

SHORT REPORT

Co-introgression of Y-chromosome haplogroups and the sickle cell gene across Africa's Sahel

Rihab E Bereir¹, Hisham Y Hassan¹, Niven A Salih¹, Peter A Underhill¹,
Luigi L Cavalli-Sforza², Ayman A Hussain¹, Dominic Kwiatkowski³ and Muntaser E Ibrahim^{*,1}

¹Institute of Endemic Diseases, Medical Campus, University of Khartoum, Khartoum, Sudan; ²Department of Genetics, School of Medicine, Stanford University, Stanford, CA, USA; ³Wellcome Trust Center for Human Genetics, Oxford, UK

The Sahel that extends from the Atlantic Ocean to the Ethiopian highland is a historical reservoir of Africa's cultures and grandest populations and a known arena of ancient and recent migrations. We are interested in the issue whether such migrations were also carriers of genetic traits and whether this introgression could be associated with population genetic markers. Based on analysis of Y-chromosome haplogroups, we present evidence that the sickle gene, one of the major protective polymorphisms known in malaria, has in fact found its way only recently to the gene pool of the populations in eastern Sahel. We discuss the possible dynamics of the process and give estimates of the age of the introduction of the S allele into eastern Sahel. *European Journal of Human Genetics* (2007) 15, 1183–1185; doi:10.1038/sj.ejhg.5201892; published online 15 August 2007

Keywords: Y-chromosome; haplogroups; sickle gene; Sudan

Introduction

The origin of malaria-resistant sickle cell genetic variants from either single or multiple events was controversial until recently.¹ The debate was partly settled in favor of multiple origins; owing to β S haplotypes in linkage disequilibrium with the sickle mutation.^{2,3} Complementary evidence for parallel origins of sickle cell is discernable in ethnicity, some notable examples being the concurrency of gene population spread attributed to expansions of Iron Age Bantu speakers both south and eastwards in Africa emerging from the area of present-day Cameroon, and subsequently across the Atlantic by events associated with the 'Middle Passage'.

Here we report Y-chromosome co-introgression with the S gene in the eastern part of Africa's Sahel, where it is known that the gene existed previously in low frequencies^{4,5} in contrast to being more prevalent among west African migrants.

Materials and methods

A total of 237 unrelated males (some of whom are known sicklers) were typed, for the sickle-cell mutations using amplification refractory mutation system PCR and the results confirmed by protein electrophoresis. The same sample was typed for the major Y chromosome haplogroup-defining mutations. The populations studied included Hausa, Masalit, Borgu as well as other groups and populations inhabiting the area across Africa's Sahel into Chad and Sudan. Populations were classified according to their linguistic affiliation, which comprised the main linguistic families of Africa (Afro-Asiatic, Niger Kordofanian and Nilo-Saharan). The population sample was selected based on preliminary knowledge of the distribution of the S gene among various groups in the Sudan.^{4–6} The sample was further classified to include the Hausa as a separate group due to the high frequency of a particular distinctive haplogroup within this group (P25) whose closest phylogenetic molecular ancestor, haplogroup R1-M173* marks a back migration to Africa from west Eurasia.⁷ The Hausa, originally of western Africa, have migrated to the eastern Sahel during the past 300 years, employed mainly in agricultural economies (Figure 1).

*Correspondence: Professor M Ibrahim, Department of Molecular Biology, Institute of Endemic Diseases, University of Khartoum, PO Box 102/Qasser Street, Kartoum, Sudan.

