

Indian Journal of Human Genetics

Volume 13, Issue 3, September-December 2007

Contents

Editorial

- Bombay phenotype in Orissa: What could we make out of it
Kanjaksha Ghosh, K. Vasantha 79

Review Articles

- Coagulation disorders seen through the window of molecular biology
Kanjaksha Ghosh 81
- DNA profiling: Social, legal, or biological parentage
A. K. Sharma 88

Commentary

- Can parallel mutation and neutral genome selection explain Eastern African M1 consensus
HVS-I motifs in Indian M haplogroups
Clyde Winters 93

Original Articles

- PvuII polymorphism of estrogen receptor- α gene in breast cancer
D. Surekha, S. Vishnupriya, D. Nageswara Rao, K. Sailaja, D. Raghunadharao 97
- Possible risk factors for Down syndrome and sex chromosomal aneuploidy in Mysore, South India
Suttur S. Malini, Nallur B. Ramachandra 102
- Identification of a rare blood group, "Bombay (Oh) phenotype," in Bhuyan tribe of Northwestern Orissa, India
R. S. Balgir 109

Short Article

- Emergence of an unrelated highly aberrant clone in an AML patient at relapse four months after
peripheral blood stem cell transplantation
*Pratibha S. Amare Kadam, Hemani V. Jain, Purvish M. Parikh, Tapan K. Saikia, Sandhya Agarwal,
Indu Ambulkar* 114

Case Reports

- Rett syndrome molecular diagnosis and implications in genetic counseling
M. Noruzinia, M. T. Akbari, M. Ghofrani, H. Sheikha 119
- Mowat-Wilson syndrome in a Moroccan consanguineous family
*Ratbi Ilham, Elalaoui Chafai Siham, Dastot-Le Moal Florence, Goossens Michel, Giurgea Irina
Sefiani Abdelaziz* 122

Letter to the Editor

- Translocation t(2;14)(p13;q32) in a case of Ph+ acute lymphoblastic leukemia
Lily Kerketta, Babu Rao Vundinti, Kanjaksha Ghosh 125

Author Index - 2007

127

Title Index - 2007

129

List of Reviewers for 2007

130

The copies of the journal to members of the association are sent by ordinary post. The editorial board, association or publisher will not be responsible for non-receipt of copies. If any of the members wish to receive the copies by registered post or courier, kindly contact the journal's / publisher's office. If a copy returns due to incomplete, incorrect or changed address of a member on two consecutive occasions, the names of such members will be deleted from the mailing list of the journal. Providing complete, correct and up-to-date address is the responsibility of the members. Copies are sent to subscribers and members directly from the publisher's address; it is illegal to acquire copies from any other source. If a copy is received for personal use as a member of the association/society, one cannot resale or give-away the copy for commercial or library use.

Can parallel mutation and neutral genome selection explain Eastern African M1 consensus HVS-I motifs in Indian M haplogroups

Clyde Winters

COE, G-323, Governors State University, University Park, Illinois 60466, USA

Controversy surrounds the relationship between the Eastern African M1 haplogroup and Indian M haplogroups. Some researchers see a “relationship” between the Eastern African M1 and Indian M haplogroups,^[1-3] while other researchers maintain that the Eastern African HVS-I signature motif 16,129, 16,189, 16,249, and 16,311,^[4] is not found in Indian M haplogroups, and this motif when it is found results from “parallel mutation.”^[5-6]

Researchers have noted that the nucleotides shared by East African M1 and Indian M haplogroups include HG M4 at 16311, HG M5 at 16,129, and HG M34 at 16249;^[3] other researchers have identified a number of transitions for Indian M1 at 16,311, 16,129, and 16,189, that correspond to Ethiopian M1.^[1] Other Indian nodes that agree with East African M1, observed by Kivisild *et al.*, include: HG M5a 16,311; HG M5 16,189; and HG M2a 16,189.^[1]

Metspalu *et al.* noted that “another Indian-specific M clade, supported by HVS-I variation as well as coding region markers, is M6. Haplogroups M3, M4, and M5 have been discriminated preliminarily by their characteristic HVS-I mutations, but since their defining positions, 16126, 16311, and 16129, respectively, are phylogenetically unstable.”^[2]

Semino *et al.* made it clear that the Eastern African group named M1 is characterized within M by a consensus HVS-I motif defined by four transitions at nt 16,129, 16,189, 16,249, and 16,311.^[5] Their research indicated that the HVS-I signature motif of M1 was not found in the Indian samples, whereas it characterizes most of the M types sporadically observed in the Mediterranean area and in 7% of Nile Valley sequences.

