Molecular Phylogenetics and Evolution 73 (2014) 77–86

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Evolutionary dynamics of Polynucelotide phosphorylases

Upneet K. Sokhi [a,c], Rob DeSalle [b,*], Manny D. Bacolod [a,c], Swadesh K. Das [a,c], Santanu Dasgupta [a,c], Devanand Sarkar [a,c,d], Paul B. Fisher [a,c,d,*]

[a] Department of Human and Molecular Genetics, Virginia Commonwealth University, School of Medicine, Richmond, VA, United States
[b] American Museum of Natural History, New York University, New York, NY, United States
[c] VCU Institute of Molecular Medicine, Virginia Commonwealth University, School of Medicine, Richmond, VA, United States
[d] VCU Massey Cancer Center, Virginia Commonwealth University, School of Medicine, Richmond, VA, United States

ARTICLE INFO

Article history:
Received 24 August 2013
Revised 9 January 2014
Accepted 21 January 2014
Available online 3 February 2014

Keywords:
Positive selection
Purifying selection
d*N*/d*S* ratio

ABSTRACT

Polynucleotide phosphorylase (PNPase) is an evolutionarily conserved 3′ → 5′ phosphate-dependent exoribonucease belonging to the PDX family of proteins. It consists of two catalytic RNase PH domains (PNP1 and PNP2), an α-helical domain and two RNA-binding domains. The PNP1 and PNP2 domains share substantial sequence and structural homology with RNase PH (RPH), which is another PDX family member found in all the three major kingdoms of life, suggesting that these three domains originated from a common ancestor. Phylogenetic analysis (based on the PNPase/RNase PH sequence information for 43 vertebrate taxa) shows that PNP2 and RPH are sister taxa which arose through duplication of the ancestral PNP1 domain. Also, all three domains (PNP1, PNP2 and RPH), along with the KH and S1 domains have undergone significant and directional sequence change, as determined by branch and site-specific d*N*/d*S* analyses. In general, codons that show d*N*/d*S* ratios that are significantly greater than 1.0 are outside the ordered regions (α-helices and β-sheets) of these protein domains. In addition, sites that have been selected for mutagenesis in these proteins lie embedded in regions where there is a preponderance of codons with d*N*/d*S* values that are not significantly different from 0.0. Overall, this report is an attempt to further our understanding of the evolutionary history of these three protein domains, and define the evolutionary events that led to their refinement in the vertebrate lineage leading to mammals.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

RNA metabolism plays a significant role in the post-transcriptional control of gene expression in both prokaryotic and eukaryotic organisms (Arraiano et al., 2010). Key players involved in the decay and processing of various RNA species are a special class of enzymatic proteins called ribonucleases (RNases) (Deutscher and Li, 2001; Deutscher, 1993). Based on their degradative functions, RNases are divided into endo- and exo-ribonucleases (Sarkar and Fisher, 2006; Sokhi et al., 2013). One such exoribonuclease is polynucleotide phosphorylase (PNPase), which is an ancient enzyme that has been evolutionarily conserved in all the kingdoms of life, although it is absent in Archaea and certain single cell eukaryotes (Leszczyniecka et al., 2004). Functionally, it is a 3′ → 5′ phosphorolytic exoribonuclease belonging to the PDX family of proteins (Andrade et al., 2009; Ibrahim et al., 2008). The PDX family includes RNase PH, another phosphate-dependent exoribonuclease. Considered the simplest member of the PDX family, RNase PH was first identified in *E. coli* (Baginsky et al., 2001; Deutscher et al., 1988), and its orthologues are found in all kingdoms (Leszczyniecka et al., 2004). The main difference between RNase PH and PNPase is that *in vivo*, RNase PH is involved in the processing of tRNA precursors and in the digestion of structured RNA in bacteria (Deutscher et al., 1988; Jain, 2012), whereas PNPase mainly plays a role in mRNA degradation (Grunberg-Manago et al., 1955; Sarkar et al., 2003).

Sequence alignment shows that PNPases from various kingdoms share a classical domain structure consisting of: two N-terminal RNase PH domains (containing the PDX catalytic sites, designated PNP1 and PNP2), an α-helical domain between PNP1 and PNP2, and the RNA-binding domains KH (K-homology) and S1 located at the C-terminal (Fig. 1) (Chen et al., 2007; Raijmakers et al., 2002; Sarkar and Fisher, 2006). Both RNase PH and PNPase form ring-like structures (hexameric and homotrimeric respec-

* Corresponding authors at: Department of Human and Molecular Genetics, VCU Institute of Molecular Medicine, Virginia Commonwealth University, School of Medicine, 1101 East Marshall Street, Sanger Hall Building, Room 11-015, Richmond, VA 23298-0033, United States. Tel.: +1 804 628 3506 (O); fax: +1 804 827 1124 (P.B. Fisher).

E-mail addresses: desalle@amnh.org (R. DeSalle), pbfisher@vcu.edu (P.B. Fisher).

http://dx.doi.org/10.1016/j.ympev.2014.01.020
1055-7903/© 2014 Elsevier Inc. All rights reserved.

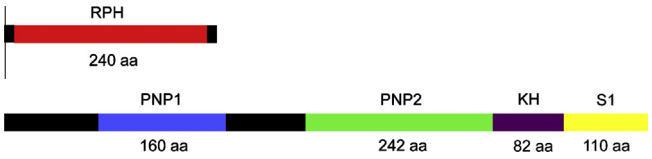

Fig. 1. Schematic diagrams of PNPase (old-35) and RPH showing the various domains in these proteins. The domain name is given above the diagram and the length of the domain (or interdomain) is given below the diagram.

tively) with a central channel in their active forms. The level of structural and sequence similarity between the PDX domains in PNPase and RNase PH raises the possibility that these domains may have originated from a common ancestor (in the most recent common ancestor of life; MRCA) via gene duplication events (Leszczyniecka et al., 2004; Sokhi et al., 2013).

