FEBS Letters 581 (2007) 5526–5532

# Phylogenetic and evolutionary analysis of the septin protein family in metazoan

Lihuan Cao[1], Xiangming Ding[1], Wenbo Yu, Xianmei Yang, Suqin Shen, Long Yu*

*State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai 200433, PR China*

Received 9 August 2007; revised 11 October 2007; accepted 19 October 2007

Available online 29 October 2007

Edited by Takashi Gojobori

**Abstract** **Septins, a conserved family of cytoskeletal GTP-binding proteins, were presented in diverse eukaryotes. Here, a comprehensive phylogenetic and evolutionary analysis for septin proteins in metazoan was carried out. First, we demonstrated that all septin proteins in metazoan could be clustered into four subgroups, and the representative homologue of every subgroup was presented in the non-vertebrate chordate *Ciona intestinalis*, indicating that the emergence of the four septin subgroups should have occurred prior to divergence of vertebrates and invertebrates, and the expansion of the septin gene number in vertebrates was mainly by the duplication of pre-existing genes rather than by the appearance of new septin subgroup. Second, the direct orthologues of most human septins existed in zebrafish, which suggested that human septin gene repertoire was mainly formed by as far as before the split between fishes and land vertebrates. Third, we found that the evolutionary rate within septin family in mammalian lineage varies significantly, human SEPT1, SEPT 10, SEPT 12, and SEPT 14 displayed a relative elevated evolutionary rate compared with other septin members. Our data will provide new insights for the further function study of this protein family.**


*Keywords:* Septin; Phylogenetic analysis; Evolution rate; Tissue expression profile

## 1. Introduction

The septin family was originally discovered through genetic screening for budding yeast mutants defective in the cell-cycle progression [1,2]. The septins are a family of conserved GTP-binding proteins that act as dynamic regulatable scaffolds for recruitment of other proteins [3–5]. In humans, growing evidence has linked abnormalities in septin expression to cancer and a variety of other diseases [6]. The septin proteins possess a conserved GTP-binding domain and they fall into the large superclass of P-loop GTPases [7]. All septins have a P-loop [8,9] which is defined by the Walker A motif (GxxGxGKST), Walker B motif (DxxG), and the GTP-specificity motif (xKxD). Septins differ from most other members of this subclass in that they lack an asparagine at position 1 of the GTP-specificity motif (i.e. not NKxD, but AKAD or GKAD) [7]. Recently, the crystal structures of the human SEPT2 G domain and the heterotrimeric human SEPT2-SEPT6-SEPT7 complex had been determined. The septin complex structure revealed a universal bipolar polymer building block, composed of an extended G domain, which forms oligomers and filaments by conserved interactions between adjacent nucleotide-binding sites and/or the amino- and carboxy-terminal extensions [10]. Septins are found in all animals and fungi but not in plants [4,7]. It was found that human septins have been implicated in many cellular processes including cytokinesis [11–13], membrane associations, cell movement [14,15], vesicle trafficking [16], and also microtubule and actin function [17–19]. It is worthy to note that considerable diversity has been generated within each species; for example, other human septins are 39–63% identical to human SEPT2 at the amino acid level [4]. Based on the amino acid sequence similarity, human septins had been classified into four subgroups [4,20,21]. However, the previous phylogenetic analyses of the septin protein family usually only covered the septin sequences from human (*Homo sapiens*), nematode (*Caenorhabditis elegans*), fruit fly (*Drosophila melanogaster*) and yeast [4,20,21]. Recently, Pan and colleagues [22] reported the phylogenetic analysis of septin family across kingdoms (fungi, microsporidia, and animals). Based on their analysis, all septins from fungi, microsporidia, and animals were clustered into five groups: Group 1 and Group 2 contain septin sequences from fungi and animals, Group 3 and Group 4 contain septin sequences from fungi and microsporidia, and Group 5 contains septin sequences from filamentous fungi. However, much attention of their analysis was paid on septins from fungi, microsporidia, thus their analysis did not cover any septin sequences from invertebrate deuterostomes, and also did not cover enough septin sequences from vertebrates such as chicken (*Gallus gallus*), frog (*Xenopus laevis*), and zebrafish (*Danio rario*). For example, they only collected 6 septin sequences from zebrafish and 2 septin sequences from frog collected (in contrast, we identified 17 septin sequences from zebrafish, 9 septin sequences from frog). Therefore, their data were not enough to describe the detailed information about the evolutionary relationship of septin family in animal.

In this study, we carefully identified 78 septin proteins from 9 metazoan organisms (human (*H. sapiens*), mouse (*Mus musculus*), chicken (*G. gallus*), frog (*X. laevis*), zebrafish (*D. rario*),

*Corresponding author. Fax: +86 21 65643250.
*E-mail address:* longyu@fudan.edu.cn (L. Yu).

[1]These authors contributed equally to this work.

