BJMG 15/2 (2012) 7-14
10.2478/bjmg-2013-0001

ORIGINAL ARTICLE

DETECTING SEX-BIASED GENE FLOW IN AFRICAN-AMERICANS THROUGH THE ANALYSIS OF INTRA- AND INTER-POPULATION VARIATION AT MITOCHONDRIAL DNA AND Y- CHROMOSOME MICROSATELLITES

Battaggia C[1], Anagnostou P[1,3], Bosch I[4], Brisighelli F[4], Destro-Bisol G[1,3], Capocasa M[2,3,*]

***Corresponding Author:** Dr. Marco Capocasa, Dipartimento di Biologia e Biotecnologie "Charles Darwin", Università "La Sapienza", Roma, Italy, P.le A. Moro 5, 00185 Roma, Italia; Tel.: +39-6-4991-2725; Fax: ++39-6-4991-2771; E-mail: marco.capocasa@uniroma1.it

ABSTRACT

This study reports on variations at the mitochondrial DNA (mtDNA) hypervariable region 1 (HVR-1) and at seven Y-chromosome microsatellites in an African-American population sample from Chicago, IL, USA. Our results support the hypothesis that the population studied had undergone a European male-biased gene flow. We show that comparisons of intra- and inter-population diversity parameters between African-Americans, Europeans and Africans may help detect sex-biased gene flow, providing a complement to quantitative methods to estimate genetic admixture.

Keywords: Admixture, African-Americans, Mitochondrial DNA (mtDNA), Sex-biased gene flow, Y-chromosome

INTRODUCTION

The European contribution to the gene pool of African populations deported to the United States of America in the course of the Atlantic slave trade may be regarded to as a paradigmatic case of gene flow in human populations [1,2]. Coherently, studies of genetic structure of present day African-Americans have attracted a particular interest in molecular anthropology [3]. The earliest investigations based on protein loci estimated a 4.0-30.0% proportion of European genes in African-Americans, pointing to a higher admixture in northern rather than in southern US regions [6]. More recent studies, based on autosomal DNA poly-morphisms, highlighted the level of admixture for northern US populations to be lower than previously thought, and the lack of a close relationship between latitude and extent of European admixtures [7]. The introduction of uni linear DNA polymorphisms of mitochondrial DNA (mtDNA) and non recombining portions of Y-chromosome, has made it possible to separately study the male and female European contribution to the African-American gene pool [7-9]. Finally, further refinements have been obtained with the introduction of genome wide approaches [10,11].

In a previous study, we used 10 autosomal microsatel-lites and an Alu polymorphism to explore the genetic structure of an African-American population from Chicago, IL, USA [12]. In this study, we analyzed the variation at the mtDNA hypervariable region 1 (HVR-1) and at seven microsatellites of the Y-chromosome in the same population sample. Using a broad population dataset, we showed that comparisons of intra- and inter-population diversity

Supplementary material avalible online at:
http:\\bjmg.edu.mk/tekstovi/bjmg15(2),2012/15(2)-01.pdf

[1] Dipartimento di Biologia Ambientale, Università "La Sapienza", Roma, Italia

[2] Dipartimento di Biologia e Biotecnologie "Charles Darwin", Università "La Sapienza", Roma, Italia

[3] Istituto Italiano di Antropologia, Roma, Italia

[4] Istituto di Medicina Legale, Università Cattolica, Roma, Italia

parameters between African-Americans, Europeans and Africans may help detect sex-biased gene flow, providing a complement to quantitative methods to estimate genetic admixtures.

MATERIALS AND METHODS

A total of 50 individuals (23 males and 27 females) was available for this study. Appropriate informed consent was obtained from all individuals participating in the study. Sequencing of the HVR-1 of mtDNA from nucleotide (nt) positions 16024 to 16383, and determination of size variations at loci DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393, was performed as previously described [13,14]. The amplified products were analyzed in polyacrylamide denaturing gels using a semi-automated DNA sequencer (A.L.F. express; Pharma-cia Biotech, Uppsala, Sweden). Allelic and internal standards were used for microsatellite typing.

The dataset for comparison was built selecting data from the mtDNA and Y-chromosomal literature for European and 10 African populations. We have preferentially used widely dispersed populations which could have contributed to the gene pool of present day African-Americans and/or have not undergone drift events in their evolutionary history (Supplementary Table 1). Moreover, data for four African-American samples were considered [15-17].