Tel: +24 99 12576418; Fax: +24 91 83793263;

E-mail: mibrahim@iend.org

Received 30 March 2007; accepted 11 May 2007; published online 15 August 2007

We employed an extension of the coalescence to estimate the possible age and date of introduction of the S gene into the eastern Sahel, using the following formula

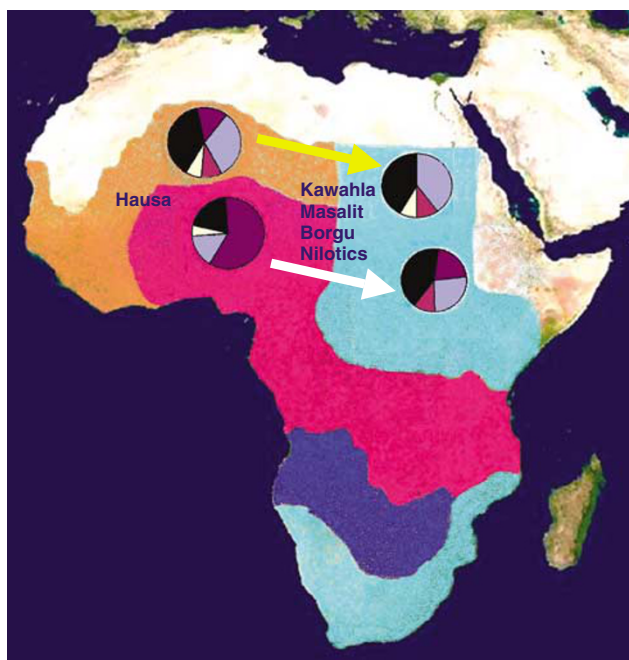
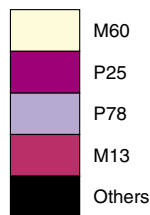


Figure 1 Map of Africa showing the hypothetical movement of Y-chromosome haplogroups across the Sahel. The yellow color of the arrow is for the haplogroups in non-sicklers and the white arrow is for sicklers. The various colors represent approximate estimations of the frequency of the S gene in the continent with red representing the highest frequency followed by the dark and light blue for lower frequencies.

Table 1 The distribution of Y-chromosome haplogroups among sicklers and non-sicklers classified according to their ethnic/linguistic/affiliations to two of the major linguistic families in Africa, AA and NS

| Y Haplogroup | Sicklers | | | Non-sicklers | | |
|--------------|----------|------------|-----------|--------------|----------|-----------|
| | Hausa | OAA | NS | Hausa | OAA | NS |
| P25 | 26 (59%) | 8 (24.2%) | 4 (9%) | 5 (22.7%) | 0 (0.0%) | 0 (0.0%) |
| M78 | 7 (16%) | 5 (15%) | 15 (33%) | 6 (27%) | 28 (64%) | 8 (18%) |
| M13 | 0 (0.0%) | 0 (0.0%) | 7 (15.5%) | 3 (14%) | 0 (0.0%) | 1 (24.4%) |
| M267 | 3 (7%) | 5 (15%) | 1 (2%) | 0 (0.0%) | 1 (0.0%) | 4 (9%) |
| M60 | 3 (7%) | 0 (0.0%) | 1 (2%) | 3 (13.6%) | 3 (0.0%) | 1 (1.8%) |
| Others | 5 (11%) | 15 (45.4%) | 17 (38%) | 5 (22.7%) | 16 (36) | 21 (47%) |
| Total | 44 | 33 | 45 | 22 | 48 | 45 |

Abbreviations: AA, Afro-Asiatic; OAA, other Afro-Asiatic speakers; NS, Nilo-Saharan.

The Hausa although AA speaker were placed separate from OAA speakers due to the high frequency of a signature haplogroup (P25) among the Hausa.

Pearson χ^2 44.33 $P < 0.000$ d.f. = 5. Tests are adjusted for all pair-wise comparisons within each inner most sub-table using the Bonferroni correction.

provided an equal assortment and no gender preferential introduction of the gene:

$$X(t) - Y = (1 - a)^t(1 - Y)$$

where $X(t)$ is the frequency of the Y haplogroup among bearers of S mutation in generation t and Y is the frequency of Y haplogroup in the chromosome and a is the decay in association due to assortment.