*Abbreviations:* NJ, neighbor-joining; ML, maximum likelihood; *Hs*, *Homo sapiens*; *Mm*, Mus musculus; *Gg*, *Gallus gallus*; *Xl*, *Xenopus laevis*; *Dr*, *Danio rario*; *Ci*, *Ciona intestinalis*; *Sp*, *Strongylocentrotus purpuratus*; *Ce*, *Caenorhabditis elegans*; *Dm*, *Drosophila melanogaster*

0014-5793/$32.00 © 2007 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2007.10.032

ascidian (*Ciona intestinalis*), sea urchin (*Strongylocentrotus purpuratus*), nematode (*Caenorhabditis elegans*), and fruit fly (*Drosophila melanogaster*)) and performed detailed phylogenetic and evolutionary analysis. The echinoderms deuterostome sea urchin (*S. purpuratus*) [23] and invertebrate urochordate ascidian *C. intestinalis* [24] were selected for analysis, and this is because they all occupy an important evolutionary position with respect to vertebrates. The echinoderms sea urchin [23] and their sister group hemichordates are the only other deuterostome animals beside the chordates. The sea urchin is thus more closely related to humans than the other major invertebrate models *D. melanogaster* and *C. elegans*. The ascidian *C. intestinalis* [24], a urochordate and one of the closest invertebrate relatives of vertebrates, provides a unique opportunity to gain insight into the complete set of septins available in chordates prior to the large-scale or whole genome duplication events that many believe to be associated with the early stages of vertebrate evolution [25,26]. The zebrafish (*D. rerio*) genome has also been selected since comparative studies provide insight into the likely timing of duplications occurring during vertebrate evolution. Gene duplications shared by fish and man are likely to have occurred prior to the tetrapod/teleost divergence [26,27].

Proteins, even in the same protein family, may have different evolutionary rates, suggesting that they are subject to various degrees of selection pressure [28]. Thus, we used three substitution rates to compare the evolution of the septin genes in mammalian lineages: $K_a$, $K_s$, and $K_a/K_s$. $K_a$ is the substitution rate per non-synonymous site which measures protein evolution, while $K_s$ is the substitution rate per synonymous site for DNA mutations which do not affect protein sequence. $K_a/K_s$ is the ratio of the non-synonymous and synonymous substitution rates, which measures the rate of protein evolution relative to the mutation rate and is a useful indicator of selection pressures [29].

Our detailed evolutionary analysis of septin family in metazoan will provide new insights on our understanding about the evolution of septin family.

## 2. Materials and methods

### 2.1. Database searches

Using the sequence of the human SEPT2 as a query, PSI-Blast [30] search was carried out at the National Center for Biotechnology Information (NCBI) Web site (http://www.ncbi.nlm.nih.gov/) to screen the non-redundant protein database from the various organisms, default values for PSI-Blast searches were used, with an E-value cutoff of 1e−40. The complete sequence of the human SEPT2 was also used to probe the genome of *C. intestinalis* using TBLASTN at the TIGR Gene Indices database [31] (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi).

### 2.2. Sequence alignments and phylogenetic analysis

Preliminary multiple sequence alignments were carried out by using CLUSTALX 1.8 [32]. Phylogenetic trees based on neighbor-joining (NJ) were generated for septin sequences using programs PROTDIST (JTT setting) [33] and NEIGHBOR (NJ setting) in PHYLIP (Version 3.63) [34]. Bootstrap replicates (1000) were generated with SEQBOOT (PHYLIP) and evaluated with NJ analysis in combination with the CONSENSE program (PHYLIP) for obtaining a majority rule consensus tree. Maximum likelihood (ML) trees were constructed by using PHYML V.2.4 [35], with 1000 bootstrap resamplings, JTT setting, and Gamma parameter values were estimated by using PHYML software. We did not use maximum-parsimony method because this method tends to yield unreliable results when highly divergent sequences were included. Tree files were viewed by using MEGA [36]. NJ tree and ML tree are shown with bootstrap values.

### 2.3. Evolution rate analysis

For calculation of mouse/human $K_a/K_s$ ratios, orthologous amino acid sequences were aligned using CLUSTAL W, and the alignment obtained was transferred to the cDNA sequences. $K_a$ and $K_s$ substitution rates were then calculated using methods of Nei and Gojobori [37] implemented in Dnasp package [38].

### 2.4. Serial analysis of gene expression

We collected more than 7,489,572 million experimental tags from 88 publicly available human serial analyses of gene expression (SAGE) libraries retrieved from the SAGE Genie repository [39]. All SAGE and tag-to-gene mapping information from SAGE Genie were parsed. SAGE septin gene tags were identified using the best tag information provided by SAGE Genie. Only tags found at least twice in libraries were considered. Tags corresponding to each septin member were counted from SAGE libraries that were derived from 23 normal tissues.

