

## Convergence pattern studies of folate receptor

Kalidoss Ramamoorthy<sup>1</sup> and Rama Shanker Verma<sup>1,\*</sup>

<sup>1</sup>Stem cell and Molecular biology Laboratory, Department of Biotechnology, Indian Institute of Technology Madras, Chennai 600 036 TN India; Rama Shanker Verma\* - Email: vermars@iitm.ac.in; Phone: 91 44 2257 4109; \* Corresponding author

received April 29, 2008; revised October 15, 2008; accepted October 30, 2008; published December 06, 2008

### Abstract:

Gene patterns and sequences of folic acid synthesizing genes that are converged as meaningful patterns during evolution in the higher eukaryotes has been identified using sequence alignment and pattern analysis. Based on the finding, we are postulating that part of genes that are involved in synthesis of folic acid in lower eukaryotes, are converged into meaningful similar functional domains of folate receptor in higher eukaryotes.

**Keywords:** PROSITE patterns; MUSCLE algorithm; Folate receptors (FRs); convergence

### Background:

Folate, the water-soluble vitamin of B complex family is an essential cofactor in several eukaryotic metabolic pathways [1]. It facilitates methylation reaction and transfer of "one carbon units" for purines and pyrimidines biosynthesis. Folate is also involved in the biosynthesis of methionine, histidine and serine. Availability of folate is essential for various metabolic pathways in higher eukaryotes [2]. Prokaryotes have the ability to synthesize folate through metabolic processes. However, eukaryotes are unable to synthesize folates and hence they require an external source as supplements. Mammals have the ability to synthesize the pteridine ring, but are unable to link it with other compounds such as p-Amino benzoate and glutamic acid [3]. In general there are two types of folate transporters present in higher eukaryotes; membrane channels and carriers; they vectorially move molecules into endocytic vesicles and are internalized. Transport systems can be distinguished by their preferences of various folates as substrates, as well as differences in physiological conditions such as temperature and pH. The main mechanism by which folate is transported from the extracellular space to the cytoplasm are the carrier mediated or receptor, which operate efficiently at natural physiological pH [2]. Here in this report, we propose a hypothesis that the genes which are present in lower eukaryotes, partially converged into a meaningful pattern in higher eukaryotes of similar function. We tested this hypothesis using folate synthesizing genes from lower eukaryotes that converged into receptor and binding proteins in higher eukaryotes.

### Methodology:

#### Dataset

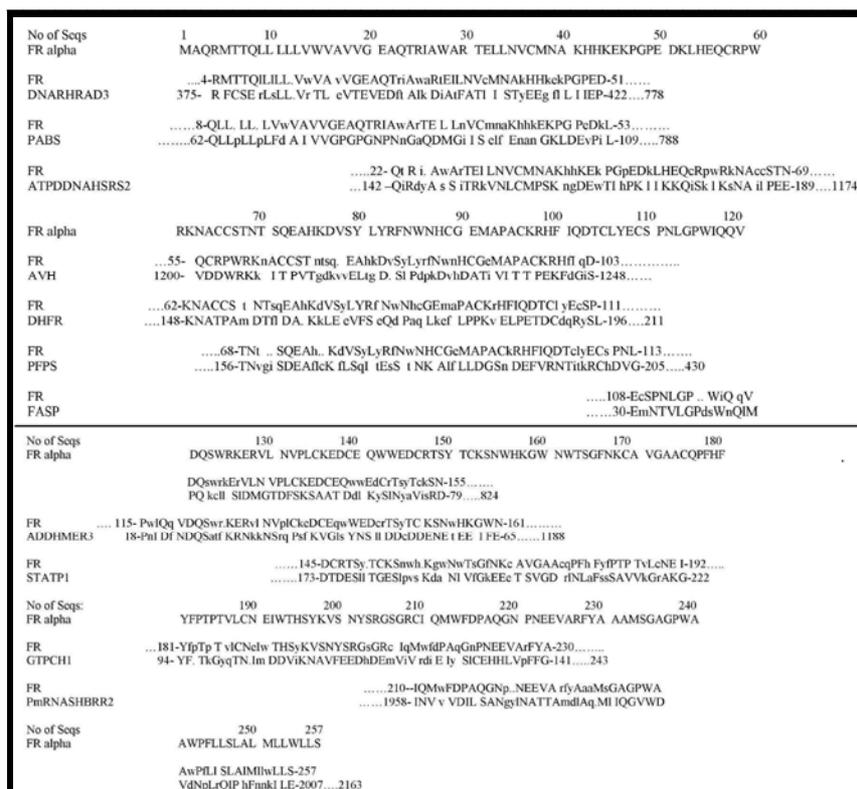
Nucleic acid sequences of Folate receptor were collected from the Swissprot (<http://www.expasy.org/>). Three types of folate receptor variants sequences were obtained, which are named as alpha, beta, and gamma. The common folate receptors found in humans belong to the alpha variant. Even within the alpha variants, many sub variants are reported [4].

