Asian Journal of Psychiatry 6 (2013) 56–59

Contents lists available at SciVerse ScienceDirect

Asian Journal of Psychiatry

journal homepage: www.elsevier.com/locate/ajp

Evolutionary genetic analyses of MEF2C gene: Implications for learning and memory in *Homo sapiens*

Sunil V. Kalmady [a,b], Ganesan Venkatasubramanian [a,b,*], Rashmi Arasappa [a,b], Naren P. Rao [a,b]

[a] Department of Psychiatry, National Institute of Mental Health & Neurosciences, Bangalore, India
[b] Translational Psychiatry Laboratory, Cognitive Neurobiology Division, Neurobiology Research Centre, National Institute of Mental Health & Neurosciences, Bangalore, India

ARTICLE INFO

Article history:
Received 15 June 2011
Received in revised form 7 August 2012
Accepted 20 August 2012

Keywords:
Learning
Memory
Evolution
Hippocampus
MEF2C

ABSTRACT

MEF2C facilitates context-dependent fear conditioning (CFC) which is a salient aspect of hippocampus-dependent learning and memory. CFC might have played a crucial role in human evolution because of its advantageous influence on survival of species. In this study, we analyzed 23 orthologous mammalian gene sequences of MEF2C gene to examine the evidence for positive selection on this gene in *Homo sapiens* using Phylogenetic Analysis by Maximum Likelihood (PAML) and HyPhy software. Both PAML Bayes Empirical Bayes (BEB) and HyPhy Fixed Effects Likelihood (FEL) analyses supported significant positive selection on 4 codon sites in *H. sapiens*. Also, haplotter analysis revealed significant ongoing positive selection on this gene in Central European population. The study findings suggest that adaptive selective pressure on this gene might have influenced human evolution. Further research on this gene might unravel the potential role of this gene in learning and memory as well as its pathogenetic effect in certain hippocampal disorders with evolutionary basis like schizophrenia.

1. Introduction

Myocyte Enhancer Factor 2C (MEF2C) gene has been demonstrated to play a crucial role in hippocampal memory functions as well as neurodevelopment (Barbosa et al., 2008). MEF2C is highly expressed in the frontal cortex, entorhinal cortex, dentate gyrus, and amygdala (Leifer et al., 1994; Lyons et al., 1995). MEF2C facilitates context-dependent fear conditioning which is a salient aspect of hippocampus-dependent learning and memory (Barbosa et al., 2008). When animals or humans experience repeated pairings of a neutral conditional stimulus (CS) (for example, tone) and an unconditional stimulus (US) (for example, foot-shock), they subsequently display fear responses to the CS and the context in which the US occurred (Kim and Fanselow, 1992; Phillips and LeDoux, 1992). This latter form of learning is known as contextual fear conditioning, and it occurs whether the US is paired with a CS, unpaired with a CS in the context, or administered in the absence of any CS (Phillips and LeDoux, 1992). The lateral amygdala plays a critical role in context-dependent fear conditioning and this is regulated by hippocampus (Maren and Hobin, 2007). CFC might have played a crucial role in human evolution because of its advantageous influence on survival of species (Ohman, 2007). Since MEF2C underlies CFC, it is possible that this gene might have undergone adaptive evolution in *Homo sapiens*.

In this study, we present evidence (supported by different comparative genomic analyses) for site-wise positive selection on the MEF2C gene in humans. We analyzed 23 orthologous placental-mammalian gene sequences for human MEF2C gene to examine the evidence for positive selection using the improved branch-site model of Phylogenetic Analysis of Maximum Likelihood (PAML) as well as dual rate analyses and fixed effects likelihood analyses using HyPhy software.

2. Materials and methods

2.1. Gene sequences

Twenty two orthologous placental-mammalian gene coding sequences for human MEF2C genes (i.e. totally 23 sequences including that of human) were analyzed in the study, obtained from the Ensembl database (Release 50) (http://www.ensembl.org/) (Common Chimpanzee, Bornean Orangutan, Rhesus Macaque, Northern Greater Galago, Gray Mouse Lemur, Northern Treeshrew, Mouse, Brown Rat, Guinea Pig, Thirteen-Lined Ground Squirrel, American Pika, West European Hedgehog, Common Shrew, Little Brown Bat, Cat, African Bush Elephant, European Rabbit, Dog, Cow, Horse, Lesser Hedgehog Tenrec and Nine-Banded

* Corresponding author at: Department of Psychiatry, National Institute of Mental Health and Neurosciences, Bangalore 560029, Karnataka, India.
Tel.: +91 80 26995256; fax: +91 80 26564830.
E-mail addresses: venkat.nimhans@yahoo.com, gvs@nimhans.kar.nic.in (G. Venkatasubramanian).

1876-2018/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ajp.2012.08.007

Armadillo). The Ensembl gene/peptide IDs and accession numbers are given in Table 1.

2.2. Pre-processing of gene sequences

The reading frames of the orthologous gene sequences were checked and corrected in Bioedit software – version 7.0.9 (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html). To preserve the reading frame and prevent any alignment artifacts, the translated sequences were aligned instead of Coding DNA Sequences (CDS) sequences themselves. The multiple sequence alignment was performed using the software Multiple sequence Alignment using the Fast Fourier Transform (MAFFT) version-6 (http://align.bmr.kyushu-u.ac.jp/mafft/software/). MAFFT has been shown to produce good quality multiple sequence alignment with high sequence similarity and considerable structure similarity (Katoh et al., 2005). The aligned protein sequences were reverse translated to CDS, with reference to unaligned CDS using PAL2NAL server (version: v12; http://coot.embl.de/pal2nal/). A species tree file (for both PAML as well as HyPhy analyses) was constructed from the consensus mammalian tree (Figure-1) [(Murphy et al., 2001) and Tree of Life web project – http://tolweb.org/tree] essentially by pruning branches that were not represented by the orthologous sequences of the MEF2C gene. The processed sequences were analyzed by comparative genomic methods using PAML and HyPhy software.