Previously, our primary focus has been to analyze how the three PDX domains (PNP1, PNP2 and RPH) have evolved in the tree of life. However, recent structural and phylogenetic information on these domains has made possible the examination of the evolutionary dynamics of their function. One common tool that is used to examine functional significance of proteins over evolutionary time is to characterize the impact of natural selection on sequence change. The methodology for statistically assessing patterns of natural selection in a phylogenetic context on a residue by residue basis, as well as on a branch by branch basis in a phylogeny are well developed (dos Reis et al., 2012; Murphy et al., 2001; O'Leary et al., 2013; Pentony et al., 2012). They are based on corrected assessment of the ratio of nonsynonomous to synonymous substitutions over evolutionary time. The null hypothesis in such studies is the situation where the nonsynonymous to synonymous changes are equal in abundance. A statistically significant overabundance of nonsynonymous change relative to synonymous change is usually taken as an indication of positive natural selection. And a significant lack of nonsynonymous change to synonymous change has been taken as evidence for negative or purifying selection by researchers.

These approaches have been criticized on statistical grounds and on the grounds (Hughes, 2008) that statistical demonstration of selection is only the first step in understanding natural selection on proteins. In other words, once natural selection has been suggested by statistical methods it should also be tested in a functional or biochemical context (Dean and Thornton, 2007; Hughes, 2008).

Here, we use the well known functional and structural characteristics of the PNP and RPH domains to examine the validity of statistically determined measures of natural selection. Several sites in PNPases have been mutagenized and the functional significance of those mutations has been well characterized. In addition, sequence integrity is an invariable factor (Hughes, 2008), but recent advances in structural biology show that secondary structure is equally important in not only determining the functionality but also in dictating the rates of site-specific evolution of proteins (Nilsson et al., 2011). These residues under natural selection (positive or negative) may occur in ordered (α-helices and β-sheets) or disordered (random coils and turns) regions (Ridout et al., 2010). However, from our analyses of PNP1, PNP2 and RPH sequences from 43 eukaryotic taxa (Table S1), we can infer that the residues with dN/dS measures greater than 1 are almost entirely found outside of α-helices and β-sheets. Finally, through phylogenetic analysis, we have been able to discern the possible time course of evolution and the molecular events when PNPase (PNP1, PNP2, KH and S1) and RPH domains might have been under positive natural selection among vertebrates.

2. Materials and methods

2.1. Sequences and matrix construction

PNPase and RNase PH sequences were obtained from Genbank using a taxonomic sampling strategy to cover the eukaryotic tree of life. Only PNPase and RNase PH sequences from fully sequenced and accurately annotated eukaryotic genomes were used. A list of the taxa used in this study is given in Table S1. Codon sequences were obtained in all cases and accession numbers for the sequences are given in Table S1. Sequences were aligned in TranslatorX (Abascal et al., 2010), a program that aligns codons using amino acids as a guide. For the various analyses in this paper we generated several different phylogenetic matrices that were used in tree building and natural selection detection programs. The first kind of matrix we generated was a taxonomic based matrix with the 43 taxa we list in Table S1. This matrix consists of full-length PNP sequences concatenated with RNase PH sequences for the same 43 taxa in a NEXUS formatted matrix. The different regions of the PNPase genes and the RPH gene (Fig. 1) were then partitioned using NEXUS commands in the NEXUS matrix. The root of this matrix was established using the three plant taxa (Arabidopsis, Selaginella and Physcomytrella) as outgroups.

The second kind of matrix combined PNP1, PNP2 and RPH domains from all 43 taxa used in this study. This matrix was constructed using a separate alignment procedure from the taxonomic-based matrix. In addition two RNase II genes (one from human and the other from Trichoplax) were used to root trees for this latter matrix.

2.2. Phylogenetic analysis

Phylogenetic trees were constructed using parsimony, maximum likelihood and Bayesian approaches (Huelsenbeck and Ronquist, 2001). The model used was a GTR with invariants rates and gamma distribution (estimated from the data set). 100 bootstrap replicates were performed when bootstrapping (either MP or ML) (Stamatakis et al., 2008). Bayesian analysis was performed using 1 million generations of simulation and the model described above. For the selection analyses we used a predetermined tree topology based on the well-accepted taxonomy of the organisms in the study (O'Leary et al., 2013).

2.3. Detecting the statistical imprint of skewed dN/dS ratios

Two tests were used to detect statistically significant patterns of skewed d*N*/d*S* ratios. The first test determines the branches in the accepted phylogeny where statistically significant departure from neutral evolution occurs. The BREL (Branch-site Random Effects Likelihood) test in the HyPhy (Pond et al., 2005) package was used on the three domains for the taxon based matrices (PNP1, PNP2 and RPH) separately. In addition, the BREL option was also used to analyze branch specific natural selection in the matrix with all three-domain sequences as terminals. The default settings and the accepted tree topology were used in these tests (Delport et al., 2010; Kosakovsky Pond et al., 2011).

The second test implemented the MEME (Mixed Effects Model of Evolution) (Murrell et al., 2012) option in DataMonkey, the web version of HyPhy (Delport et al., 2010; Murrell et al., 2012). This option uses mixed model approaches to detect departure from neutrality at individual codons. This latter test was performed individually on each of the three domains – PNP1, PNP2 and RPH.

Download English Version:

https://daneshyari.com/en/article/5919516

Download Persian Version:

https://daneshyari.com/article/5919516

Daneshyari.com