### PROSITE pattern search

Pattern set were obtained from prosite database. All other sequences for folate synthesizing gene were obtained from the KEGG (Koyoto Encyclopedia of Genes and Genomes) database (<http://www.genome.ad.jp/kegg/pathway.html>). To get the data out from the kegg data base we used Folate, folate synthesizing gene and folate metabolism and folate synthesis as key words. After getting the information about pathways, several enzymes involved in the pathways were noted and search for pattan analysis using swiss pro data base. The SWISSPROT ID for the enzymes are P12954-ATP dependent DNA Helicase SRS2, P37254-Probable polyglutamate synthase, P06839-DNA repair helicase RAD-3, P13587-Sodium transport ATPase 1, Q12691-ATP dependent DNA helicase SRS-2, P53848-Folic acid synthesis protein FOL-1, P07807-Dihydrofolate reductase, P51601-GTP Cyclohydrolase 1, P51979-ATP dependent DNA helicase MER-3, P36001-Probable folypolyglutamate synthase, P53327-Antiviral helicase SLH-1, P32639-Pre mRNA splicing helicase BRR-Genes involved in synthesis of folate in lower eukaryotes (*Saccharomyces cerevisiae*) are summarized in Table 1 (supplementary material). Prosite database also helped in identifying functions of uncharacterized proteins translated from genomic or cDNA sequences and in the determination of protein function with biologically meaningful patterns and profiles [5-9].

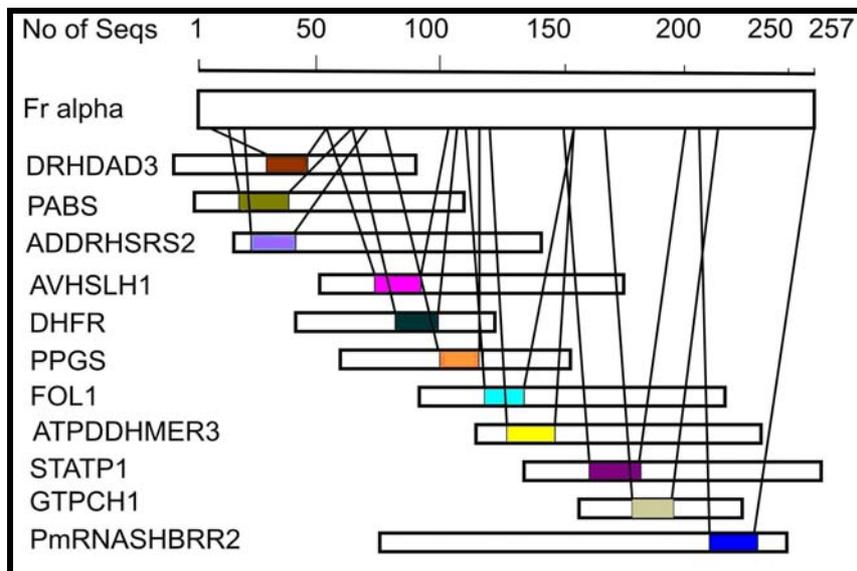
### MUSCLE alignment

MUSCLE is a comparatively new program and used in creating multiple alignments of protein sequences (10, 11). Multiple alignments of protein sequences are important in many applications, including phylogenetic tree estimation, structure prediction and critical residue identification. In this experiment, we used MUSCLE (MUltiple Sequence Comparison by Log-Expectation), a Multiple Sequence Alignment program with better average accuracy and speed compared to ClustalW2 or T-Coffee and gives output as a summary of the algorithm [12].



