Journal of Advanced Bioinformatics Applications and Research
ISSN 0976-2604.Online ISSN 2278 – 6007
Vol 4, Issue 1, 2013, pp15-19
http://www.bipublication.com

DISTRIBUTION AND PHYLOGENETIC ANALYSIS OF MITOCHONDRIAL ENZYME AMONG DIFFERENT TAXA.

Tribhuwan Chandra[1], Ashish Thapliyal[1], Prashant Anthwal[1], Madhu Thapliyal[2],

Promila Sharma[1]and Kumud Pant[1]

[1]Department of Biotechnology, Graphic Era University, Dehradun, Uttarakhand, India
[2] Departments of Zoology, R C U PG College Uttarkashi, Uttarakhand, India
tribhuwan.chandra@gmail.com

[Received-01/12/2012, Accepted-21/02/2013]

ABSTRACT:

Comparative genomics of whole genome data of different taxa has led to the suggestion that the process of gene transfer is a more influential evolutionary mechanism in the evolution and distribution of GDH genes. GDH sequences were identified and sequences from the four different classes of the enzyme as queries. The amino acid sequences were aligned and clearly aligned regions were removed. Protein maximum likelihood bootstrap values were calculated by analysis of re-sampled datasets using the default parameters followed by total rearrangements of each replicate. The phylogenetic distribution pattern of the gdh genes have showed that all classes of gdh genes have been found in eubacteria, gdh4 have been found in eukaryotes while only gdh2 have been found in archea. The correlation between eubacterial phylogeny and distribution of gdh genes indicates that a gene transfer must have played a significant role in the evolution of this gene family. The gdh1 and gdh2 phylogenetic analyses strongly indicate gene transfer in evolution of gdh gene in proteobacteria. The phylogenetic analysis fails to support the distribution of gdh gene in archaea by simple evolutionary process. The phylogenetic reconstructions of gdh1 and gdh2 in combination with the phylogenetic analyses provide strong support for multiple inter and intra domain gene transfer in eukaryotes and lead present occurrence of gdh genes among different taxa.

Keywords: Glutamate Dehydogenase, Lateral gene transfer (LGT), Hexameric, Homologs, Polyphyletic, Phylogenetic

INTRODUCTION

Glutamate dehydrogenase (GDH) is an enzyme, present in most microbes and the mitochondria of eukaryotes, as are some of the other enzymes required for urea synthesis, that converts glutamate to α-ketoglutarate, and vice versa. Glutamate dehydrogenase enzyme is very diverse and can be divided into four distinct classes. GDH-1 and GDH-2 are small hexameric enzymes with a broad taxonomic distribution [1, 2]. Previously GDHs have been found only in fungi and protists, function in glutamate catabolism. The evolution and distribution of GDH among different taxa is thought to be occurred by the certain evolutionary events.

Lateral gene Transfer emerged as a potentially important evolutionary mechanism in evolution of GDH. With the availability of more and more genomic sequences from eukaryotes become available and new insight provided by technology and researcher across the world involving many different lineages have indicated LGT as a potentially important evolutionary mechanism also in eukaryotic organisms. In eukaryotes there is two types of gene transfer one involves the transfer of genes from the organelles to the nucleus of the eukaryotic cell (endosymbiotic gene transfer) and LGT between unrelated species. In eukaryotes endosymbiotic gene transfer is widely accepted as an important source of genetic material [3]. In determining the phylogenetic relationships among domains of life, transfer of genes between organisms, and the bringing together of different genomes into a single organism, remain important evolutionary factors. Lateral gene transfer (LGT) is always been considered to be significant evolutionary mechanism in prokaryotic genome evolution. It may be the most important mechanism for evolutionary innovation in Eubacteria and Archaea [4, 5]. Eukaryotes are less studied for occurrence of LGT as compared to the prokaryotes because of the lack of complete genome sequences. There are several cases of gene transfer between prokaryotes and eukaryotes have been recorded. Recently it has been reported that a number of transfers involving prokaryote to-eukaryote direction, and also between different eukaryotic lineages [6, 7, 8]. Finally, enzymes of a fourth class have been discovered in Eubacteria [9]. Phylogenetic analyses in a more recent study were unfortunately based on an alignment of all four classes of the enzyme, that are very distantly related [9] making them difficult to interpret. The phylogenetic distribution pattern of the different GDH classes between species and the phylogenetic analyses of the classes themselves clearly indicated that gene transfer was likely to be a significant evolutionary mechanism in the evolutionary history of these enzymes. The glutamate dehydrogenase (GDH) gene families have been selected to investigate the distribution of this gene and its transfer in prokaryotic versus eukaryotic genome evolution.