Parameters of intra-population diversity [haplotype diversity (HD) and mean number of pairwise differences (MNPD)] and genetic distances between populations were calculated using the Arlequin software ver. 3.5 [18]. Given the relatively high differentiation expected between African and European haplotypes, molecular methods were used in addition to the haplotypic ones to detect further signatures of admixture.

RESULTS AND DISCUSSION

Mitochondrial DNA and Y-Chromosomal Variation. DNA sequences of the HVR-1 were determined in 50 unrelated individuals (Table 1). We found nucleotide differences at 63 out of 360 positions with respect to the 'consensus' sequence [19], five of which were transversions [16114 (A>C); 16183 (A>C); 16188 (C>G); 16258 (T>A)]. In total, we were able to define 40 different lineages, with frequencies ranging from 0.002 to 0.008. We obtained a HD value of 0.987 ± 0.008 and a MNPD value of 7.673 ± 3.637.

Y-Chromosomal haplotypes, built using seven microsatellite loci, were determined in 23 individuals (Table 2). Nineteen different haplotypes were observed, with frequencies ranging from 0.042 to 0.125. We obtained a HD value of 0.996 ± 0.014 and a MNPD value of 4.383 ±2.246.

Intra- and Inter-population Diversity in African-American and African Populations. Since the seminal work by [20], the approach based on the simultaneous use of mtDNA and Y-chromosomal polymorphisms has been used in numerous studies of human genetic diversity [21]. This approach may also turn out to be useful in the case of African-American populations. In fact, the availability of uni linear DNA polymorphisms offers the opportunity to detect possible signatures of a sex-biased gene flow from Europeans to African-Americans. Intriguingly, whereas a male-biased gene flow was expected on the basis of historical knowledge, demographic data indicate that since 1960, most of the mixed marriages involved African-American males to females of European origin [22].

In a study carried out among African-American students at the Texas University at Austin, Austin, TX, USA, Hsieh and Sutton [23] compared admixture estimates based on mtDNA and protein loci, concluding that there is no evidence of a European male-biased gene flow. On the other hand, the latter case was supported by more recent studies carried out in a number of African-American populations from different areas and based on population specific mtDNA and Y-chromosome alleles [7-10].

Estimates of sex-specific admixture suffer from two different orders of limitations. Obtaining reliable African parental populations is difficult for three main reasons. First, present day African-Americans originate from populations scattered in a wide area from the central and western parts of the continent. Second, it is tricky to identify the populations in continuity with those from which slaves were taken to the US, as their demography was reshaped by slave capture and forced migrations [24]. Third, there is also substantial variation among sub-Saharan groups for mtDNA and Y-chromosome polymorphisms [25,26], so that results may vary substantially depending on