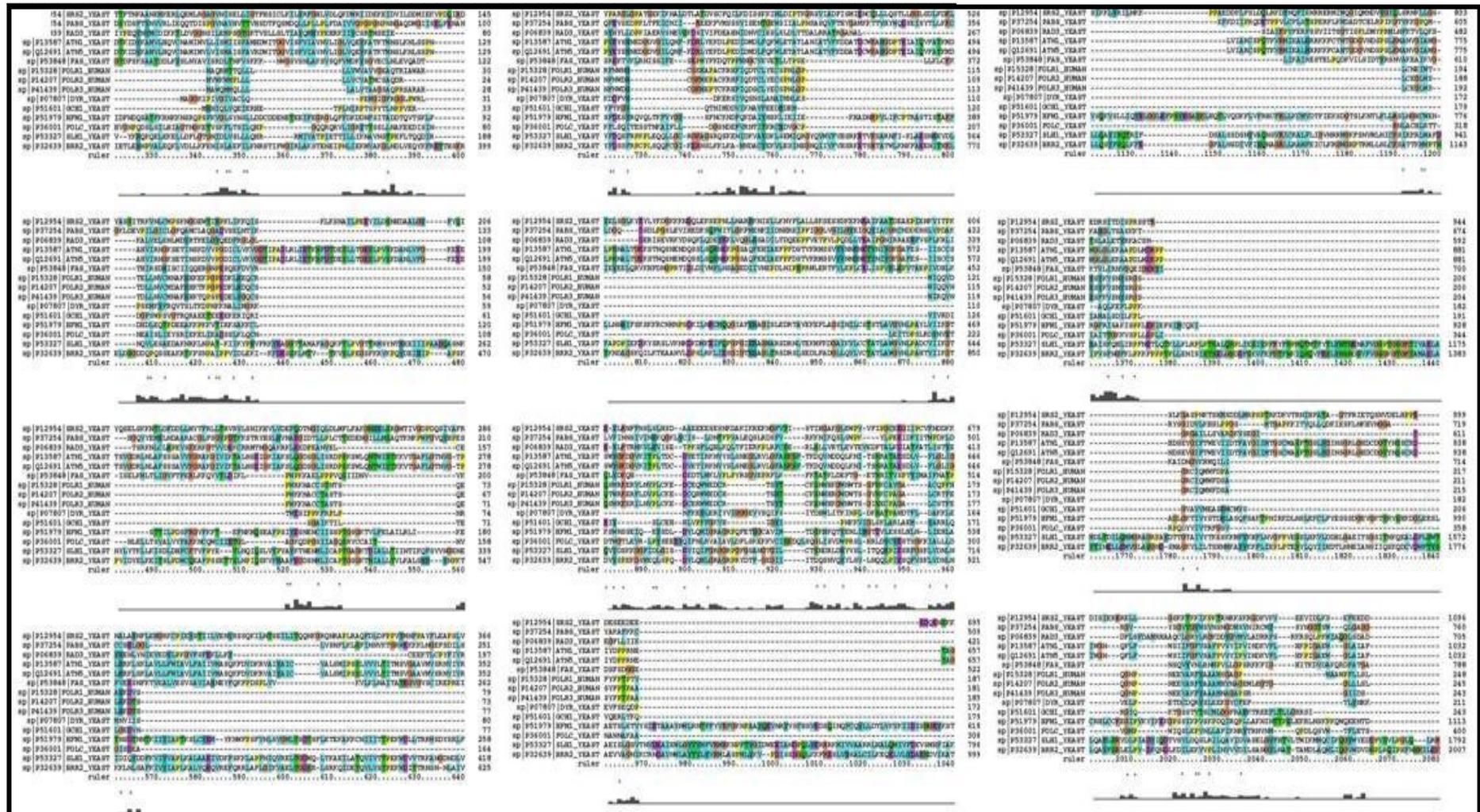
**Figure 1:** Prosite scan of folate receptor alpha (Human) and folate synthesizing genes (Yeast)

FR = Folate receptor; DNARHRAD3 = DNA repair helicase RAD-3; ATPDDNAHSRS2 = ATP dependent DNA helicase SRS-2; AVH = Anti viral helicase SLH-1; DHFR = Dihydrofolate reductase; PFPS = probable folypolyglutamate synthase; FASP = Folic acid synthesis protein; ADDHMER3 = ATP dependent DNA helicase MER-3; STATP1 = Sodium transport ATPase-1; GTPCH-1 = GTP cyclohydrolase-1; PmRNASHBRR2 = Pre mRNA splicing helicase BRR-2.