## 3. Results and discussion

### 3.1. Septin sequences identification

Using the human SEPT2 sequence as a query, we performed PSI-Blast searches at the National Center for Biotechnology Information (NCBI) Web site (http://www.ncbi.nlm.nih.gov/) to screen the non-redundant protein database from nine organisms (*H. sapiens*, *M. musculus*, *G. gallus*, *X. laevis*, *D. rario*, *C. intestinalis*, *S. purpuratus*, *C. elegans*, and *D. melanogaster*). The sequence of the human SEPT2 was also used to probe the genome of *C. intestinalis* using TBLASTN at the TIGR Gene Indices database [31] (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). Only full-length or nearly full-length protein sequences which contain almost the entire septin GTP domain were selected for analysis. Some retrieved sequences were discarded on the basis of the following criteria: (1) partial sequences or sequences resulting from frameshifts in the underlying mRNA as a result of cloning artifacts or possibly aberrant alternative splicing; (2) duplicated database submissions of the same sequence; and (3) alternatively spliced isoforms. At last we identified total 78 septin sequences from 9 organisms. 14 sequences from human (*H. sapiens*), 13 from *mouse* (*M. musculus*), 10 from chicken (*G. gallus*), 9 from frog (*X. laevis*), 17 from zebrafish (*D. rario*), 4 from sea urchin (*S. purpuratus*), 4 from Ciona (*C. intestinalis*), 2 from nematode (*C. elegans*), and 5 from fruit fly (*D. melanogaster*).

### 3.2. Evolutionary relationship of the septin protein family in metazoan

Previously, human septins are classified into four subgroups [4,20,21]: SEPT3 subgroup (SEPT3, SEPT9, and SEPT 12), SEPT2 subgroup (SEPT2, SEPT1, SEPT4, and SEPT 5), SEPT7 subgroup (SEPT7), and SEPT6 subgroup (SEPT 6, SEPT8, SEPT10, and SEPT11). This homology-based classification agreed well with their compatibility in the recombinant complex formation tested in insect cells. (1) A SEPTx/6/7 complexes of ~1:1:1 stoichiometry ($x$ = 1, 2, 4, or 5) can form. (2) SEPT6 in the SEPT2/6/7 complex is replaceable with SEPT8, SEPT10, and SEPT11 [20,21,40]. Human septins of SEPT2 subgroup, SEPT7 subgroup and SEPT6 subgroup all were predicted to have a coil-coil domain in the C-terminal [4,40]. However, SEPT3 subgroup (SEPT3, SEPT9 and SEPT12)

