ORIGINAL ARTICLE

Microcephaly: Tracing the Evolutionary Lineage of ASPM gene

Mushtaq Hussain[1], Muhammad Asif Qureshi[2], Muhammad Rafiq khanani[1], Abdul Wajid[3], Shaheen Shoaib Sharafat[1] and Talat Mirza[1]

ABSTRACT

Objective: Microcephaly, in the form of congenital autosomal recessive disorder (MCPH), is characterized by the reduced occipital frontal head circumference >3 standard deviation of otherwise normal population of matching age and sex. The disease is primarily associated with mild to severe mental retardation. Earlier studies have unravelled that among Pakistani population, mutations in *ASPM* gene is strongly associated in MCPH. In the present study, we have explored the ancestral root of this disease and the process involved in its evolution using tools of bioinformatics. **Experimental Methods:** cDNA gene and protein sequences of *ASPM* gene were retrieved from NCBI database and subjected to the non-redundant BLAST. Consensus phylogenetic tree was developed after multiple sequence alignment and bootstrapping of the protein sequences of *ASPM* gene from different mammals using Neighbour Joining method, selecting non mammals as an out group. Comparisons of the gene synteny and exon and intron patterns of *ASPM* gene were also undertaken to investigate chromosomal changes during the course of human evolution. Different statistical evolutionary models namely, Codon Based Z test and Maximum Composite Likelihood Estimate were used in order to estimate the nature of nucleotide substitution and the type of selection pressure the gene has undergone. **Results:** Phylogenetic tree based on *ASPM* gene clearly segregated all non mammalian members as an out group. Mammalian in group holds the established evolutionary lineage, based on morpho-genetic attributes of mammalian evolution, segregating monotremes at the beginning followed by the members of rodentia and finally radiation of the primates including humans. Orientation of the *ASPM* gene remains conserved between human and chimpanzee, however, it was found reversed along with two flanking genes, a zinc finger binding domain 41 and coagulation factor XIII, which suggest relatively recent event of gene inversion. Some earlier and, in comparison, more intricate chromosomal changes have also been detected among the lower order of mammals. Aligning *ASPM* gene exons with the primates and lower order mammals indicates transitional bias of mutation over transversion (R value= 1.563). Holistically, codon based Z test revealed positive selection pressure on of *ASPM* gene from rodentia to primates. **Conclusion:** Briefly, the studies highlights the evolutionary events of *ASPM* gene in mammals especially primates including humans. Further studies in connection to correlating the cranial cavity size and ancestral gene sequences and in depth sequence comparison would be more insightful in this regard and studies in this connection are ongoing and will be reported shortly.

Key words: Microcephaly, Human Evolution, ASPM gene.

INTRODUCTION

Human brain, in its normal physiological capacity, is arguably the most fascinating milestone achieved during the course of human evolution. The notion is self evident if one compare the sheer intellectual and communication fire power and in many instances dexterity of the humans with other closely and distantly related organisms. Increased size of the brain in comparison to body (encephalization)[1] is the most important attribute in this regard and any anomaly that leads to the reduction of the brain size is responsible for the partial and/or complete loss of such distinctive characteristics. One such illness is called as microcephaly.

Microcephaly (MC) is defined as a significant reduction in head circumference (>3 SD below the mean) for an individual's age and sex.[2] It is a multifactorial diseases which can either present itself as a single disease or as part of a syndrome. Dianne Abuelo has previously reviewed various symptoms and disease associations of MC in detail.[3] Several environmental and genetic factors have been found responsible for the microcephaly. Common environmental factors that cause MC include antenatal exposure to teratogens including alcohol, hydantoin and radiation amongst others. Genetically, the disease may present as an autosomal dominant, autosomal recessive or X-linked traits.[3] On the basis of its onset, the disease is broadly classified into primary MC, which arises due to the impairment of the brain development during gestation period, and secondary MC, which develop during the postnatal period.[4]

1. Department of Molecular Pathology, Dow University of Health Sciences, Ojha Campus, Karachi, Pakistan.
2. Department of Biochemistry, Dow University of Health Sciences, Ojha Campus, Karachi, Pakistan.
3. HEJ Research Institute of Chemistry, University of Karachi, Karachi, Pakistan.
Correspondence: Dr. Muhammad Rafiq Khanani, Department of Molecular Pathology, Dow University of Health Sciences, Suparco Road, Ojha Campus, Karachi, Pakistan.
E-mail: r.khanani@duhs.edu.pk