MATERIAL AND METHOD

GDH sequences were identified using BLAST searches against a variety of databases with sequences from the four different classes of the enzyme as queries. BLAST, is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences [10]. Some published sequences were retrieved from the National Center for Biotechnology Information and unpublished eukaryotic sequences were retrieved from NCBI using the other eukaryotes BLAST service in their genomic BLAST pages and the "microbial genomes" BLAST service at NCBI were searched for homologs. Unpublished GDH-3 sequences GDH-1 and GDH-2 sequences also being retrieved to check whether eukaryotic groups originated by endosymbiotic gene transfer from the mitochondria and chloroplast. The amino acid sequences within each dataset were aligned using CLUSTALW [11] and overlapping regions were identified and removed. Sequences with high similarity (>80% amino acid sequence identity) within the unambiguously aligned regions were excluded from the dataset to reduce the computational time. TREE-PUZZLE, version 4.02 [12] were applied to the datasets and sequences that failed the test were excluded from further phylogenetic analysis.

PHYLIP (the *PHYL*ogeny *I*nference *P*ackage) is a package of programs for inferring phylogenies (evolutionary trees) were being used to create phylogram. PMBML [13] within the PHYLIP package have been used to infer protein phylogeny. Output is written onto special files "resultf" and "resultt". Trees written onto "resultt" are in the Newick format, standard used by authors of a number of major phylogeny packages.

Protein maximum likelihood bootstrap values were calculated by analysis of 500 re-sampled sets using the default parameters followed by total length rearrangements were performed for each set. Protein maximum likelihood distance bootstrap values for bipartitions were calculated by analysis of 500 re-sampled datasets using PUZZLEBOOT.

Complete list of accession numbers for the sequences used in the analyses, are available in GenBank under the accession numbers AF533881-AF533889.

RESULTS AND DISCUSSION

All available homologs of GDH were downloaded from public databases to study the phylogenetic analysis of evolution of *gdh* genes among different taxonomic units. It is being known that GDH-1 is found in eubacteria and eukaryotes, GDH-2 is found in all domains of life, and GDH-4 is only found in eubacteria [1, 14]. However, we identified two GDH-3 genes from the proteobacteria *Desulfovibrio vulgaris* and *Geobacter sulfurreducens*; GDH-3 was previously found only in eukaryotes including fungi, euglenozoa and apicomplexa [15]. Five new GDH-1 and two new GDH-2 sequences were obtained. GDH-1 and GDH- 2 cDNA clones from the red alga *Porphyrayezoensis*, GDH- 1 cDNA clones. The oomycete *Phytophthora infestans*, diplomonad *Spironucleus barkhanus*, and the parabasalid *Trichomonas vaginalis*, and a GDH-2 cDNA clone from the green alga *Chlamydomonas reinhardtii* were made available from the various EST projects [2, 16].

PHYLOGENETIC ANALYSES

The outcome of the phylogenetic analysis was used to infer the pattern emerged among different groups (Fig. 1, 2 and 3). The inferred GDH amino acid sequences were aligned individually and clearly aligned regions were identified. Sequences with deviant amino acid composition were excluded to reduce the redundancy relative to the phylogenetic signal in the dataset. In previous studies, the different families of GDH have been aligned and combined phylogenetic analyses have been performed

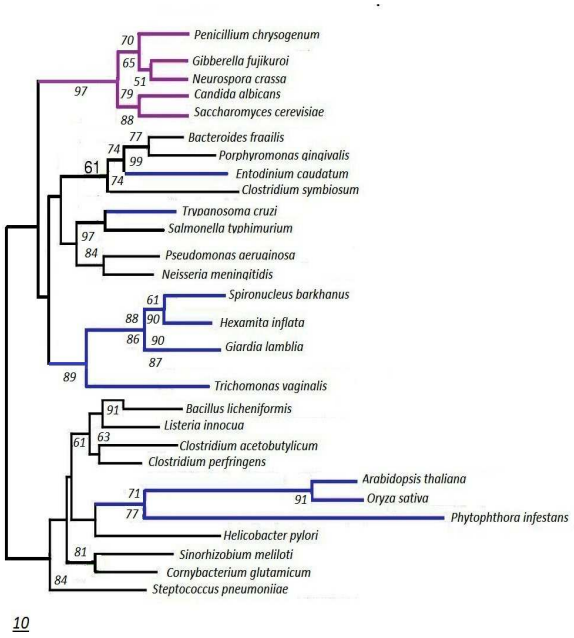

Fig. 1: Phylogenetic tree of GDH1gene distributed among different taxa.

