

# Molecular Evolution of *PMS2* Gene Family in Higher Primates and Its Possible Biological Implications

**George V. Shpakovski**  
gvs@ibch.ru

**Elena K. Shematorova**  
elenashe@ibch.ru

**Dmitry G. Shpakovski**  
mati@mail.ru

Laboratory of Mechanisms of Gene Expression, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 16/10 Miklukho-Maklaya Str., Moscow 117997, Russia

## Abstract

We have performed phylogenetic analysis of *PMS2*-like cDNA and genomic sequences in the primate lineage and have shown that molecular evolution of *PMS2* protein, the essential component of the eukaryotic mismatch repair and gene conversion systems, correlates well with the route of evolution of the ape lineage leading to hominids. Our data suggest that a novel tripartite polypeptide system of *PMS2*-like proteins, which could replace the original *PMS2* protein in some of its novel or substantially diverged functions, have been established at least in *Homo sapiens* – with possible impact on such vitally important genetic processes as gene conversion, somatic hypermutation and class switch recombination.

**Keywords:** mismatch repair, gene conversion, *PMS2* gene family, Primates, molecular evolution

## 1 Introduction

Mismatch repair and gene conversion systems provide several genetic stabilization functions: they correct DNA biosynthetic errors, ensure the fidelity of genetic recombination, and participate in the earliest steps of checkpoint and apoptotic responses to several classes of DNA damage [1, 2]. In eukaryotes, these systems are much more complex than in prokaryotes, in particular, with respect to detection and elimination of mispaired DNA bases and repair of short insertions and deletions (two to about ten nucleotides) in microsatellite sequences (MMR, mismatch repair). For example, in *Escherichia coli*, only one MutS and one MutL protein participate in these processes at the initial, most important stages, whereas the similar components of the MMR system in yeast, plants and man are represented by a family of proteins (four to seven various polypeptides of the each type) [3]. One of the human homologues of MutL protein is *PMS2* (an orthologous protein in yeast is called *PMS1*). Moreover, in contrast to the majority of eukaryotes, not only the major *PMS2*-encoding master gene *PMS2* but also a whole family of *PMS2*-like sequences that encode numerous and various mRNAs have been discovered on chromosome 7 in the human genome [4–8]. The evolutionary origin, detailed structure, and the biological role of these so-called *PMS2* pseudogenes remain unclear. The goal of this study was a phylogenetic analysis of *PMS2* gene family in the primate lineage and molecular characterization of all *PMS2*-like human genes and corresponding proteins they possibly encode.

## 2 Method and Results

Using bioinformatic (BLAST, BLAT, ClustalW, DS Gene and PHYLIP software packages) and experimental (PCR cloning and sequencing of cDNA and genomic fragments from various species) approaches we carried out a molecular characterization and phylogenetic analysis of the sixteen *PMS2*-like genes present at several loci on chromosome 7 of *Homo sapiens* and corresponding *PMS2* paralogues in other primate species (*Callithrix jacchus*, *Macaca mulatta*, *Nomascus leucogenys*, *Pongo abelii*, *Gorilla gorilla* and *Pan troglodytes*). Our results indicate that amplification of the *PMS2*-related genes is characteristic only for higher primates and was originated as duplication of the *PMS2CL* region (exons 9–15; also known as  $\psi 0$  pseudogene [7, 8]) approximately 18 Mya [million years ago] in lineage leading to modern Lesser (gibbon)

and Great Ape species after its separation from branch of Old World Monkeys. The evolution follows by multiple rounds of amplification of the region *PMS2NL* (exons 1-5; exemplified by  $\psi 1$ – $\psi 14$  pseudogenes in *Homo sapiens* [8]) in orangutan (only one *PMS2NL* gene subgroup present), gorilla and chimpanzee (appearance of the second *PMS2NL* subgroup with only two different members). *Homo sapiens* represents an extreme case of amplification of the second *PMS2NL* gene subgroup with eight for the most part clusterly arranged new members:  $\psi 2$ – $\psi 3$ ,  $\psi 5$ ,  $\psi 6$ – $\psi 8$ ,  $\psi 11$ – $\psi 12$ .