Autosomal recessive primary microcephaly (MCPH) is a subtype of primary MC in which reduction of head circumference is associated with mild to severe mental retardation and/or other neurological symptoms such as seizures cognitive deficits etc.[5] Alarmingly, compared to 1 in a million in British population, incidence of MCPH in Pakistan is 1 in 10,000.[6] More recently various genes have been found associated with the occurrence of MCPH. These genetic loci include *MCPH1, MCPH3, MCPH5/ASPM* and *MCPH6*. [7,8,9,10,11] Amongst all the mutations identified in different microcephaly associated genes, *ASPM* mutations have been reported to be the most common cause of MCPH in Pakistan.[12,13]

ASPM (Abnormal spindle like microcephaly-associated protein) gene encodes for a protein which expressed in the cerebral cortical ventricular and proliferative zones of medial and lateral ganglionic eminences during neurogenesis. Postnatally, *ASPM* also expresses in the regions of constant neurogenesis. The gene regulates the normal mitotic spindle function in the embryonic neuroblasts and therefore it is required for proliferation of neuronal cells to attain the required brain size. Mutations in the genes generally results in the production of truncated non functional protein which impairs the normal brain development and causing micocephalic conditions.[14,15]

Several studies have been conducted which relates the sequential transitions of *ASPM* gene in different mammals with their brain size.[6,14,15,16] However, the nature of selection pressure and transformation of genomic synteny underlying the very process of *ASPM* evolution is still debatable or largely unaddressed. In the present study we have explored the ancestral lineage of *ASPM* genes among different representative mammals including human. The study has been further extended to investigate the changes which occurred in the *ASPM* gene order of the compared organism while going through evolution. Sequences were also analysed using suitable mathematical and/or statistical models to unravel the nature of selection pressure in this connection. It is our belief that the present study will provide more insights to the evolutionary history of microcephaly with reference to *ASPM* gene. To the best of our cognizance the study is the first report to illustrate the genomic architectural changes of *ASPM* during its evolution.

METHODS

Sequence Analysis and Evolutionary Tree Construction

Protein and cDNA (CCDS 1389.1) sequences of Human *ASPM* gene were retrieved from dBest database of the NCBI (National Centre for Biotechnology Information;).[17] The sequences were subjected to non-redundant BLAST (Basic Local Alignment Search Tool) in FASTA format.[18,19] The homologues of the *ASPM* were retrieved and multiple sequence alignment were generated using default parameters (with some manual adjustment) of program Clustal X and MEGA4 (Molecular Evolutionary Genetic Analysis version.4).[20,21] The consensus boot strapped (500 replicas) phylogenetic tree was constructed from ASPM protein sequences using Neighbour Joining method.

Genomic Organization

Information regarding *ASPM* and its flanking genes organization of selected animals were taken from NCBI gene database. For the sake of comprehensibility, the schematic representation were developed using Microsoft power point. Additionally, the ASPM gene exons and intronic regions were also compared.

Evolutionary Rate Selection Pressure Test

Both gene and protein sequence alignments were subjected appropriately to various statistical tests in order to delve out the nature of selection pressure directing the adaptive evolutionary course of the gene. To explore the nature of selection pressure, the gene sequence alignment were subjected to Null hypothesis of all three selections, negative selection(HA:dN>dS), positive selection (HA:dN<dS) and neutrality (HA:dN=dS), where dN is the number of non synonymous mutation which dS is the number of synonymous mutations. The hypothesis was tested using Codon-Based Z-Test along with a mathematical model of Nei-Gojobori (p-distance).[22, 23]