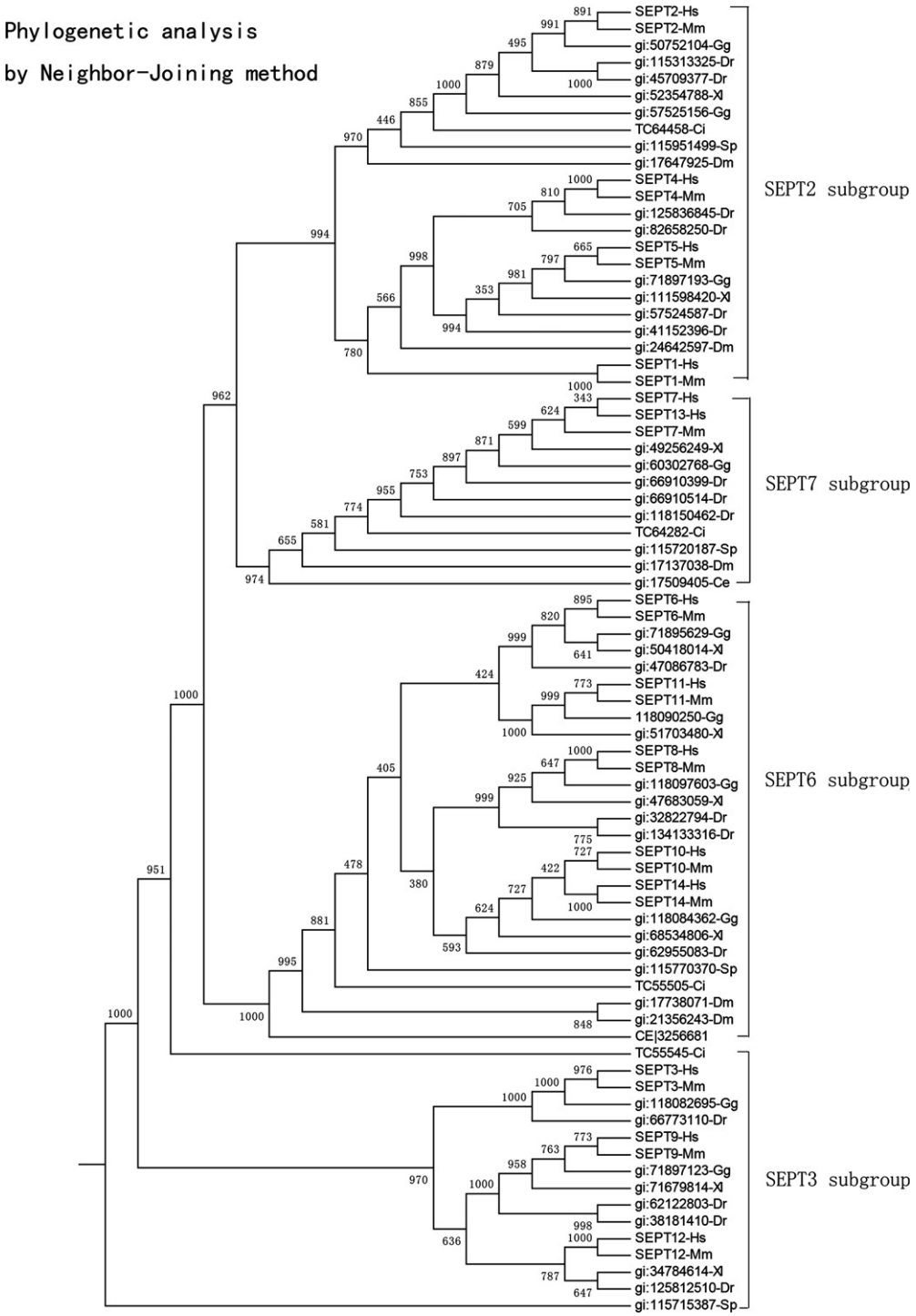


Fig. 1. Phylogenetic relationships of septin family in metazoan. The septin family was separated into four subgroups (SEPT2 subgroup, SEPT3 subgroup, SEPT7 subgroup, and SEPT6 subgroup), based upon the clades produced and phylogenetic analyses performed. The trees shown were inferred by neighbor-joining from a gapped alignment. The values on the tree nodes are neighbor-joining bootstrap. The gi number of Human and mouse septin sequences was indicated below: SEPT1-Hs, gi|20178107|; SEPT2 -Hs, gi|2500769|; SEPT3-Hs, gi|13124527|; SEPT4-Hs, gi|3287733|; SEPT5-Hs, gi|6685760|; SEPT6-Hs, gi|20178343|; SEPT7-Hs, gi|67472677|; SEPT8-Hs, gi|45645200|; SEPT9-Hs, gi|93141311|; SEPT10-Hs, gi|21945064|; SEPT11-Hs, gi|8922712|; SEPT12-Hs,gi|21389409|; SEPT13-Hs, gi|113418512|; SEPT14-Hs, gi|153252198|; SEPT1-Mm, gi|1169343|; SEPT2-Mm, gi|1346679|; SEPT3-Mm, gi|117606375|; SEPT4-Mm, gi|114978|; SEPT 5-Mm, gi|83305642|; SEPT6-Mm, gi|20178348|; SEPT7-Mm, gi|28173550|; SEPT8-Mm, gi|45477305|; SEPT9-Mm, gi|56749655|; SEPT10-Mm, gi|67461572|; SEPT11-Mm, gi|57634518|; SEPT12-Mm, gi|21312734|; SEPT14-Mm, gi|149254604|. Other septin sequences were indicated their gi accession numbers followed by their organism abbreviations. The organism abbreviations used were listed below: *Hs*, *Homo sapiens*; *Mm*, *Mus musculus*; *Gg*, *Gallus gallus*; *Xl*, *Xenopus laevis*; *Dr*, *Danio rario*; *Ci*, *Ciona intestinalis*; *Sp*, *Strongylocentrotus purpuratus*; *Ce*, *Caenorhabditis elegans*; *Dm*, *Drosophila melanogaster*. The sequences from *C. intestinalis* was identified at the TIGR Gene Indices database [31] (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). and did not have genbank accession number (The amino acid sequences of septin sequences from *C. intestinalis* can be found in additional file 1). All of the amino acid sequences and accession number of sequences we analyzed can be obtained from additional file 1.

was predicted to lack this coiled-coil region, which is unlike the other septins [40]. It was reported that SEPT9 isoforms (SEPT3 subgroup) associate with other septins through their N-terminal domain [41].

Owing to limited genome sequences available, the previous phylogenetic analysis of the septin protein family did not cover any septin sequences from invertebrate deuterostomes, and also did not cover enough septin sequences from vertebrates such as chicken (*G. gallus*), frog (*X. laevis*), and zebrafish (*D. rario*) [20–22], To get a better evolutionary history of septin family in metazoan, especially the evolutionary history of septin proteins from invertebrates to vertebrates, we identified 78 septin proteins from 9 metazoan organisms.

The 78 septein protein sequences were then aligned with ClustalX version 1.81, using BLOSUM 30 as the protein weight matrix. Because the N-terminus and C-terminus of septin sequences are poorly aligned, only multi-alignment of the GTP domain of septin sequences was selected for the phylogenetic analysis. Phylogenetic trees based on neighbor-joining (NJ) were generated by PHYLIP (Version 3.63) [34]. And maximum likelihood (ML) trees were constructed using PHYML V.2.4 [35].

Neighbor-joining (NJ) phylogenetic analysis with high supporting bootstrap values revealed that all metazoan septin proteins can be clustered into four subgroups, which was well in agreement with previous four subgroup classification of human septins (Fig. 1). Maximum likelihood (ML) phylogenetic analysis also gave out the same subgroup classification for all metazoan septin proteins analyzed (additional file 2). Neighbor-joining (NJ) and maximum likelihood (ML) phylogenetic approaches generated almost identical overall tree topologies, inferring our evolutionary analysis is reliable. As mentioned before, both the echinoderms deuterostome sea urchin (*S. purpuratus*) [23] and invertebrate urochordate ascidian *C. intestinalis* [24] occupy important evolutionary positions with respect to vertebrates.