**Figure 2:** Schematic representation of convergence of Folate receptor (Human) from folate synthesizing genes (Yeast).

Proteins from colored boxes represents patterns which related to folate receptor colored boxes shows corresponding patterns which converged from folate synthesizing genes from folate receptor. FR = Folate receptor; DNARHRAD3 = DNA repair helicase RAD-3; ATPDDNAHSRS2 = ATP dependent DNA helicase SRS-2; AVH = Anti viral helicase SLH-1; DHFR = Dihydrofolate reductase; PFPS = probable folypolyglutamate synthase; FASP = Folic acid synthesis protein; ADDHMER3 = ATP dependent DNA helicase MER-3; STATP1 = Sodium transport ATPase-1; GTPCH-1 = GTP cyclohydrolase-1; PmRNASHBRR2 = Pre mRNA splicing helicase BRR-2.



**Figure 3:** Alignment of the amino acid sequence of folate receptor from Human (alpha, beta and gamma) and folate synthesizing genes from yeast. Asterisk (\*) shows similarity between folate receptors and folate synthesizing genes.

### Discussion:

To evaluate the similar motifs (amino acid), between folate receptor and folate synthesizing gene, we used the protein sequence of Folate receptor- $\alpha$  and scanned it with DNA repair helicase RAD 3 (EC 3.6.1.-) enzyme, which forms a pattern that showing the region of 4 to 51 of the folate receptor alpha (Figure 1), of which 9 amino acid exactly matches and 24 were same group of amino acids rest are different residues. This observation led us to scan more patterns of folate synthesizing gene with folate receptor alpha protein sequence. Para amino benzoate synthase (E.C 6.3.5.8) scanning results gave a meaningful pattern with in the region of 8 to 53 of FR $\alpha$ , out of which 11 exact amino acid residues are matched, 23 amino acids were in similar group. ATP dependent DNA helicase SRS-2 (E.C 3.6.1.-) shows a pattern in the region of 23 to 69 of FR $\alpha$ , out of which 11 amino acid residues were exactly matched in both the sequence and 19 were similar group of amino acids. Antiviral helicase SLH-1 (E.C 3.6.1.-) shows a pattern in the region of 55 to 103 of FR $\alpha$ , of which 8 amino acid residues exactly matched in both sequences, and 24 were similar groups of amino acids. Dihydrofolate reductase (E.C 1.5.1.3) showed the pattern in between the amino acid residues 62 to 111 of FR $\alpha$ , of which 9 amino acid residues were exactly matched and 25 were similar groups of amino acids.

Probable poly glutamate synthase (E.C 6.3.2.17) exhibited the pattern in between the amino acid residues of 68 to 113 of FR $\alpha$ , of which 11 amino acid residues were exactly matched and 22 were similar groups of amino acids. Folic acid synthesis protein (E.C 2.5.1.15) also showed a pattern in between the amino acid residues of 108 to 155 of FR $\alpha$ , of which 8 amino acid residues were exactly matched and 24 were similar groups of amino acids. Similarly, ATP dependent DNA helicase MER-3 (E.C 3.6.1.-) showed a pattern in between the amino acid residues of 115 to 161 of FR $\alpha$ , of which 7 amino acid residues were exactly matched and 24 were similar groups of amino acids. Sodium transport ATPase 1 (E.C 3.6.3.7) showed a pattern in between amino acid residues of 145 to 192 of FR $\alpha$ , of which 8 amino acid residues were exactly matched and 21 were similar groups of amino acids. GTP cyclohydrolase1 (E.C 3.5.4.16) shows a pattern in between the amino acid residues of 181 to 230 of FR $\alpha$ , of which 7 amino acid residues were exactly matched and 27 were similar groups of amino acids. Pre mRNA splicing helicase BRR-2 (E.C 3.6.1.-) showed a pattern in between the amino acid residues of 210 to 257 of FR $\alpha$ , of which 11 amino acid residues were exactly matched and 22 were similar groups of amino acids. We have combined all the patterns that were obtained in the scanning results and created complete length of converged folate receptor alpha sequence with their synthesizing protein sequences. These results indicate that part of the synthesizing genes was partially converged into folate receptor alpha sequence in higher eukaryotes (human) (Figure 2).