Table 1. Variable nucleotide positions of the HVR1 of mtDNA control region in 50 AfricanAmerican samples compared to the Cambridge reference sequence. The "i" stands for insertion.

```
          1111111111111111111111111111111111111111111111111111111111111111111
          6666666666666666666666666666666666666666666666666666666666666666666
          0000001111111111111111111112222222222222222222222222233333333333333
          3588990112224456678888899901123344455667778899999990011122256666
          8116235144692836821378902393430915968450184601234690918905770128
                                    i
Ref       AAATTTTCCTTGCCGACTAACCTCC-TGCCACACTCACACTCACCCCACCAAATAGCTCTCGTT
AAm01     ............................T...........T......T...G.......... 1
AAm02     ...........A.T..TC..TGC......TG.......................C..T....... 1
AAm03     ............................T............................T....... 1
AAm04     ............................T...........T......T...G..........C 1
AAm05     ..........C.........T.C......T.......T.T.T.....G.....C.......... 1
AAm06     ............................T...........T......T...G.......... 1
AAm07     .....C.T..C.........T.C......T.T.......T.T.....G.....C.......... 1
AAm08     ...C..............G..........T...........T......T...G.......... 1
AAm09     ...........A.T..TC..TGC......TG.......................C..T....... 1
AAm10     ..G.........................T...........T....................... 1
AAm11     .....C...........C.C.........T............................T....... 1
AAm12     .....C...CC.................T....................................... 1
AAm13     .........C.....G............T....................................... 1
AAm14     .........C..................T...........T.....................C. 1
AAm15     ........................C..T.......................C.......... 1
AAm16     ........TC..................TT..........CT.....................C. 1
AAm17     ........A..A..........C.T....T...........T......T...G.......... 1
AAm18     ............................T................................... 1
AAm19     ...........T.........C......T.............................C....C. 1
AAm20     .G.....................T.....T.......T....................... 1
AAm21     ..................C....C......T.........................T....... 1
AAm22     ......................C......T.........................T....... 1
AAm23     ............................T......T.............G.....T....... 1
AAm24     .........C..................T...........T.....................C. 1
AAF1      ............................T...........T......T...G.......... 1
AAF2      .........C..................T..........CT.....................C. 1
AAF3      ...C......C...A.....................................TT.............. 1
AAF4      ....C.......................T...........T.TT...T.........T..... 1
AAF5      ............................T...........T.....G......G......... 1
AAF6      .....C.T....................T....C.........T........C.A......C. 1
AAF7      ............................T............................T....... 1
AAF8      ..........C.........T.C......T.......T.T.T.....G.....C.......... 1
AAF9      .........C..................T...........T.....................C. 1
AAF10     ......................C......T...........T......T...G......C.... 1
AAF11     ........................C..T.......................C.......... 1
AAF12     ............................T...T.....T.T....................... 1
AAF13     ...C.......A.T.....-..C..C..............T.........G............ 1
AAF14     ............................T...........T......T...G.......... 1
AAF15     ...........A.T..TC..TGC......TG....T..................C..T....... 1
AAF16     ...........A........T.C.....TT........C..T.A.T..T....C......T... 1
AAF17     G..C.......A........T.C......T...........TG....GT....C......T... 1
AAF18     ......C..............C.....................T......T.............. 1
AAF19     .....C.........G.............T......C.............................. 1
AAF20     ......................C.T....T...........T......T...G.......... 1
AAF21     ............................T........T..............C.......... 1
AAF22     ......................C.T....T...........T......T............C.. 1
AAF23     ...C.......A........T.C......T..G........T...T..T....C......T... 1
AAF24     ........A..A..............A.T...........T....................... 1
AAF25     .........C............C......T...........T....T.................. 1
AAF26     .........C..................T...........T.....................C. 1
```

the populations chosen as the parental ones. A further source of confusion may be created by geneticists and molecular anthropologists when they use inaccurate ethnolinguistic labels and/or assume particular populations as representative of wider groups [27].

The second considerable problem is in the large statistical uncertainty of admixture estimates based on mtDNA and Y-chromosome polymorphisms. Given their intrinsic nature of single loci in evolutionary terms, both the uni linearly transmitted ge-

Table 2. Seven loci Y chromosome haplotype frequencies in 23 African-American subjects.

Sample	Number of Haplotypes	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	Frequency (%)	Standard Deviations
AA01	2	17	14	31	21	10	11	15	0.0870	0.0601
AA02	3	16	13	30	21	10	11	14	0.1304	0.0718
AA03	1	14	13	29	24	10	13	13	0.0435	0.0435
AA04	1	15	14	33	21	10	11	13	0.0435	0.0435
AA05	1	16	14	31	25	10	11	13	0.0435	0.0435
AA06	1	14	12	28	25	10	11	13	0.0435	0.0435
AA07	1	15	14	32	21	10	11	13	0.0435	0.0435
AA10	1	14	12	28	24	11	13	13	0.0435	0.0435
AA11	1	15	14	31	21	10	10	13	0.0435	0.0435
AA12	1	17	13	30	21	11	11	15	0.0435	0.0435
AA13	1	15	13	30	22	10	13	13	0.0435	0.0435
AA14	1	14	13	29	24	11	13	14	0.0435	0.0435
AA16	1	14	12	28	22	10	11	13	0.0435	0.0435
AA17	1	17	14	31	20	10	11	14	0.0435	0.0435
AA18	1	15	12	27	23	10	14	13	0.0435	0.0435
AA19	2	15	13	31	21	10	11	13	0.0870	0.0601
AA22	1	14	14	30	24	11	13	13	0.0435	0.0435
AA223	1	14	13	30	24	11	13	13	0.0435	0.0435
AA24	1	14	12	28	23	10	11	13	0.0435	0.0435

Table 3. Haplotype and molecular intra-population diversity measurement in African-Americans and Africans. Mitochondrial DNA estimates refer to nucleotide positions 16090 to 16365. Y-chromosome estimates refer to the five loci haplotypes (DYS389I, DYS3910, DYS391, DYS392 and DYS393).