We found that representative homologue of every subgroup can be found in *C. intestinalis* and *S. purpuratus* (Fig. 1). This indicated that the emergence of the four septin subgroups can be dated back to before the split of vertebrate and invertebrate. It has been also proposed that large-scale gene duplications occurred after the vertebrates diverged from the cephalochordates and urochordates [42]. Therefore, we speculated that the expansion of the septin gene number in vertebrates has occurred by the duplication of pre-existing genes rather than by the appearance of new septin subgroup, though it is also possible that septin genes evolved earlier in chordate evolution but the orthologues were subsequently ablated in urochordates *C. intestinalis* and invertebrate deuterostomes sea urchin (*S. purpuratus*). Second, our phylogenetic analysis revealed that most human septin protein (except human SEPT1, SEPT 13, and SEPT 14) can found their distinct orthologues in fish. And it was speculated that human septin gene repertoire was mainly formed by as far as before the split between fishes and land vertebrates. It is interesting that there are 17 distinct septin proteins in zebrafish, but only 14 septin proteins in human. The increased number of septin paralogues in the zebrafish genome compared to human genome is what would be predicted if a teleost-specific genome duplication had occurred after their divergence from the tetrapod lineage [43].

The evolutionary relationship within every subgroup was also revealed by our phylogenetic analysis. The SEPT7 subgroup includes human SEPT7 and human SEPT13 (gi|113418512|). Human SEPT13 was first predicted by Dr. Hall [40]. The representative homologue of human SEPT7 subgroup presented in protostomes (*D. melanogaster* (gi: 17137038) and *C. elegans* (gi: 17509405)), invertebrate deuterostomes (*C. intestinalis* (tc64282) and *S. purpuratus* (gi: 115720187)) and all vertebrates were investigated. Interestingly, the direct homologue of human SEPT13 was not found in *M. musculus* and *G. gallus*.

SETP2 subgroup included human SETP1, SETP2, SETP4, SETP5, and the representative homologue of this subgroup can be found in *D. melanogaster* (gi: 17647925 and gi: 24642597), but not in *C. elegans.* It is worthy to mention that, the direct homologue of SEPT2 can be found in *S. purpuratus* (gi: 115951499), *C. intestinalis* (TC64458), and all vertebrates, while the direct orthologue of human SEPT1 can only be found in mammals, but not in other vertebrates (such as *G. gallus*, *X. laevis*, and *D. rario*). It is not sure if the emergence of SEPT1 is later than other human SETP2 subgroup members or the orthologue of SEPT1 in other vertebrates is lost during evolution.

SEPT6 subgroup includes human SEPT6, 8, 10, 11, 14, and the representative homologue of this subgroup can be found in the protostomes (*D. melanogaster* (gi: 17738071, gi: 21356243) and *C. elegans* (gi: 32566810)) and invertebrate deuterostomes (*C. intestinalis* (TC55505) and *S. purpuratus* (gi: 115770370)). The direct homologue of human SEPT14 (gi|153252198|) was only found in *M. musculus* (gi|149254604|), but not in *G. gallus* and *X. laevis*. Human SEPT 14 was recently reported to share closest homology to human SEPT 10 [44], our phylogenetic analysis also indicated this.

SEPT3 subgroup, which includes human SEPT3, SEPT9, and SEPT12, is different with other subgroups. First, all members of SEPT3 subgroup did not have the coil-coil domain. Second, the representative homologue of SEPT3 subgroup did not present in the protostomes, but only present in *C. intestinalis* (TC55545) and *S. purpuratus* (gi: 115715387). So it was speculated that the emergence of SEPT3 subgroup was later than the other three septin subgroups.

### 3.3. The evolutionary rate analysis of the septin family

For protein coding sequences, the synonymous rate ($K_s$) is often regarded as a measure of the underlying mutation rate [45], though it may be influenced by other factors [46]. By contrast, the non-synonymous rate ($K_a$) or the ratio $K_a/K_s$ (which corrects for variation in $K_s$ among proteins) is often regarded

Table 1
Synonymous ($K_s$) and non-synonymous ($K_a$) nucleotide substitution rates for GTPase domains of human and mouse septins

| | SEPT1 | SEPT2 | SEPT3 | SEPT4 | SEPT5 | SEPT6 | SEPT7 | SEPT8 | SEPT9 | SEPT10 | SEPT1 | SEPT12 | SEPT14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $K_a$ | 0.048 | 0.005 | 0.000 | 0.004 | 0.005 | 0.016 | 0.002 | 0.005 | 0.008 | 0.058 | 0.000 | 0.116 | 0.102 |
| $K_s$ | 0.419 | 0.454 | 0.311 | 0.354 | 0.470 | 0.35 | 0.199 | 0.355 | 0.432 | 0.366 | 0.469 | 0.455 | 0.527 |
| $K_a/K_s$ | 0.114 | 0.0180 | 0.0 | 0.01 | 0.0180 | 0.049 | 0.008 | 0.013 | 0.018 | 0. 157 | 0.000 | 0.255 | 0.193 |