### 3 Discussions

To summarize the above data and taking into account extensive coding capacities of many human-specific *PMS2* paralogues, we have developed the following model of the protein complexes encoded by the aforementioned evolutionarily young *PMS2*-like human genes [9]. All components of this *PMS2* system can be divided into three groups of polypeptides (I, II, and III), specific of solely the MMR and/or gene conversion complexes of the higher primates (first of all *Homo sapiens*, but probably also, and to a lesser extend, African Great apes). In the proposed triple *PMS2* system, class III proteins are responsible for the interaction with other MutL-like proteins of eukaryotic repair and gene conversion systems (primarily with MLH1 and possibly with PMS1 and MLH3 as well). Class II proteins seem to be involved in the recognition of the complexes formed by MutS-like proteins and mismatched or heteroduplexed DNA regions and targeting some other specific protein components of the repair or gene conversion systems to these structures. Class I proteins, owing to the nucleotide (ATP-, GTP?-)-binding and/or nucleotide-hydrolyzing activity, can switch on or off certain stages of DNA repair and gene conversion processes. Besides MMR and gene conversion, the outlined above properties of these novel tripartite h*PMS2*-like complexes could be used in other vitally important molecular genetic processes, such as meiotic nuclear division, somatic hypermutation and class-switch recombination of immunoglobulin genes.

The work was supported by the program "Molecular and Cell Biology" (direction 'Functional Genomics') of the Russian Academy of Sciences and by grant No. 07-04-01167 of the Russian Foundation for Basic Research (RFBR).

### References

- [1] Modrich, P., and Lahue, R., Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.*, 65:101-133, 1996.
- [2] Modrich, P., Doi, K. and Imai, H., Mechanisms in eukaryotic mismatch repair, *J. Biol. Chem.*, 281:30305-30309, 2006.
- [3] Lin, Z., Nei, M., and Ma, H., The origins and early evolution of DNA mismatch repair genes – multiple horizontal gene transfers and co-evolution, *Nucleic Acids Res.*, 35:7591-7603, 2007.
- [4] Horii, A., Han, H.J., Sasaki, S., Shimada, M., and Nakamura, Y., Cloning, characterization and chromosomal assignment of the human genes homologous to yeast *PMS1*, a member of mismatch repair genes, *Biochem. Biophys. Res. Commun.*, 204:1257-1264, 1994.
- [5] Nicolaides, N.C., Carter, K.C., Shell, B.K., Papadopoulos, N., Vogelstein, B., and Kinzler, K.W., Genomic organization of the human *PMS2* gene family, *Genomics*, 30:195-206, 1995.
- [6] Kondo, E., Horii, A., and Fukushige, S., The human *PMS2L* proteins do not interact with hMLH1, a major DNA mismatch repair protein, *J. Biochem.*, 125: 818–825, 1999.
- [7] Nakagawa, H., Lockman, J.C., Frankel, W.L., Hampel, H., Steenblock, K., Burgart, L.J., Thibodeau, S.N., and de la Chapelle A., Mismatch repair gene *PMS2*: disease-causing germline mutations are frequent in patients whose tumors stain negative for *PMS2* protein, but paralogous genes obscure mutation detection and interpretation, *Cancer Res.*, 64:4721-4727, 2004.
- [8] De Vos, M., Hayward, B.E., Picton, S., Sheridan, E., and Bonthron, D.T., Novel *PMS2* pseudogenes can conceal recessive mutations causing a distinctive childhood cancer syndrome, *Am. J. Hum. Genet.* 74:954-964, 2004.
- [9] Shpakovski, D.G., Shematorova E.K., and Shpakovski, G.V., Human *PMS2* gene family: origin, molecular evolution, and biological implications, *Dokl. Biochem. Biophys.*, 408:175-179, 2006.