		mtDNA				Y-Chromosome			
		n	HD (SE)	MNPD (SE)	MNPD/HD	*n*	HD (SE)	MNPD (SE)	MNPD/HD
African-Americans 1	AA1	50	0.984 (0.009)	7.291 (3.471)	7.410	23	0.976 (0.0201)	2.818 (1.543)	2.887
African-Americans 2	AA2	230	0.993 (0.002)	7667 (3.587)	7.721	106	0.978 (0.0045)	2.839 (1.508)	2.902
African-Americans 3	AA3	78	0.987 (0.006)	6.651 (3.173)	6.739	426	0.984 (0.0022)	3.034 (1.584)	3.082
Bakaka	BAK	50	0.983 (0.008)	9.457 (4.414)	9.620	49	0.796 (0.0270)	1.232 (0.796)	1.548
Bamileke	BAM	48	0.988 (0.007)	7.869 (3.725)	7.965	50	0.850)0.0375)	1.875 (1.091)	2.207
Bassa	BAS	46	0.991 (0.007)	9.194 (4.306)	9.277	49	0.843 (0.0355)	1.764 (1.040)	2.093
Beti	BET	48	0.965 (0.012)	8.395 (3.955)	8.699	36	0.916 (0.0232)	1.738 (1.035)	1.898
Cabinda	CAB	110	0.988 (0.003)	8.960 (4.159)	9.069	72	0.900 (0.0168)	1.993 (1.138)	2.215
Ewondo	EWO	53	0.983 (0.008)	9.666 (4.500)	9.833	39	0.804 (0.0449)	1.505 (0.926)	1.871
Fulbe	FUL	34	0.975 (0.016)	6.460 (3.134)	6.626	27	0.946 (0.0253)	2.644 (1.456)	2.795
Guinea Bissau	GUB	372	0.985 (0.002)	7.466 (3.496)	7.580	162	0.937 (0.0110)	2.183 (1.215)	2.330
Mozambique	MOZ	107	0.967 (0.007)	8.216 (3.839)	8.496	112	0.907 (0.0152)	1.975 (1.125)	2.177
Ngoumba	NGO	44	0.984 (0.007)	8.798 (4.137)	8.941	36	0.929 (0.209)	2.324 (1.301)	2.502

HD: haplotype diversity; SE: standard error; MNPD: mean number of pairwise differences.

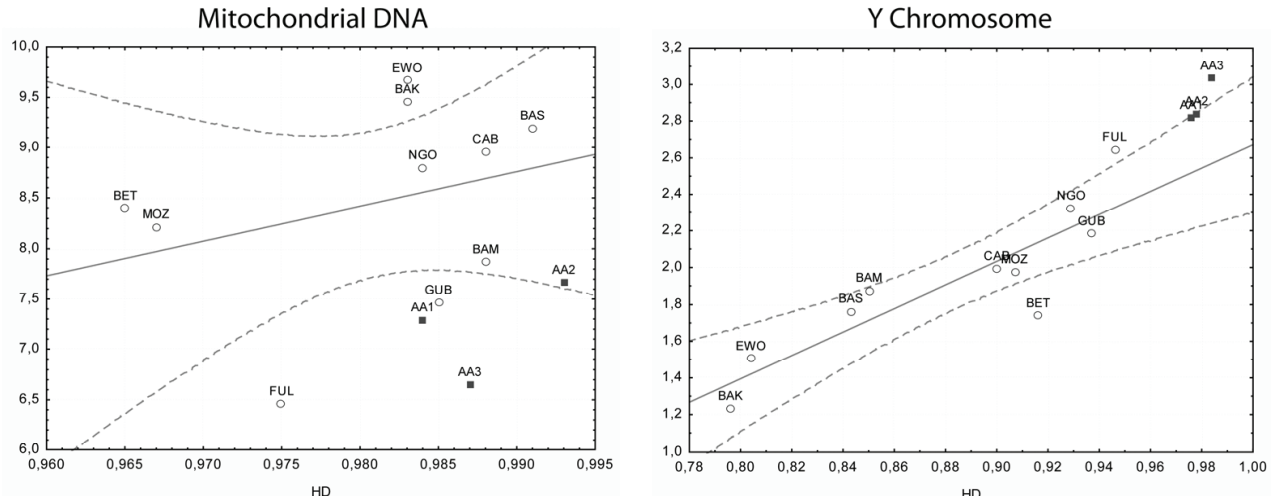

Figure 1. Scatterplots of MNPD and HD values for mtDNA and Y-chromosomes in the populations under study. Abbreviations are given in Table 3. Dotted lines indicate the 95.0% confidence interval for the regression line obtained for African populations.

netic systems generally produce values with extended confidence intervals that substantially overlap (*e.g.,* see estimates for Maryland, Texas and Virginia of table 5 in [9]).