Table 2
SAGE analysis of septin mRNA expression in human tissues

| Human tissue | SEPT1 | SEPT2 | SEPT3 | SEPT4 | SEPT5 | SEPT6 | SEPT7 | SEPT8 | SEPT9 | SEPT10 | SEPT1 | SEPT12 | SEPT13 | SEPT14 | Total tags |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Brain | | 119 | 71 | 382 | 119 | 22 | 274 | 11 | 80 | 13 | 22 | | | 2 | 463420 |
| Retina | | 26 | 16 | 473 | 43 | 60 | 148 | 50 | 12 | 2 | 6 | | 9 | | 820326 |
| Thyroid | | 129 | | 26 | | | 129 | 26 | 17 | 9 | 9 | | | 2 | 115938 |
| Lung | | 69 | | 6 | | 18 | 263 | 13 | 119 | 19 | 44 | | | | 159917 |
| Heart | | 36 | | 12 | 36 | | 157 | 12 | 12 | | 12 | | | | 83063 |
| Breast | 2 | 100 | | 196 | 10 | 4 | 72 | 4 | 103 | 12 | 18 | | | 4 | 509596 |
| Stomach | | 94 | | 18 | | | 76 | 12 | 12 | | 36 | | | | 170675 |
| Pancreas | | 140 | | 23 | | | 93 | 12 | 36 | 12 | | | | | 85728 |
| Liver | | 45 | | | | | 30 | 15 | 15 | 15 | 30 | | | | 66308 |
| Kidney | | 187 | | | | | 94 | 19 | 56 | | 47 | | | | 106476 |
| Colon | | 30 | | | 10 | | 30 | | 173 | 10 | | | | | 98089 |
| Peritoneum | | 186 | | | | | 95 | | 76 | 19 | 76 | | | | 53527 |
| Spinal Cord | 18 | 274 | | 292 | | | 310 | | 54 | 36 | 36 | | 13 | | 54785 |
| Ovary | | 252 | | 21 | 11 | 11 | 221 | | 126 | 21 | 32 | | | | 94887 |
| Placenta | | 35 | | 5 | | 15 | 130 | 20 | 20 | | 50 | | | | 207348 |
| Prostate | 4 | 86 | | | | | 56 | 11 | 34 | 15 | | | | | 266949 |
| Bone Marrow | 15 | 83 | | 5 | 53 | | 73 | | 15 | | | | | | 204563 |
| Muscle | 9 | 9 | | 18 | | | 9 | | | 9 | 9 | | | | 107836 |
| Skin | | 27 | | | 27 | | 55 | | 136 | | 27 | | | | 36656 |
| Lymph node | | 111 | | | | 21 | 82 | | 60 | 30 | 10 | 50 | 10 | 20 | 99426 |
| White blood cells | 70 | 115 | | 1 | 50 | 8 | 27 | 4 | 128 | 7 | 6 | | | 3 | 846268 |
| Embryonic stem cell | 3 | 22 | 10 | 3 | 1 | 5 | 90 | 11 | 81 | 60 | 36 | | | 2 | 2594061 |
| Vascular | 4 | | 21 | 48 | 28 | 8 | 131 | 12 | 94 | 24 | 49 | | 4 | 8 | 243730 |

*Note*: mRNA is expressed as positive tags per million sequenced tags (total tags). Only tags found at least twice were considered. Unfilled cells indicate too low a level to be estimated.

either as a measure of the amount of purifying selection on the protein or as a measure of the amount of positive selection. For most genes, non-synonymous rates are lower than synonymous rates and are much more variable from gene to gene, and this is thought to reflect differences in the extent of selective constraints and purifying selection among proteins [47,48].

We investigated the evolutionary rate of septin family in mammalian lineage. Because the GTP-binding domain of septin family proteins was most conserved and the lengths of N-terminal and C-terminal of septin proteins vary greatly, we investigated $K_a/K_s$ data for GTP-binding domain of the septin family in human and mouse (Table 1). The values of $K_a/K_s$ for all septins were below 1, so all septin proteins was under negative selective purification pressure. It was found that the $K_a$ values of most septins, especially SEPT3, SEPT7, and SEPT11 were very low, indicating these septins all evolved very slowly. Moreover, the values of $K_a/K_s$ of these septins also were very low, indicating that they all were under very high selective purification pressure. However, the $K_a$ and $K_a/K_s$ values for SEPT1, SEPT 10, SEPT 12, and SEPT14 were a little higher, and this indicates that SEPT1, SEPT10, SEPT12, and SEPT14 display relative evaluated evolutionary rates compared to other septin members. Why the members of septin family have different evolutionary rate and are subject to various degrees of selection pressure will be interesting for further study.

