, Princeton, Princeton University Press.

Daunn, D. S., (2001). Statistics and data analysis for the behavioural sciences, New York, McGraw-Hill Higher Education.

Foley, R., (1998). Humans before Humanity: An Evolutionary Perspective, Massachusetts, Blackwell Publishers Ltd.

Hartl, D.L., Clark, A. G., (1977) Principles of Population Genetics, Third Edition, Sunderland, Massachusetts, Sinauer Associates, Inc. Publishers.

Mange, E. J., Mange, A. P., (1999), Basic Human Genetics, Second Edition, Sunderland, Massachusetts, Sinauer Associates, Inc., 199Nei, M. (1987). Molecular Evolutionary Genetics, Columbia, Columbia University Press.

Smith, J. M. (1989). Evolutionary Genetics, New York, Oxford University Press.

Articles

Bah, T. M., Ghomsi, E. (février 1986) ''*Problématique des transmission des technique à travers le Sahara du 8e au 16e siecle,*'' in *AFRIKA ZAMANI* no. 16 and 17, pp.22-35.

Bradman, N., Thomas, M., (1998) "The Y Chromosome in the study of Human Evolution, Migration and Prehistory,'' *Science Spectra,* Issue 14, 1998, pp.32-37.

Vansina, J., (1995). "The Contributions of Linguistics to the Study of History in Africa", *Journal of African History*, 36 (2): 173-95.

Wilson, I. J., M. Weale, M. E. and Balding, D. J. "Inferences from DNA data: Population histories, evolutionary processes, and forensic match probabilities,'' *Nature,* November 1, 2000, pp.1-24.