A possible approach is represented by the search for possible effects of gene flow on intra- and inter-population genetic variation. Since the two genetic systems cannot be directly compared due to substantial difference in types and rate of mutations and demographic dynamics, we have compiled a paired mtDNA and Y-chromosomal population database including both European and African populations [25,28]. This makes it possible to compare the extent of mtDNA and Y-chromosome variation within a given population, by contrasting it with data from other populations included in the database.

An incoming European male-biased gene flow predicts a greater ratio of Y-chromosome to mtDNA diversity for African-Americans than for African populations. This expectation is met by our African-American sample. In fact, their mtDNA HD is equal or lower than in five African populations, whereas their Y-chromosomal value is higher (Table 3). The greater diversity of African-Americans is even more evident comparing MNPD, a measure which weighs molecular differences. In this case, the diversity between Europeans and Africans may have a greater impact than happens with HD, in which the extent of inter-haplotypic differentiation is not taken into account. In fact, there are eight African populations with a mtDNA MNPD higher than our African-Americans, and only one for the Y-chromosomal polymorphisms. It is noteworthy that the latter case involves the Fulbe, a population which has been shown to have undergone a male-biased gene flow due to African migrations [29]. Although no inference on asymmetric gene flow can be made for other African-American populations, it is remarkable that the ratio between molecular and haplotypic measures of intra-population diversity are the lowest for mtDNA and the highest for Y-chromosome polymorphisms (Figure 1).

A male-biased gene flow in African-Americans from Chicago, IL, USA is also supported by the analysis of genetic distances (Table 4). They show the lowest genetic distance (both molecular and haplotypic from Europeans for Y-chromosomal polymorphisms but not for the mtDNA ones. Furthermore, the ratio of mtDNA to Y-chromosomal genetic distances from Europeans is markedly higher (18.3-38.3% for molecular and haplotypic distances, respectively) than all African groups. The introgression of Y-chromosomes into other African-American populations is consistent with their comparatively low genetic distance from Europeans.

To sum up, our results coherently support the hypothesis that the African-American sample under study has undergone a European male-biased gene flow. On a more general note, our research showed that comparative analysis of intra- and inter-population variation for Y-chromosome and mtDNA polymorphisms in a broad dataset including African and

Table 4. Average genetic distances of Africans and African-Americans from Europeans.

		Molecular Distances			Haplotype Distances		
		mtDNA	Y Chrom.	mt/Y	mtDNA	Y Chrom.	mt/Y
African-Americans 1	AA1	0.219	0.281	0.779	0.027	0.024	0.151
Bakaka	BAK	0.240	0.624	0.385	0.028	0.120	0.232
Bamileke	BAM	0.267	0.456	0.585	0.025	0.095	0.270
Bassa	BAS	0.266	0.598	0.445	0.024	0.098	0.241
Beti	BET	0.312	0.477	0.653	0.036	0.058	0.624
Cabinda	CAB	0.250	0.435	0.575	0.025	0.066	0.374
Ewondo	EWO	0.278	0.528	0.527	0.028	0.114	0.244
Fulbe	FUL	0.198	0.399	0.496	0.032	0.038	0.832
Guinea Bissau	GUB	0.184	0.560	0.328	0.026	0.052	0.501
Mozambique	MOZ	0.254	0.415	0.612	0.035	0.067	0.528
Ngoumba	NGO	0.248	0.377	0.659	0.025	0.058	0.427

Europeans groups, may help detect admixture signatures, providing a useful complement to methods for admixture estimates.