Sun *et al.*, after complete sequencing of the ancestral motifs of the Indian M haplogroups, found none of the

variations that characterize M1, i.e., 6446, 6680, 12403, and 14110;^[6] the common mutations in the control region between Eastern African M1 and Indian M haplogroups reported by Roychoudhury *et al.*^[3] and Metspalu *et al.*^[2] reflect random parallel mutations.

Given the fact that except for Semino *et al.*,^[5] all of these authors found Eastern African HVS-I motifs in Indian M haplogroups, we have to ask the question: is this the result of parallel mutation as claimed by Sun *et al.*^[6] or evidence of a close relationship between these M haplogroups?

The M haplogroups are among the older African mtDNA category. It is among this group that selection has played a prominent role in the evolutionary process.

Parallel evolution is an adaptive behavior in which beneficial genes in one haplogroup appear in a different haplogroup, because they live in identical environments and belong to replicate populations.^[7] Generally, parallel mutations are low-probability mutations.

In human mtDNA, it is known that transitional mutations (between A and G and between T and C) occur much more frequently (10-20 times more frequently) than transversional mutations (the remaining combinations of A, T, G, and C) and that the extent of variation of the mutation rate among nucleotide sites is high. These characteristics of the nucleotide mutations in human mtDNA make parallel mutations quite likely at some sites.

Parallel evolution is usually explained by the mutational landscape model (MLM), allelic fitness, and extreme value theory (EVT), which is a branch of statistics used to explain extreme or rare events. In traditional data analysis, extremes are either ignored or considered as outliers.

Contemporary genetics theory is focused on adaptation. Adaptation theory describes the standard rate of mutations of favorable mutations exponentially as a result of beneficial fitness effects

Beneficial mutations are the foundation of evolution by natural selection.^[9] Beneficial mutations are “exceedingly rare” but they present large fitness effects that result from adaptation to new environments.

Generally, adaptation theory suggests that allelic fitness is well behaved because the distribution of beneficial mutations illustrate fitness effects distributed exponentially across various spaces with identical beneficial effects.^[10] The general theory is used to explain the direction and rate of adaptation.

Beneficial mutations reside in the domain of EVT.^[11-12] As a result, the allelic fitness distribution is “well behaved”^[11] and exponentially distributed.^[11-12] Gillespie^[12] and Orr^[8,10,13] believe these are reasonable assumptions for populations whose optimized wild-type has declined as a result of an environmental shift.

The MLM makes it clear that beneficial mutations are distributed exponentially. The MLM predicts that the effects of beneficial mutations are exponentially distributed even if the mutations are random.^[14]

Theoretically, alleles that improve fitness increase in frequency because they are beneficial to the organism. Beneficial mutations which leads to parallel mutations are rare.^[10] In relation to the beneficial mutation theory, Orr established two properties: 1) the distribution of beneficial fitness effects at a gene is exponential and 2) the distribution of beneficial effects at a gene has the same mean, regardless of the fitness of the present wild-type allele.^[10]

We would assume that if parallel mutation accounted for the distribution of African M1 HVS-I, we would see $m + 1$ alleles and a relatively high fitness for these transitions because the replicate populations live in the same environment and would present similar transitions in the M haplogroup because of the number of beneficial effects of African nt 16,111, 16,129,16,249, and 16,311. It is assumed that beneficial effects of the favorable mutant within the haplogroup will account for the absolute fitness of the shared allele, which is advantageous for identical environments pursuant to extreme value theory.

Table 1: African M1 transitions in Indian haplogroups

Haplogroup	16311	16129	16111	16249	16189
M2	X	-	-	-	X
M2a	X	-	-	-	-
M4	X	X	-	X	-
M4'30	-	X	-	X	-
M35a	X	-	X	-	X
M37	-	-	-	-	X
M38	-	-	X	-	-
M40	-	X	-	-	-
M5	-	-	-	-	X

This is not the case. The distribution of African M1 nucleotides in Indian M haplogroups is 9/27 [Table 1]. If these nucleotides were the result of parallel mutations they would be more widespread across the subclusters of the Indian M haplogroups, at least $\frac{1}{2}$, given the high probability of parallel mutations appearing in human mtDNA.