RESULTS AND DISCUSSION

The Neighbour joining consensus tree based on protein sequences of *ASPM* gene, shown in Fig.1, is supported by high boot strap values indicating the identical branch and clades segregations during boot-strapping stages. The tree clearly excluded all the out groups (non mammalian version of ASPM protein). Importantly, the tree holds the established evolutionary lineage of mammalian evolution, showing radiation of monotremes (*Ornithorhyncus*) before the rest of the advanced mammals. This is followed by the separation of rodentia (*Mus* and *Rattus*) clade. Further to it is the separation of taxa belonging to artiodactyla (*Bos*, *Hippopotamus* and *Ovis*) and cetacean (*Tursiops*) in a single clade. It is worth mentioning that although Hippopotamus is taxonomically the member of order artiodactyla but several paleontological, physiological and molecular studies have unravelled that they are indeed the closest relatives of cetaceans like Dolphins and Whales.[24,25] Moreover, both artiodactyla and cetaceans shares a common ancestor (cetaartiodactyla) and split between these two order occurred around 60-55 million years ago (mya). Taken these facts into account, the clustering of Hippopotamus with Dolphin in the tree based on ASPM protein sequences provides yet another piece of evidence regarding the evolutionary relatedness of *Hippopotamus* and cetaceans. Following that is the branch leading to *Equus* (a member of Perisodactyla) which also radiated from the common ancestors of cetaartiodactyla, namely ungulatomorpha in late cretaceous or early paleocene period (70-64 mya) however, they only started diverging around 10-15 million years later than the earlier split.[26,27] Similarly, the tree also agrees with the notion of relatively recent diversification of carnivores than perisodactyla. Indeed, the split between the common ancestors of cetaartiodactyla and ungulatomorpha and members of order carnivora (*Felis, Canis* and *Ailuropoda*)

occurred in the early paleocene period (80 mya) but their diversification had started around 55 mya.[26,27] As anticipated, the primates formed the separate clades after all the non primate mammals, which is in agreement with the established time scale of separation of primates and other mammals from their common ancestors. Albeit the split between the common ancestors of primates (including human) and all the mentioned mammalian orders occurred during late cretaceous period or perhaps much earlier (>65mya) but the primates only diversified after Eocene period (<34mya) resulting the late separation of the primate clades than rest of earlier mammals. Akin to the well accepted notion of primate evolution, which is based on paleontological and molecular data (not ASPM), the ASPM phylogeny also validates the segregation of new world monkeys (*Callithrix, Aotus and Saimiri*) before the old world monkeys (*Colobus, Chorocebus and Macaca*). According to geological time scale, the split between the common ancestors of old and new world monkeys occurred at about 45 mya (Eocene period). The later split in the common ancestors of old world monkeys and rest of relatively advanced primates (*Hylobates*; Gibbon) happened at the late Oligocene period (around 25 mya). The further separation between *Pongo* (Orang utan) and Gibbon occurred in Miocene period (18 mya). *Gorilla* and other African apes including Human share their common ancestors with Orangutan around 14 mya. Common ancestor of both Chimpanzee (*Pan)* and Human (*Homo*) splits from *Gorilla* at around 7 mya. Finally the chimpanzees and humans have taken different evolutionary course from their common ancestors around 6 mya (very late Miocene period).[26,27,28] The greater congruencies between the established evolutionary lineage of the mammals and the constructed phylogenetic tree not only strengthen the standpoint of evolution by natural selection but also verify the fidelity of tree and the algorithm on which it is based. Moreover, higher bootstrap values and clear out grouping of non mammals from the tree strongly implicates towards the validity of the phylogenetic calculation.

In human, *ASPM* gene is located on chromosome no 1, the largest of all human chromosome bearing 8% of our total genetic information.[29] Genetically, *ASPM* gene in all mammals are respectively flanked by Zn finger and BTB domain 41 (ZBTB41) and coagulation factor XIII genes (F13B) in 5' and 3' directions. However, the orientation of the gene changes during the course of evolutionary history. Both in Humans (*Homo*) and Chimpanzees (*Pan*) the orientation of all three genes remains same, however, it is found reversed in a representative of the old world monkeys (*Macaca*). The similarity of ASPM order between human and chimpanzee could be inferred in terms of the interbreeding era (roughtly 3 million years) since the divergence from the common ancestor (6 mya)[29]. As the old world monkeys and advanced apes have split from their common ancestor in the Oligocene period, thus the event of this genes inversion must be not older than 25 million years from the present time. It is interesting to note here that the orientation of all three genes (ASPM, ZBTB41 and F13B are same in old world Macaca and wolves (*Canis*; Dog family) and mice (*Mus*). However, the orientation of ZBTB41 and ASPM are found reversed in Cow (Bos). This implies that the inverted orientation in this lineage could happened during the divergence of artiodactyla (<55mya) (if only found in cow and/or few members of the clade). Alternatively, it may be ancient duplication (>65mya) (Fig. 2).[27,28,29] In addition to the gene order changes, subtle changes have also been found within the *ASPM* gene. Fig. 3 clearly represents the differences among the distribution, numbers and sizes of exons and introns in the *ASPM* gene among the compared mammals including human. Architecturally, the diversity of the eukaryotic genome is primarily mediated by the sequence divergence, insertion, deletion, duplication and recombination[30], the mentioned changes in the synteny and exon/intron pattern in *ASPM* genes could be the function of any/some or all of such processes. Additionally, expression of the gene is known to be affected directly or indirectly by gene synteny and/or orientation and pattern of splicing.[31,32] Taken together, it is possible that the order and perhaps more importantly the intragenetic architecture may influence the ASPM protein expression positively in the humans resulting in the extraordinary encephalization and increment to the intellectuality and cognitive abilities. Supportive to this inference is the study conducted by Caceres and co-workers, which has shown the exceptional transcriptional edge in human as compared to other primates.[33] Additionally, the sequential changes in *ASPM* gene after the split from the common ancestors of chimps could not be discounted as it is known that human encephalization had occurred several times in its evolutionary history, starting around 2 mya and reaching its current status approximately 500,000 years ago.[14,34]