REFERENCES

1. Curtain PD. The Atlantic Slave Trade: a Census. Madison: University of Wisconsin Press, 1969.
2. Conniff ML, Davis TJ.; Africans in the Americas: a History of the Black Diaspora. New York: St. Martin's Press, 1994.
3. Salas A, Richards M, Lareu MV, Scozzari R, Coppa A, Torroni A, *et al.* The African diaspora: mitochondrial DNA and the Atlantic slave trade. Am J Hum Genet. 2004; 74(3): 454-465.
4. Salas A, Carracedo A, Richards M, Macaulay V. Charting the ancestry of African-Americans. Am J Hum Genet. 2005; 77(4): 676-680.
5. Reed FA, Tishkoff SA. African human diversity, origins and migrations. Curr Opinion Genet Develop. 2006; 16(6): 597-605.
6. Reed TE. Caucasian genes in American Negroes. Science. 1969; 165(3895): 762-768.
7. Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, *et al.* Estimating African-American admixture proportions by use of population-specific alleles. Am J Hum Genet. 1998; 63(6): 1839-1851.
8. Parra EJ, Kittles RA, Argyropoulos G, Pfaff CL, Hiester K, Bonilla C, *et al.* Ancestral proportions and admixture dynamics in geographically defined African-Americans living in South Carolina. Am J Phys Anthropol. 2001; 114(1): 18-29.
9. Kayser M, Brauer S, Schädlich H, Prinz M, Batzer MA, Zimmerman PA, *et al.* Y Chromosome STR haplo-types and the genetic structure of US populations of African, European, and Hispanic ancestry. Genome Res. 2003; 13(4): 624-634.
10. Lind JM, Hutcheson-Dilks HB, Williams SM, Moore JH, Essex M, Ruiz-Pesini E, *et al.* Elevated male European and female African contributions to the genomes of African-American individuals. Hum Genet. 2007; 120 (5): 713-722.
11. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, *et al.* The genetic structure and history of Africans and African-Americans. Science. 2009; 324(5930): 1035-1044.
12. Destro-Bisol G, Maviglia R, Caglià A, Boschi I, Spedini G, Pascali V, *et al.* Estimating European admixture in African-Americans by using microsatellites and a micro-satellite haplotype (CD4/Alu). Hum Genet. 1999; 104(2): 149-157.
13. Kayser M, de Knijff P, Dieltjes P, Krawczak M, Nagy M, Zerjal T, *et al.* Applications of microsatellite-based Y-chromosome haplotyping. Electrophoresis. 1997; 18(9): 1602-1607.
14. Coia V, Boschi I, Trombetta F, Cavulli F, Montinaro F, Destro-Bisol G, *et al.* Evidence of high genetic variation among linguistically diverse populations on a micro-geographic scale: a case

study of the Italian Alps. J Hum Genet. 2012; 57(4): 254-260.

15. Torres JB, Doura MB, Keita SOY, Kittles RA. Y Chromosome lineages in men of West African descent. PLoS One. 2012; 7(1): e29687.
16. Kline MC, Vallone PM, Redman JW, Duewer DL, Calloway CD, Butler JM. Mitochondrial DNA typing screens with control region and coding region SNPs. J Forensic Sci. 2005; 50(2): 377-385.
17. Ely B, Wilson JL, Jackson F, Jackson BA. African-American mitochondrial DNAs often match mtDNAs found in multiple African ethnic groups. BMC Biol. 2006; 4: 34.
18. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010; 10(3): 564-567.
19. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 1999; 23(2): 147.
20. Seielstad MT, Minch E, Cavalli-Sforza LL. Genetic evidence for a higher female migration rate in humans. Nat Genet. 1998;20(3): 278-280.
21. Destro-Bisol G, Coia V, Boschi I, Verginelli F, Cagliá A, Pascali V, *et al.* The analysis of variation of mtDNA hypervariable region 1 suggests that Eastern and Western Pygmies diverged before the Bantu expansion. Am Nat 2004; 163(2);212-226.
22. United States Census Bureau 2000. url: http://www.census.gov/main/www/cen2000.html.
23. Hsieh CL, Sutton HE. Mitochondrial and nuclear variants in a US black population: origins of a hybrid population. Ann Hum Genet. 1992;56(Pt 2): 105-112.
24. Whatley W, Gillezeau R. The impact of the transatlantic slave trade on ethnic stratification in Africa. Am Econ Rev.2011; 101(3): 571-576.
25. Rosa A, Brehm A. African human mtDNA phylo-geography at-a-glance. J Anthropol Sci. 2011; 89: 25-58.
26. Montano V, Ferri G, Marcari V, Batini C, Anyaele O, Destro-Bisol G, *et al.* The Bantu expansion revisited: a new analysis of Y chromosome variation in Central Western Africa. Mol Ecol. 2011; 20(13): 2693-2708.
27. Mitchell P. Genetics and southern African prehistory: an archaeological view. J Anthropol Sci. 2010; 88: 73-92.
28. Francalacci P, Morelli L, Useli A, Sanna D. The history and geography of the Y chromosome SNPs in Europe: an update. J Anthropol Sci. 2010; 88: 207-214.
29. Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, *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(5): 1197-1214.