Sun *et al.* makes it clear that consensus Eastern African M1 nucleotides appear frequently in the phylogenetic tree of the Indian mtDNA.^[6] The highly recurrent mutations, according to Sun *et al.*, were 16,129, 16,189, and 16,311.^[6]

Sun *et al.* report that the Eastern African M1 nt in the Indian M haplogroups include M2 16,3111 and 16,189; M2a 16,311; M5 16,129; M4 16,129, 16,311, 16,249; M5 16,129; M'30 16,249, 16,129; M34 16,249; M35a 16,311, 16,189; M37 16,189; and M40 16,129.^[6] The researchers maintain that even though Indian M haplogroups bear variant M1 nucleotides in the control region, the reconstructed ancestral motif of all the Indian M haplogroups were devoid of M1 variations 6446,6680,12403 and 14110.^[6]

Selection has influenced the evolution of the human genome.^[7] It is often assumed that selection plays a limited role in the mtDNA control region. As a result, neutral evolution fails to play a significant role in mtDNA evolution.^[7]

There is a selective constraint on mutation frequencies of an mtDNA site.^[15-16] Some of the East African transitions, namely, 16129, 16189, and 163111 T → C polymorphism are the most rapidly occurring nucleotide substitutions in the human mitochondrial genome.^[16] These transitions are often referred too as “hotspots.” These hot spots of mutational activity suggest that positive selection influences mutation rates and not neutral selection which,

theoretically, would manifest parallel mutations.

The average mutation rate for “hotspots” like 16129 and 16189 is 1.25×10^{-5} substitutions per generation per substitution.^[16] The mutation rate for fast sites would be 125% sequence divergence per Myr for HVI. The average estimate for the distribution of substitution of fast sites is probably around 40.^[16]

Research makes it clear that these fast nucleotide sites demonstrate a hypermutation process.^[16] This mutation process, although rapid, is never neutral and illustrates differing rates of selection on the distribution of each genotype.

The suggestion by Sun *et al.* that parallel mutations account for the Eastern African M1 nucleotides in Indian haplogroups, lacks congruence to the theoretical models associated with adaptation and parallel evolution (EVT and MLM).^[6] Orr makes it clear in the phenotype model of adaptation or DNA sequence model of adaptation that the genes that cause adaptation should have approximately exponentially distributed effects, that is, involve many genes that have small effects and a few genes that have large effects.^[10] In the discussion of Sun *et al.*, sequencing of the Indian M haplogroups show small fitness effects as illustrated above.^[6] The failure of Sun *et al.* to meet Orr's^[8,13] fitness effects for beneficial genes rule against the presence of M1 genes in Indian haplogroups as a result of parallel evolution.

Clark recognized early that when DNA sequences are well described it is extremely unlikely that multiple parallel mutations occur.^[17] This feature was evident in the East African M1 nucleotides found in the Indian M haplogroups reported by Sun *et al.*; the results make it clear that there were few M1 mutations in the Indian haplogroups.^[6]

Sun *et al.* sequenced around 27 Indian M haplogroups. Eastern African M1 nt 16,311, 16,129, 16,111, 16,249, and 16,189 were found in only 8 of the Indian M haplogroups: M2, M2a, M4, M4'30, M35a, M37, M38, and M40.

The distribution of M1 HVS-I in the Indian haplogroups vary. We only see transition in M2 and M4'30, while M37, M38, and M40 had only one M1 nt.^[6] In relation to M4 and M35a, Sun *et al.* report 3 shared transitions.^[6] The limited distribution of M1 transitions in Indian M haplogroups failed to evidence the exponential distribution that is expected when parallel mutation takes place.

If the presence of African M1 transitions in the Indian haplogroups were the result of parallel evolution, the most plausible explanation would be that the polymorphism maintained in both lineages must date to the time of the most recent common ancestor (MRCA). But we must reject this hypothesis because Sun *et al.* made it clear that the Indian haplogroups lack the ancestral motifs characterized by M1.

In summary, the presence of Eastern African HVS-I distribution of M1 mutations across the Indian HG subclusters is not exponential. Secondly, distribution of the mean for the Eastern African M1 motifs across and within the Indian M haplogroups is not the same.

The lack of regularity in distribution of the Eastern African M1 transitions among the Indian M haplogroups lacks congruency. Since Indian M haplogroups are relatively of the same age and population they would demonstrate regular and similar distributions of neutral substitutions across the haplogroups due to the clock-like and context-dependent origination of neutral substitutions. The distribution of the African M1 mutations are not distributed across the entire spectrum of Indian M haplogroups as would be expected if they were the result of neutral origination. Their existence within the Indian M subclusters must be the result of reasons other than neutral substitution theory and parallel mutations.

This illustrates that the existence of M1 HVS-I in Indian M haplogroups fail to meet the properties associated with beneficial allele mutations that researchers believe account for parallel mutations.