### *3.4. Tissue expression profile of human septin family*

A few studies have shown that broadly expressed gene tend to evolve more slowly than narrowly expressed genes [49]. To compare the ontogeny of the septin family with physiological functions, we wished to examine the tissue distribution of each member. To this aim, we collected SAGE data from normal human tissues. The SAGE method, developed for quantitative analysis of expressed genes [50], has been widely used to compare mRNA distribution in different tissues [51]. Corresponding to each septin member, we counted its expression from SAGE libraries derived from 34 tissues. Unique tags were not considered. For each tissue, results are expressed as tags per million (Table 2). Our results were in general in agreement with the septin gene expression profile assayed by DNA microarray methods [39]. By SAGE assay, we found that SEPT2, SEPT7, and SEPT9 which were under high selective purification pressure, appeared as the most ubiquitously expressed human septin members, and detected in 80–100% of examined libraries. However, SEPT1 mostly expressed in hemopoietic tissues, and the expression levels of human SEPT12, SEPT13, and SEPT14 were very low. As noted before, SEPT1, SEPT12, and SEPT14 displayed elevated evolution rates.

## 4. Conclusions

We for the first time demonstrated that all septin proteins in metazoan could be clustered into four subgroups, and the emergence of the four septin subgroups should have occurred prior to the divergence of vertebrates and invertebrates, and the expansion of the septin gene number in vertebrates was by the duplication of pre-existing genes in early vertebrate evolution. We also proved that septin gene repertoire in human was mainly formed before the split between fishes and land vertebrates. The evolutionary rates of septin family in mammalian lineage and the tissue expression profile of human septin were also investigated. Our systematic phylogenetic and evolutionary analysis of septin family in metazoans should provide new insights for further function study of this protein family.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2007.10.032.

## References

[1] Hartwell, L.H. (1971) Genetic control of the cell division cycle in yeast. IV. Genes controlling bud emergence and cytokinesis. Exp. Cell Res. 69, 265–276.
[2] Byers, B. and Goetsch, L. (1976) A highly ordered ring of membrane-associated filaments in budding yeast. J. Cell Biol. 69, 717–721.
[3] Field, C.M. and Kellogg, D. (1999) Septins: cytoskeletal polymers or signalling GTPases? Trends Cell Biol. 9, 387–394.
[4] Kinoshita, M. (2003) The septins. Genome Biol. 4, 236.
[5] Kinoshita, M. (2006) Diversity of septin scaffolds. Curr. Opin. Cell Biol. 18, 54–60.
[6] Hall, P.A. and Russell, S.E. (2004) The pathobiology of the septin gene family. J. Pathol. 204, 489–505.
[7] Leipe, D.D., Wolf, Y.I., Koonin, E.V. and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. J. Mol. Biol. 317, 41–72.
[8] Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) The P-loop – a common motif in ATP-and GTP-binding proteins. Trends Biochem. Sci. 15, 430–434.
[9] Bourne, H.R., Sanders, D.A. and McCormick, F. (1991) The GTPase superfamily: conserved structure and molecular mechanism. Nature 349, 117–127.
[10] Sirajuddin, M., Farkasovsky, M., Hauer, F., Kuhlmann, D., Macara, I.G., Weyand, M., Stark, H. and Wittinghofer, A. (2007) Structural insight into filament formation by mammalian septins. Nature 449, 311–315.
[11] Kinoshita, M., Kumar, S., Mizoguchi, A., Ide, C., Kinoshita, A., Haraguchi, T., Hiraoka, Y. and Noda, M. (1997) Nedd5, a mammalian septin, is a novel cytoskeletal component interacting with actin-based structures. Genes Dev. 11, 1535–1547.
[12] Spiliotis, E.T., Kinoshita, M. and Nelson, W.J. (2005) A mitotic septin scaffold required for Mammalian chromosome congression and segregation. Science 307, 1781–1785.
[13] Nguyen, T.Q., Sawa, H., Okano, H. and White, J.G. (2000) The *C. elegans* septin genes, unc-59 and unc-61, are required for normal postembryonic cytokineses and morphogenesis but have no essential function in embryogenesis. J. Cell Sci. 113 (Pt 21), 3825–3837.
[14] Finger, F.P., Kopish, K.R. and White, J.G. (2003) A role for septins in cellular and axonal migration in *C. elegans*. Dev. Biol. 261, 220–234.
[15] Hsu, S.C., Hazuka, C.D., Roth, R., Foletti, D.L., Heuser, J. and Scheller, R.H. (1998) Subunit composition, protein interactions, and structures of the mammalian brain sec6/8 complex and septin filaments. Neuron 20, 1111–1122.
[16] Surka, M.C., Tsang, C.W. and Trimble, W.S. (2002) The mammalian septin MSF localizes with microtubules and is required for completion of cytokinesis. Mol. Biol. Cell 13, 3532–3545.
[17] Nagata, K., Kawajiri, A., Matsui, S., Takagishi, M., Shiromizu, T., Saitoh, N., Izawa, I., Kiyono, T., Itoh, T.J., Hotani, H. and Inagaki, M. (2003) Filament formation of MSF-A, a mammalian septin, in human mammary epithelial cells depends on interactions with microtubules. J. Biol. Chem. 278, 18538–18543.
[18] Finger, F.P. (2002) One ring to bind them. Septins and actin assembly. Dev. Cell 3, 761–763.
[19] Kartmann, B. and Roth, D. (2001) Novel roles for mammalian septins: from vesicle trafficking to oncogenesis. J. Cell Sci. 114, 839–844.