To prove further this convergence we wanted to see how many motif are actually present in folate receptor form various part of genes which are part of folate synthesis in yeast, we used muscle alignment (multiple sequence

alignment) which used Gonnet Pam 250 protein weighing matrix was helpful in identifying domain similarities in between the distantly related protein sequences. This gave a good alignment scores between protein sequences in folate receptors and folate synthesizing genes. Possibility of the gene conservation as well as convergence of folate synthesizing genes from lower eukaryotes (Yeast) into folate transport protein (folate receptor) of higher eukaryotes (human), we aligned all the folate synthesizing genes from Yeast and of folate receptor protein sequences (Human). After alignment maximum conserved domains we identified as shown in Figure 3. The aligned sequence shows the part of folate synthesizing genes are converged to form folate receptors in higher eukaryotes. In this alignment P15328, P14207 and P41439 are folate receptors ( $\alpha$ ,  $\beta$  and  $\gamma$ ) from human and other sequences are folate synthesizing genes from Yeast. Maximum residues are conserved from Sodium transport ATPase-1, sodium transport ATPase-5, folic acid synthesis protein, antiviral helicase SLH-1 and pre mRNA splicing helicase BRR-2. These results clearly validated our hypothesis that genes for folate receptors in higher organisms (human) are converged from folate synthesizing genes of lower eukaryotes (yeast) during evolutionary process.

### Conclusion:

We described gene evolution and conversance pattern of the similar functional domains. Here, we have chosen a model of receptor and synthesizing genes of similar function. Similarly, other model can be evolved to test this hypothesis. These conserved and converged region maybe used to develop as target for drug development in future.

### Acknowledgment:

This work was supported by the grant from MHRD (0506/007/RAMS) and DBT (07/08/032/RAMS) and partial support to KR from DBT is acknowledged.

### References:

- [01] M. Susanne *et al.*, *Coulson J Nutr.*, 127: 30 (1997) [PMID: 9040540]
- [02] L. Falquet *et al.*, *Nucleic acids res.*, 30: 235 (2002) [PMID: 11752303]
- [03] O. Stranger, *Curr Drug Metab.*, 3: 211 (2002) [PMID: 12003352]
- [04] S. Iwakiri *et al.*, *Ann. Surg. Oncol.*, 15: 889 (2008)
- [05] A. Gattiker *et al.*, *Appl Bioinformatics*, 1: 107 (2002) [PMID: 15130850]
- [06] A. Bairoch, *Nucleic Acids Res.*, 21: 3097 (1993) [PMID: 8332530]
- [07] N. Hulo *et al.*, *Nucleic Acids Res.*, 34: 227 (2006) [PMID: 16381852]
- [08] J. Sigrist *et al.*, *Brief. Bioinform.*, 3: 265 (2002) [PMID: 12230035]
- [09] N. Hulo *et al.*, *Nucleic Acids Res.*, 32: 134 (2004)
- [10] J. D. Thompson *et al.*, *Nucleic acids Res.*, 27: 2682 (1999) [PMID: 10373585]
- [11] R. C. Edgar, *Nucleic Acids Res.*, 32: 1792 (2004) [PMID: 15034147]
- [12] R.C. Edgar, *BMC Bioinformatics.*, 19:113 (2004) [PMID: 15318951]

---

Edited by P. Kanguane

Citation: Ramamoorthy & Verma, Bioinformatics 3(4): 168-172 (2008)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

---

### Supplementary material

S. No.	Folate synthesizing genes
1.	Antiviral helicase SLH-1(E.C 3.6.1.-)
2.	ATP dependent DNA helicase MER-3 (E.C 3.6.1.-)
3.	ATP dependent DNA helicase SRS-2 (E.C 3.6.1.-)
4.	Sodium transport ATPase-1 (E.C 3.6.3.7)
5.	Dihydrofolate reductase (E.C 1.5.1.3)
6.	DNA repair helicase RAD-3 (E.C 3.6.1.-)
7.	Folic acid synthesis protein FOL-1 (E.C 12.5.1.15)
8.	GTP cyclohydrolase-1 (E.C 3.5.4.16)
9.	Pre mRNA splicing helicase BRR-2 (E.C.3.6.1.-)
10.	Probable folypolyglutamate synthase (E.C 6.3.2.17)
11.	Yeast para amino benzoate synthase (E.C 6.3.5.8)

**Table 1:** List of folate synthesizing genes