Conclusion

A cursory examination of the data provided by Sun *et al.* evidences that the African M1 signature mutations are not exponential among Indian M HGs, nor do they have the same mean among and within the Indian M haplogroups that possess the African M1 transitions. The absence of high fitness of African M1 mutations in Indian M haplogroups does not fit the theoretical assumptions associated with the origination of parallel mutations.

The theoretical and empirical assumptions discussed above make it clear that the presence of African M1 in

Indian M haplogroups cannot be the result of parallel mutation. Parallel mutations are explained by EVT and MLM. These theoretical models imply that mutation effects result from beneficial alleles that demonstrate exponential distribution within haplogroups due to the fitness effects of the beneficial mutations.

The distribution of M1 transitions in only 30% of the Indian M subclusters suggest that they may evidence a recent migration of the Dravidian-speaking people to India. The data provided by Sun *et al.* fails to demonstrate a robust distribution of African M1 across the Indian subclusters.^[6]

Failure to meet the minimum expected distribution pattern of mutations predicted by parallel mutants, and the present limited distribution of African M1 across and within Indian haplogroups, is probably explained by a recent African origin of the Dravidian-speaking people, who probably migrated to Indian during the Neolithic period. This explanation has some merit, given the anthropological, linguistic, and osteological evidence of an African origin of the Dravidian speakers.^[18-19]

References

1. Kivisild, Toomas, Kaldman K, Metspalu M, Juriparik, Papiha S. *In: Genomic diversity*. Papiha Deka R, editor. SS Kluwer/ Plenum Publishers: 1999. p. 135-52. Available from: <http://evolutsioon.ut.ee/publications/Kivisild1999b.pdf>.
2. Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, *et al.* Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet* 2004;5:26.
3. Roychoudhury S, Roy S, Basu A, Banerjee R, Vishwanathan H, Usha Rani MV, *et al.* Genomic structures and population histories of linguistically distinct tribal groups of India. *Hum Genet* 2001;109:339-50.
4. Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nat Genet* 1999;23:437-41.
5. Semino O, Santachiara-Benerecetti A, Falaschi F, Cavalli-Sforza LL, Underhill PA. Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *Am J Hum Genet* 2002;70:265-8.
6. Sun C, Kong QP, Palanichamy MG, Agrawal S, Bandelt HJ, Yao YG, *et al.* The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes. *Mol Biol Evol* 2006;23:683-90.
7. Elson JL, Turnbull DM, Howell N. Comparative genomics and the evolution of human mitochondrial DNA: Assessing the effects of selection. *Am Soc Hum Genet* 2004;74:229-38.
8. Orr HA. The probability of parallel evolution. *Evolution Int J Org Evolution* 2005;59:216-20.
9. Cowperthwaite MC, Bull JJ, Meyers LA. Distributions of beneficial effects in RNA. *Genetics* 2005;170:1449-57.
10. Orr HA. The distribution of fitness effects among beneficial mutations. *Genetics* 2003;163:1519-26.
11. Weinreich DM. The rank ordering of genotypic fitness values predates genetic constraints on natural selection on landscapes lacking sign epistasis. *Genetics* 2005;171:1397-405.
12. Gillespie EJ. A simple stochastic gene substitution model. *Theor Popul Biol* 1983;23:202-15.
13. Orr HA. The population genetics of adaptation on correlated fitness landscapes: The block model. *Evolution Int J Org Evolution* 2006;60:113-24.
14. Betancourt A, Bollback JP. Fitness effects of beneficial mutations: Thew mutational landscape model in experimental evolution. *Curr Opin Genet Dev* 2006;16:618-23.
15. Howell N, Smejkal CB. Persistent heteroplasmy of a mutation in the human mtDNA control region: Hypermultiplication as an apparent consequence of simple-repeat expansion/contraction. *Am J Hum Genet* 2000;66:1589-98.
16. Excoffier L, Yang Z. Substitution rate variation among sites in mitochondrial hypervariable region 1 of Humans and Chimpanzees. *Mol Biol Evol* 1999;16:1357-68.
17. Kwang DG, Green P. Bayesian Markov chain Monte Carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution. *Proc Nat Acad Sci USA* 2004;101:13994-4001.
18. Sergeant, Bernard. *Genèse de L'Inde*. Payot: Paris; 1992.
19. Winters C. Did the Dravidian speakers originate in Africa? *Bioessays* 2007;29:497-8.

Source of Support: Nil, **Conflict of Interest:** None declared.