Codon based Z test clearly shows the expected positive selection of the protein as one move from rodentia to primates. Importantly, most of the significant values lies in the clade of old world monkeys (clade where humans reside). The values imply that non synonymous mutations (change in amino acid) in the ASPM protein are tolerated and in certain cases even favoured in the nature. However, this case of positive selection could not be implemented to the mutations of ASPM gene causing microcephaly, as mutations responsible for microcephaly are protein truncation mutations thus producing an incomplete protein.[35] Our finding is in line with some earlier investigations regarding ASPM evolutionary pressure[6,14,16], however, the studies have mainly confined to the selection pressure present in only primate lineage. Maximum Composite Likelihood Estimate has shown that the transition over transversion, however, with few exceptions this trend considered to be universal among metazoans.[36]

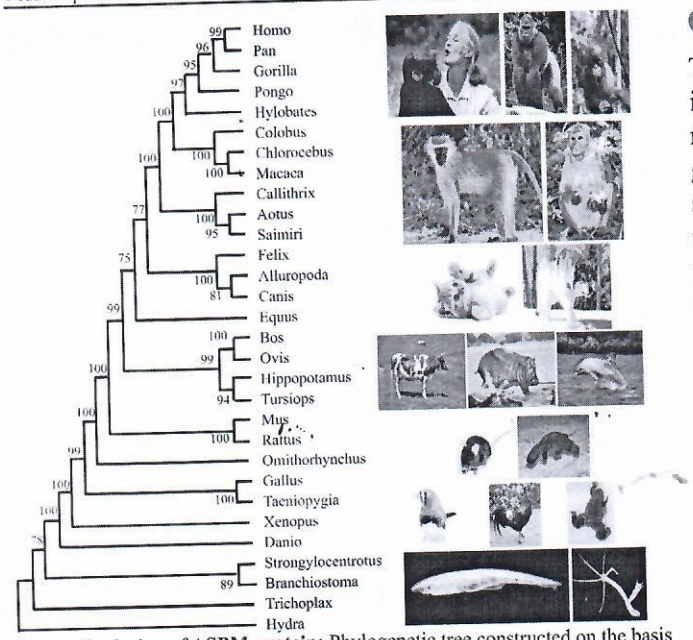

Fig. 1. Evolution of ASPM protein: Phylogenetic tree constructed on the basis of ASPM protein sequence alignment using neighbour joining method. Clades of different taxa are differentially coloured. Orange branches are of out groups (non mammalian ASPM), brown branch represents monotremes, Purple clade include members of cetaartiodactyla, blue for perisodactyla, orange for members of carnivore and green clade constituted on the members of primates including human. Boot strap values are shown at each branch point. (*For details please see text*).