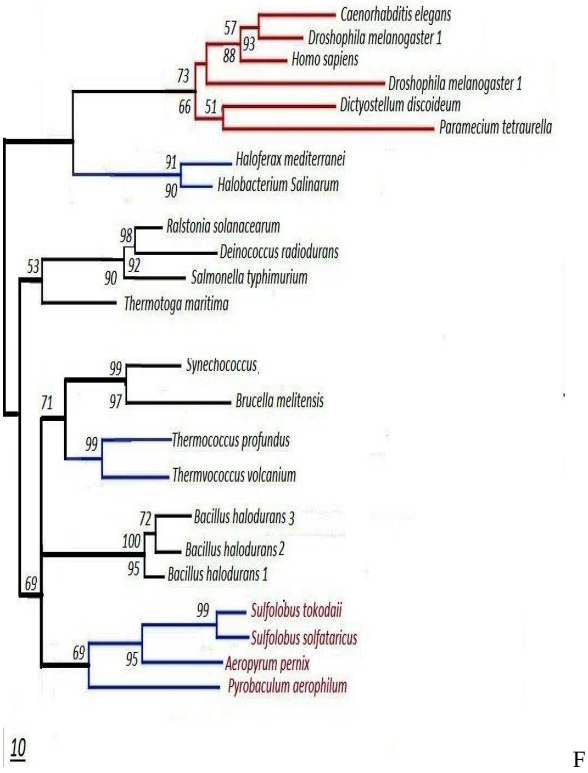

Fig.2: Phylogenetic tree of GDH2 gene distributed among different taxa.

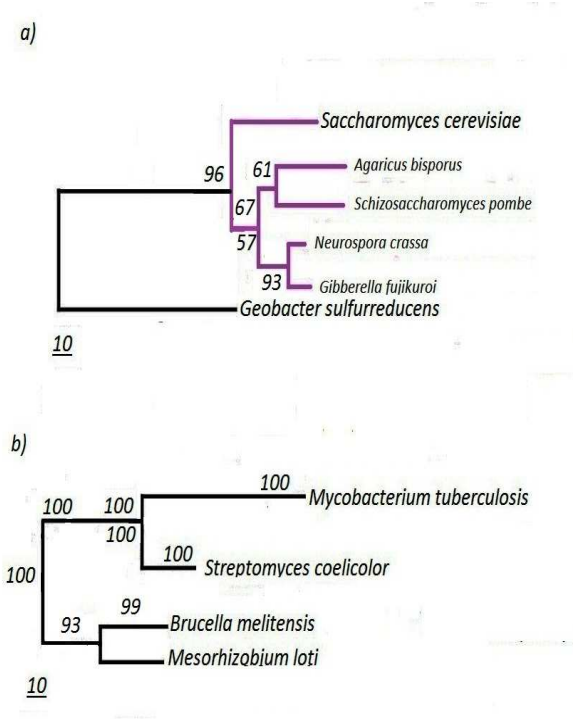

Fig. 3: Phylogenetic tree of a) GDH3 and b) GDH4 gene distributed among different taxa.

[1, 2]

The GDH-1 and GDH-2 phylogenetic analyses indicate that Gene transfer has played an important role in the evolution of these gene families. For example, proteobacterial sequences are found in five groups in GDH-1, three of which are separated with statistical support values >90% in both bootstrap analyses, and five groups in GDH-2, both analyses. In the GDH-1 phylogenetic reconstruction the *Sinorhizo biummeliloti* groups with *Corynebacterium glutamicum*, and the *Pasteurella multicoda* and *Salmonella typhimurium* form a strong group with *Deinococcusra diodurans* and the unicellular eukaryote *Trypanosoma cruzi*, to the exclusion of the other proteobacterial sequences in the tree. Similarly, in the GDH-2 reconstruction the *Brucellamelitensis* groups with two cyanobacterial sequences, while the two other sequences group with a large eukaryotic cluster. Also, the two sequences fail to group together in the GDH-2 tree – one is an immediate outgroup to a group with low G+C gram positive sequences, while the other is in a strongly supported cluster with a proteobacterial sequence and the *Deinococcus* sequence. The inferred pattern explain the distribution of the proteobacterial GDH sequences. The polyphyletic pattern is not unique to proteobacteria within the eubacterial domain; the three cyanobacterial GDH-2 sequences are separated into two distinct clusters with bootstrap support values of 90% and 88%, respectively and the low G+C gram positives sequences are split into at least two groups each for GDH-1 and GDH-2.