Results

Table 1 shows the distribution of the S gene in 122 cases and 115 controls with Y chromosome genotyped for the major haplogroups. Two haplogroups, R1b-P25 and E3b-M78, display the highest frequency, and were differentially distributed among the various groups and between sicklers and non-sicklers. The distribution of the observed values were significantly different from the expected (Pearson χ^2 44.9 $P < 0.000$). The association of the P25 with the S gene in Hausa, which barely holds ($P = 0.07$) in spite of the relatively large effective population size of the Hausa (Hassan *et al*, unpublished), might indicate that the S mutation may have been affiliated with a male founder belonging to Y haplogroup R1b-P25. The observed frequency among S-gene carriers in populations other than Hausa makes a strong case for gene flow from the Hausa to local groups in Sudan.

The dating estimates based on a formula that is an extension of the coalescence gives a recent figure of 1–3 generations for the introduction of the gene and associated haplotypes to eastern Sahel.

Discussion

Patterns of genetic diversity are the result of locus-specific forces (natural selection) and population-level forces (eg demographic growth and range expansion). When multiple independent loci correlate with geography, population-level forces are likely responsible. Conversely, when

patterns diverge, natural selection is implicated in modulating observed diversity.

Mitochondrial DNA and Y-chromosome diversity closely correlate with geography and these loci have recently become the markers of choice in defining both provenance and trajectory of population movements and also provide insights into population substructure and group membership. By typing both the haploid Y chromosome and the S gene in the same samples it should be possible to test for common demography as well as detect affinities of particular S variants observed in Sudan to other regions of Africa.

Our results suggest that the sickle cell gene may have been preferentially introduced through males of migrating west African tribes (Figure 1), particularly Hausa-Fulani, and Bagara in the large migrations that began in the eighteenth century and escalated during the nineteenth and early twentieth century. The estimates of a recent figure of 1–3 generations for the introduction of the gene and associated haplotypes to eastern Sahel, is consistent with demography during the past 100 years and with a hypothesis of a recent origin of malaria as a major human infection.^{6,8,9}

Acknowledgements

This work is part of the activities of the BioMalPar European Network of Excellence supported by a European grant (LSHP-CT-2004–

503578) from the Priority 1 'Life Sciences, Genomics and Biotechnology for Health' in the 6th Framework Programme.

References

- 1 Pagnier J, Mears JG, Dunda-Belkhdja OD *et al*: Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proc Natl Acad Sci USA* **81**, (1984); 1771–1984.
- 2 Lanclos KD, Oner C, Dimovski AJ, Gu YC, Huisman TH: Sequence variations in the 5' flanking and IVS-II regions of the G gamma- and A gamma-globin genes of 8S chromosomes with five different haplotypes. *Blood* (1991); **77**: 2488–2496.
- 3 Oner C, Dimovski A, Olivieri N *et al*: β s haplotypes in various world population's. *Hum Genet* (1992); **89**: 99–104.
- 4 Elhassan AM, Godber MG, Kopec AC, Mourant AE, Tills D, Lehmann H: The inherited blood factors of the Beja of the Sudan. *Man* (1968); **3**: 272.
- 5 Bayoumi RA, Taha TS: A study of some genetic characteristics of the Fur and Baggara tribes of the Sudan. *AM J Phys Anthropol* (1985); **67**: 363–370.
- 6 Mohammed AO, Attalla B, Bashir FMK *et al*: The relationship of the sickle cell gene to the ethnic and geographic groups populating the Sudan. *Community Genet* (2006); **9**: 113–120.
- 7 Cruciani F, Santolamazza P, Shen P *et al*: A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* (2002); **70**: 1197–1214.
- 8 Tishkoff SA, Varkonyi R, Cahinhinan N *et al*: Haplotype diversity and linkage disequilibrium at human G6PD: Recent origin of alleles that confer malaria resistance. *Science* (2001); **293**: 455–462.
- 9 Volkman SK, Barry AE, Lyons EJ *et al*: Recent origin of *Plasmodium falciparum* from a single progenitor. *Science* (2001); **293**: 482–483.