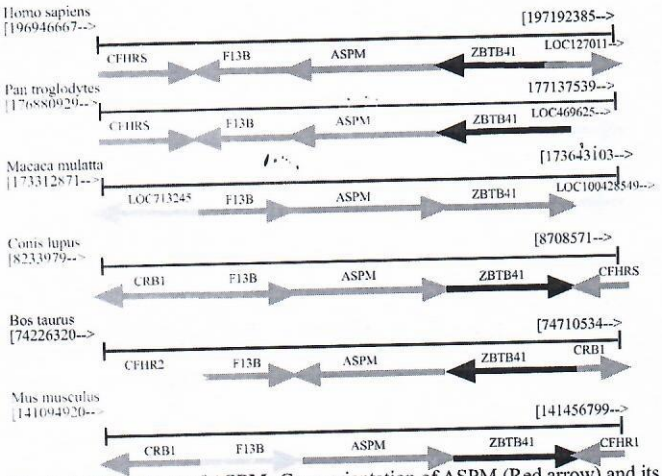

Fig. 2. Gene synteny of ASPM: Gene orientation of ASPM (Red arrow) and its flanking genes Zn finger and BTB domain 41 (ZBTB41; black arrow) and coagulation factor XIII genes (F13B; green arrow) is shown, head of arrows indicates the direction of transcription. Note the change in the orientations among different mammals which may affect the rate of expression of the ASPM, details are furnished in the text.

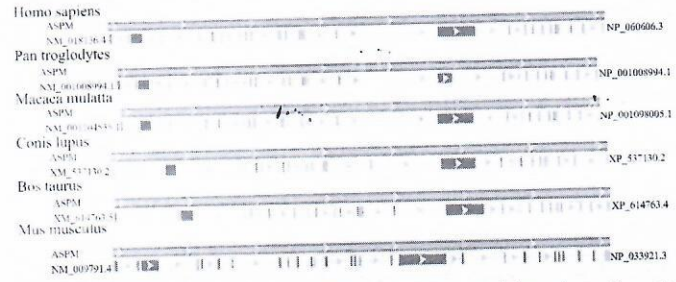

Fig. 3. Distribution of exons and introns of ASPM gene: The pattern of exons and introns in the ASPM varies considerably in terms of size, number and location within the gene among different mammals. Green bars, rectangles and squares indicated underneath the gene are the exons while the lines in between them are introns.

CONCLUSION

Taken together, we believe that substantially high intelligence in the primates, especially in humans, is the result of gradual modifications within and outside the brain development genes like *ASPM*. Such modifications are rigorously scrutinized by the forces of natural selection and favoured if found benevolent for the survival of the organism individually and in many cases for the entire species. We realize that the empirical data in connection to *ASPM* for such notion is lacking. Additionally, we are broadly unaware of the structure-function aspects of ASPM protein, once known, it may clarify the significance of minute sequential changes among the ASPM protein of different mammals in terms of their structure and consequently their functions. However, intellectual superiority of humans over other living species could not be fully answered by single gene or set of genes. Evolutionary Medicine is a new emerging field of biology, it is plausible to expect that soon in combination with the system biology and transgenic animal models we would be able to answer this great mystery of nature.

REFERENCES

1. Harvey PH, Krebs JR. Comparing brains. Science 1990; 249: 140–6.
2. Ross JJ, Frias JL. Congenital malformations of the brain and skull. Elsveir Holland Biomedical press: 1977; 30: 507-24.
3. Abuelo D. Microcephaly Syndromes. Seminars in Pediatric Neurology 2007; 14: 118-27.
4. Woods CG. Human Microcephaly. Curr opin Neurobiol 2004; 14: 112-7.
5. Roberts E, Hampshire DJ, Pattison L, Springell K, Jafri H, Corry P, et al. Autosomal recessive primary microcephaly: an analysis of locus heterogeneity and phenotypic variation. J Med Genet 2002; 39:718-21.
6. Woods CG, Bond J, Enard W. Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings. Am J Hum Genet 2005; 76:717-28.
7. Jackson AP, McHale DP, Campbell DA, Jafri H, Rashid Y, Mannan J, et al. Primary autosomal recessive Microcephaly (MCPH1) maps to chromosome 8p22-pter. Am J Hum Genet: 1998; 63; 541-6.
8. Jamieson CR, Govaerts C, Abramowicz MJ. Primary autosomal recessive Microcephaly: homozygosity mapping of MCPH4 to chromosome 15. Am J Hum Genet 1999; 65: 1465-9.
9. Roberts E, Jackson AP, Carradice AC,Deeble VJ, Mannan J, Rashid Y, et al. The second locus for autosomal recessive Microcephaly (MCPH2) maps to chromosome

19q13.1-13.2. Eur J Hum Genet 1999; 7: 815-20.