The GDH-2 gene family is found in Archaea. At first glance, this might be taken as evidence that an ancestral archaeon encoded this class of the gene and that it was passed on to extant Archaea by vertical inheritance. However, the phylogenetic analysis of GDH-2 argues against this simple interpretation. The archaeal sequences are split into three distinct groups the *Thermoplasma* and *Thermococcus* sequences form a cluster with a cyanobacterial/α-proteobacterial group which is nested within a cluster of low G+C gram positive eubacteria with a bootstrap support value of 66% in the maximum likelihood analysis the crenarchaeote sequences group with another cluster of low G+C gram positives with 74% bootstrap support in the same analysis, and the *Halobacterium* and the *Haloferax* sequences form a group that is excluded from the two other archaeal clusters. The phylogenetic analysis of GDH1 suggests that interdomain transfer may not be limited to prokaryotes the *T. cruzi* and the *Entodinium caudatum* sequences are phylogenetically distant from the other eukaryotic clusters. The *Trypanosoma* sequence forms a group with two proteobacterial sequences and a *Deinococcus* sequence, with a bootstrap support values of >85% for the bipartition. A second gene transfer event inferred to explain the presence of a *gdh1* gene sequence in the ciliate *E. caudatum* which groups with sequences from *Bacteroides*

and *Porphyromonas* with a bootstrap support values of >75%. The phylogenetic analysis of GDH-1 suggests additional gene transfer events; eukaryotes emerge in five different places in the tree. Three of the eukaryotic groups, the plant/oomycete cluster, the large protist cluster and the fungi/red algal cluster could indeed represent a large eukaryotic GDH1 group. Two eukaryotic groups are found in the GDH2 phylogenetic tree, one cluster with two green algal sequences and an *Arabidopsis* sequence, and a second larger cluster. As mentioned, two α-proteobacterial sequences form a strongly supported group with the larger eukaryotic cluster. This proteobacterial/eukaryotic cluster is a sister to the green algal/land plant cluster. Several different evolutionary scenarios could have produced this pattern. In this case, the eukaryote lineage would be one large clade with the eubacterial grouping arising from within them.

The phylogenetic reconstructions of GDH-1 and GDH-2 in combination with the phylogenetic distribution analyses provide strong support for multiple inter and intra domain gene transfer events. The phylogeny of GDH-3 and GDH-4 fail to indicate transfer events within organismal groups.

CONCLUSIONS

The phylogenetic reconstructions showed that analyses of distribution patterns of the four *gdh* gene families provide strong support for gene transfers involving prokaryotes, as well as microbial and eukaryotes. Differential gene loss does not seem to have played an important role in the evolution of *gdh* genes in any of the three domains of life.

REFERENCES

[1]. Benachenhou-Lahfa, N, et al., (1993), Evolution of glutamate dehydrogenase genes: evidence for two paralogous protein families and unusual branching patterns of the archaebacteria in the universal tree of life. *J. Mol. Evol.*, 36, 335-346.

[2]. Minambres, B., et al., **(**2000**),** A new class of glutamate dehydrogenases (GDH). Biochemical and genetic characterization of the first member, the AMP-requiring NAD-specific GDH of *Streptomyces clavuligerus*. *J Biol Chem* 275, 39529-39542.

[3]. Martin, W. et al., (2002), Evolutionary analysis of *Arabidopsis*, cyanobacterial and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA, 99, 12246–12251.

[4]. Doolittle, W. F. (1999), Phylogenetic classification and the universal tree. Science, 284, 2124–2129.

[5]. Lawrence, J. G, (2000), Lateral gene transfer and the nature of bacterial innovation. Nature, 18, 405(6784), 299-304.

[6]. Andersson, J.O, et al., (2003), Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr. Biol.*, 13, 94-104.

[7]. Andersson, J. O., et al., (2001), Are there bugs in our genome? Science, 292, 1848–1850.

[8]. Andersson, J.O, et al., (2002), Evolutionary analyses of the small subunit of glutamate synthase: gene order conservation, gene fusions and prokaryote-to-eukaryote lateral gene transfers. *Eukaryot Cell*, 1, 304-310.

[9]. Doolittle, W. F., (1998), You are what you eat: a gene transfer could account for bacterial genes in eukaryotic nuclear genomes. Trends. Genet., 14, 307–311.

[10]. Altschul, S.F, et al., (1990), Basic local alignment search tool. J. Mol. Biol., 215, (3), 403-10.

[11]. Thompson, J. D, et al., (1994), CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *N. A. Res*, 22, 4673-4680.

[12]. Strimmer, K, et al., (1996), Quartet puzzling: a quartet maximum-likelihood method for reconstructing treetopologies. *Mol. Biol. Evol.*, 13, 964-969.

[13]. Veerassamy S, et al., A transition probability model for amino acid substitutions from BLOCKS *J. Comput. Biol*.

[14]. Brown, J. R, et al., (1997), Archaea and the prokaryote-to eukaryote transition. *Microbiol. Mol. Biol. Rev.*, 61, 456-502.

[15]. Asamizu E, et al., (2000), Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of *Chlamydomonas reinhardtii*. *DNA Res.*, 7, 305-307.

[16]. Kamoun, S, et al., (1999), Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences. *Fungal Genet. Biol.*, 28, 94-106.