2.3. Comparative genomic analyses

Comparative genomic methods analyze the non-synonymous (dN) and synonymous (dS) nucleotide substitution rates in orthologous gene sequences. Analyzing dN and dS are among the most direct ways to obtain evidence for positive selection on a protein coding gene (Anisimova and Liberles, 2007). We used improved branch-site model analysis (Yang, 2007; Zhang et al., 2005) in Phylogenetic Analysis of Maximum Likelihood (PAML) software (version-4b) (http://abacus.gene.ucl.ac.uk/software/paml.html). To examine for definitive evidence for positive selection, the log likelihood of the maximum-likelihood model A was compared to similar models lacking selection (i.e. null-hypothesis models) using two Log-likelihood Ratio Tests (LRT).

Model-A is the alternate hypothesis model for both the LRTs. For LRT-1, the null hypothesis is the site model (also referred to as M1a model (Yang et al., 2005)). For LRT-2, the null hypothesis is the branch-site model-A, but with $\omega_2 = 1$ fixed. This null model allows sites evolving under negative selection on the background lineages to be released from constraint and to evolve neutrally on the foreground lineages. Bayes Empirical Bayes (BEB) procedure was used to calculate the posterior probabilities that each site belongs to the site class of positive selection on the foreground lineages using the previously described method (Nielsen et al., 2005).

Additional tests were done using the software *Hy*pothesis testing using *Phy*logenies (HyPhy package – version 0.99 beta; http://www.hyphy.org/ (Pond et al., 2005)). Nucleotide substitution bias can have a significant effect on the proportions of synonymous and nonsynonymous substitutions, and, by extension, they can affect the estimates of dN and dS. To avoid this potential confound, we estimated the substitution bias model specific for these study data (model 010020) and composed it with MG94 to extend them to further MEF2C codon data analyses in HyPhy. Furthermore, to examine for site-wise positive selection in the MEF2C gene (to obtain further support for the sites identified in PAML analysis), we performed a two rate Fixed Effects Likelihood (FEL) analysis.

3. Results

Four Amino Acid (AA) sites on coding region of human MEF2C gene are independently identified by both BEB and dual rate FEL analysis as positively selected (Table 2). These sites fall into single cluster from 272AA to 278AA marked with phylogenetic variation while rest of the coding region is very well conserved among mammalian species. Moreover, haplotter analysis revealed ongoing positive selection on the MEF2C gene particularly in a central European population [Haplotter phase-1 data analysis: Region (Start) – 88100239; Region (End) – 88263097; Central European population $p = 0.049735$; Yeruba population $p = 0.999955$; Asian population $p = 0.359281$].

4. Discussion

The study findings show evidence for significant positive selection on 4 codon position in human MEF2C gene. The significant positive selection in human MEF2C gene was also supported by HyPhy analyses that incorporated synonymous rate

Table 1
Table lists the gene and protein Ensembl IDs and accession numbers for MEF2C coding sequences from 23 placental-mammalian species.

Species	Common name	Ensembl gene IDs	Ensembl peptide IDs
Homo sapiens	Human	ENSG00000081189	ENSP00000340874
Pan troglodytes	Common Chimpanzee	ENSPTRG00000024214	ENSPTRP00000042298
Pongo pygmaeus	Bornean Orangutan	ENSPPYG00000015621	ENSPPYP00000017465
Macaca mulatta	Rhesus Macaque	ENSMMUG00000019836	ENSMMUP00000026081
Otolemur garnettii	Northern Greater Galago	ENSOGAG00000006790	ENSOGAP00000006073
Microcebus murinus	Gray Mouse Lemur	ENSMICG00000005703	ENSMICP00000005204
Tupaia belangeri	Northern Treeshrew	ENSTBEG00000006794	ENSTBEP00000005866
Mus musculus	House Mouse	ENSMUSG00000005583	ENSMUSP00000105188
Rattus norvegicus	Brown Rat	ENSRNOG00000033134	ENSRNOP00000046667
Cavia porcellus	Guinea Pig	ENSCPOG00000011104	ENSCPOP00000009986
Spermophilus tridecemlineatus	Thirteen-Lined Ground Squirrel	ENSSTOG00000001883	ENSSTOP00000001678
Ochotona princeps	American Pika	ENSOPRG00000012052	ENSOPRP00000010995
Erinaceus europaeus	West European Hedgehog	ENSEEUG00000002968	ENSEEUP00000002719
Sorex araneus	Common Shrew	ENSSARG00000012570	ENSSARP00000011351
Myotis lucifugus	Little Brown Bat	ENSMLUG00000008230	ENSMLUP00000007513
Felis catus	Cat	ENSFCAG00000011595	ENSFCAP00000010764
Loxodonta africana	African Bush Elephant	ENSLAFG00000013761	ENSLAFP00000011521
Oryctolagus cuniculus	European Rabbit	ENSOCUG00000015467	ENSOCUP00000013292
Canis familiaris	Dog	ENSCAFG00000008302	ENSCAFP00000012227
Bos taurus	Cow	ENSBTAG00000020701	ENSBTAP00000027589
Equus caballus	Horse	ENSECAG00000024716	ENSECAP00000022307
Echinops telfairi	Lesser Hedgehog Tenrec	ENSETEG00000015637	ENSETEP00000012682
Dasypus novemcinctus	Nine-Banded Armadillo	ENSDNOG00000000377	ENSDNOP00000000286