10. Moynihan L,Jackson AP, Roberts E, Karbani G, Lewis I, Corry P, et al. A third novel locus for primary autosomal recessive Microcephaly maps to chromosome 9q34. Am J Hum Genet. 2000; 66: 724-7.

11. Pattison L, Crow YJ, Deeble VJ, Jackson AP, Jafri H, Rashid Y, et al. A fifth locus for primary autosomal recessive Microcephaly maps to chromosome 1q31. Am J Hum Genet. 2009; 67: 1578-80.

12. Muhammad F, Mahmood SH, Hansen L, Sajid MH, Anjum II, Aslam M, et al. Compound heterozygous ASPM mutations in Pakistani MCPH families. Am J Med Genet.Ê2009; 149: 926-30.

13. Kousar R, Nawaz H, Khurshid M, Ali G, Khan SU, Mir H, et al. Mutation analysis of the ASPM gene in 18 Pakistani families with autosomal recessive primary microcephaly. Child NeurolÊ2010; 25:715-20.

14. Bond J, Roberts E, Mochida GH, Hampshire DJ, Scott S, Askham JM, et al. ASPM is a major determinant of cerebral cortical size. Nat Genet 2002; 32: 316-20.

15. Ponting C, Jackson AP. Evolution of primary microcephaly genes and the enlargement of primate brains. Curr Opin Genet DevÊ2005; 15: 241-8.

16. Mekel-Bobrov N, Gilbert SL, Evans PD, Vallender EJ, Anderson JR, Hudson RR, et al. Ongoing adaptive evolution of ASPM, a brain size determinant in Homo sapiens. ScienceÊ2005; 309 : 1720-2.

17. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K,Church DM, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2005; 33: 39–45.

18. Altschul SF, Madden TL, Sch¨affer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSIBLAST:a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389–402.

19. Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 1990; 183: 63–98.

20. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997; 25: 4876–82.

21. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 2007; 24: 1596-9.

22. Tajima F. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics 1989; 123:585-95.

23. Nei M, Kumar S. Molecular Evolution and Phylogenetics. Oxford University Press, New York. 2000.

24. Gingerich PD, Ul-Haq M, von Koenigswald W, Sanders WJ, Smith BH, Zalmout IS. New protocetid whale from the middle eocene of pakistan: birth on land, precocial development, and sexual dimorphism. PLoS One 2009;4: 4366.

25. Spaulding M, O'Leary MA, Gatesy J. Relationships of Cetacea (Artiodactyla) among mammals: increased taxon sampling alters interpretations of key fossils and character evolution. PLoS One 2009; 23:4: 7062.

26. Hasegawa M, Thorne JL, Kishino H. Time scale of eutherian evolution estimated without assuming a constant rate of molecular evolution. Genes and Genetic Systems 2003; 78: 267–83.

27. Springer MS, Murphy WJ, Eizirik E, O'Brien SJ. Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc Nat Acad Sci USA 2003; 100: 1056–61.

28. Springer MS, Murphy WJ. Mammalian evolution and biomedicine: new views from phylogeny. Biological Reviews 2007; 82: 375-92.

29. Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. The DNA sequence and biological annotation of human chromosome 1. Nature 2006; 441: 315-21.

30. De S, Lopez-Bigas N, Teichmann SA. Patterns of evolutionary constraints on genes in humans. BMC Evol Biol 2008; 8: 275.

31. Kemkemer C, Kohn M, Cooper DN,Froenicke L, Högel J, et al. Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution. BMC Evol Biol 2009; 9:84.

32. Webster MT, Smith NG, Lercher MJ, Ellegren H. Gene expression, synteny, and local similarity in human noncoding mutation rates. Mol Biol Evol 2004 21:1820-30.

33. Caceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, et al. Elevated gene expression levels distinguish human from non-human primate brains. Proc Natl Acad Sci USA 2003; 100: 13030–5.

34. McHenry HM. Tempo and mode in human evolution. Proc Natl Acad Sci USA 1994; 91:6780–6.

35. Bond J, Scott S, Hampshire DJ, Springell K, Corry P, Abramowicz MJ, et al. Roberts E, Woods CG. Protein-Truncating Mutations in *ASPM* Cause Variable Reduction in Brain Size. Am J Hum Genet 2003; 73: 1170–7.

36. Keller I, Bensasson D, Nichols RA. Transitional-Transversion Bias is not universal: A counter example