[20] Kinoshita, M. (2003) Assembly of mammalian septins. J. Biochem. (Tokyo) 134, 491–496.
[21] Spiliotis, E.T. and Nelson, W.J. (2006) Here come the septins: novel polymers that coordinate intracellular functions and organization. J. Cell Sci. 119, 4–10.
[22] Pan, F., Malmberg, R.L. and Momany, M. (2007) Analysis of septins across kingdoms reveals orthology and new motifs. BMC, Evol. Biol. 7, 103.
[23] Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R.C., Angerer, L.M., Arnone, M.I., Burgess, D.R., et al. (2006) The genome of the sea urchin *Strongylocentrotus purpuratus*. Science 314, 941–952.
[24] Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., et al. (2002) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science 298, 2157–2167.
[25] McLysaght, A., Hokamp, K. and Wolfe, K.H. (2002) Extensive genomic duplication during early chordate evolution. Nat. Genet. 31, 200–204.
[26] Holland, P.W. (1999) Gene duplication: past, present and future. Semin. Cell Dev. Biol. 10, 541–547.
[27] Meyer, A. and Van de, P.Y. (2005) From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays 27, 937–945.
[28] Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562.
[29] Yang, Z. and Bielawski, J.P. (2000) Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. 15, 496–503.
[30] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
[31] Lee, Y., Tsai, J., Sunkara, S., Karamycheva, S., Pertea, G., Sultana, R., Antonescu, V., Chan, A., Cheung, F. and Quackenbush, J. (2005) The TIGR Gene Indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. Nucleic Acids Res. 33, D71–D74.
[32] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
[33] Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) A new approach to protein fold recognition. Nature 358, 86–89.
[34] Felsenstein, J. (1989) PHYLIP – phylogeny inference package (Version 3.2). Cladistics 5, 164–166.
[35] Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
[36] Kumar, S., Tamura, K. and Nei, M. (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. 5, 150–163.
[37] Nei, M. and Gojobori, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.
[38] Rozas, J. and Rozas, R. (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175.
[39] Boon, K., Osorio, E.C., Greenhut, S.F., Schaefer, C.F., Shoemaker, J., Polyak, K., Morin, P.J., Buetow, K.H., Strausberg, R.L., De Souza, S.J. and Riggins, G.J. (2002) An anatomy of normal and malignant gene expression. Proc. Natl. Acad. Sci. USA 99, 11287–11292.
[40] Hall, P.A., Jung, K., Hillan, K.J. and Russell, S.E. (2005) Expression profiling the human septin gene family. J. Pathol. 206, 269–278.
[41] Nagata, K., Asano, T., Nozawa, Y. and Inagaki, M. (2004) Biochemical and cell biological analyses of a mammalian septin complex, Sept7/9b/11. J. Biol. Chem. 279, 55895–55904.
[42] Holland, P.W., Garcia-Fernandez, J., Williams, N.A. and Sidow, A. (1994) Gene duplications and the origins of vertebrate development. Dev. Suppl., 125–133.
[43] Brunet, F.G., Crollius, H.R., Paris, M., Aury, J.M., Gibert, P., Jaillon, O., Laudet, V. and Robinson-Rechavi, M. (2006) Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. Mol. Biol. Evol. 23, 1808–1816.
[44] Peterson, E.A., Kalikin, L.M., Steels, J.D., Estey, M.P., Trimble, W.S. and Petty, E.M. (2007). Characterization of a SEPT9 interacting protein, SEPT14, a novel testis-specific septin. Mamm. Genome 2007 Oct 6 (Epub ahead of print).
[45] Miyata, T., Yasunaga, T. and Nishida, T. (1980) Nucleotide sequence divergence and functional constraint in mRNA evolution. Proc. Natl. Acad. Sci. USA 77, 7328–7332.
[46] Williams, E.J. and Hurst, L.D. (2002) Is the synonymous substitution rate in mammals gene-specific? Mol. Biol. Evol. 19, 1395–1398.
[47] Kimura, M. (1983) The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge.
[48] Graur, D. and Li, W.-H. (2000) Fundamentals of Molecular Evolution. Sinauer Associates, Sunderland, MA.
[49] Duret, L. and Mouchiroud, D. (2000) Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. Mol. Biol. Evol. 17, 68–74.
[50] Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995) Serial analysis of gene expression. Science 270, 484–487.
[51] Harbers, M. and Carninci, P. (2005) Tag-based approaches for transcriptome research and genome annotation. Nat. Meth. 2, 495–502.