Published in final edited form as:

J Eukaryot Microbiol. 2011 ; 58(6): 537–538. doi:10.1111/j.1550-7408.2011.00577.x.

Guidelines for the Naming of Genes, Gene Products, and Mutants in the Opportunistic Protists

Andrew H. Limper^{a,1} and Louis M. Weiss^{b,1}

^aThoracic Diseases Research Unit, Department of Medicine, Division of Pulmonary and Critical Care Medicine, Mayo Clinic College of Medicine, Rochester, Minnesota 55905

^bDepartment of Pathology, Division of Infectious Disease, Albert Einstein College of Medicine, Bronx, New York 104615

Abstract

The opportunistic protists encompass a wide diversity of organisms including *Pneumocystis*, *Toxoplasma*, cryptosporidia, microsporidia, and related genera. Recent advances in the molecular biology and cellular biochemistry of these organisms have led to the identification of an ever growing numbers of key genes and their cognate proteins. Until now, these molecules have not been designated using any consistent nomenclature system, leading to considerable confusion. The participants of the 11th International Workshop on Opportunistic Protists met on August 3, 2010 to reach consensus of a nomenclature system for genes, gene products, and mutants in the opportunistic protists. The following summary reports the consensus agreement to move toward a unified nomenclature system for these organisms. The system is adapted from that used for *Saccharomyces cerevisiae*.

Keywords

Cryptosporidia; genes; microsporidia; nomenclature; *Pneumocystis*; *Toxoplasma*

The last decade has seen dramatic expansion of genetic, genomic, and protein characterization studies across all of the opportunistic protists investigated by members of their respective research communities. This has been facilitated both by large genome projects, such as the *Pneumocystis carinii* genome project (<http://pgp.cchmc.org/>), as well as by individual laboratory efforts (Cushion and Smulian 2001). In addition, improved biochemical methods have led to isolation and characterization of protistan proteins and their functional activities. Occasional confusion has arisen by naming proteins by their molecular weights or by selecting gene names that are not consistent with orthologous genes in related species. In addition, due to the lack of consistent nomenclature for naming genes and cognate proteins, the literature now contains the same molecular species bearing a variety of names across multiple publications. Accordingly, the International Workshop on Opportunistic Protists-11 conducted a roundtable discussion on August 3, 2010 to find a common set of ground rules for naming genes and related gene products in the opportunistic protistan organisms, including *Pneumocystis*, *Toxoplasma*, cryptosporidia, microsporidia, and related genera.

On Behalf of the participants of the 11th International Workshop on Opportunistic Protists, held August 3, 2010, Hilo, HI.

© 2011 The Author(s)

Corresponding Author: A. Limper, Thoracic Diseases Research Unit, Department of Medicine, Division of Pulmonary and Critical Care Medicine, Mayo Clinic, Gonda-18 South, Rochester, Minnesota 55905—Telephone number: +1507 284 4348; FAX number: +1507 255 4372; limper.andrew@mayo.edu.

It was generally agreed that any nomenclature system should be convenient, easily referable, and should be consistent with the publication requirements of the major journals in which work is published, including the journals of the American Society of Microbiology. The attendees were strongly supportive of adopting or modifying a previously accepted system of nomenclature. In this light, consensus rapidly focused on adopting the general system of nomenclature utilized in the *Saccharomyces cerevisiae* research community. *Saccharomyces cerevisiae* represents one of the first organisms for which a complete genome was cloned and annotated (Iwasaki et al. 1991). This system has further been used extensively for genetic modeling and protein characterization. As such, it represents a suitable prototype for gene and gene product nomenclature for the opportunistic protists. The *Neurospora* and *Toxoplasma* research communities have adopted previously a similar approach (http://www.yeastgenome.org/sgdpub/Saccharomyces_cerevisiae.pdf) (Sibley et al. 1991). This type of approach has also been previously proposed for *Pneumocystis*-related genes and proteins (Stringer and Cushion 1998).

The following ground rules were proposed, discussed, and adopted by consensus by the Participants of the 11th International Workshop on Opportunistic Protists held on August 3, 2010 in Hilo, HI.

1. Gene names will be designated by a three-letter designation followed by an integer if multiple members of that gene sequence are known. In general, the gene should be noted in all lower case italicized letters. For instance, the gene for the cell wall biosynthesis kinase of *P. carinii* would be designated *cbk1* (Kottom and Limper 2004).
2. It was agreed that for most publications it is not necessary to designate the species of origin. Accordingly, the cell wall biosynthesis kinase of *P. carinii* would be designated *cbk1* rather than *Pccb1*. However, in a particular manuscript when orthologous genes of multiple species are discussed together, then it becomes acceptable to use the species designation. For instance, in a paper discussing cell wall biosynthesis kinases of both *P. carinii* and *Pneumocystis jirovecii*, it would be appropriate to designate them as *Pccb1* and *Pjcb1*, respectively, in that manuscript.
3. If the gene is known, the cognate protein should be non-italicized with the first letter capitalized. For example, the cognate protein of *P. carinii cbk1* would be designated as Cbk1. The former practice of appending a lower case “p” to designate the cognate protein (i.e. Cbk1p) has decreased in usage in many nomenclatural systems over the past few years (Fundel and Zimmer 2006).
4. Whenever possible, the three-letter designation should be based upon orthologues that are present in other species. Thus, if an orthologue has been designated for *S. cerevisiae*, the same three-letter designation should be applied for the gene in the opportunistic protist.
5. The designation of a newly isolated protein, for which the corresponding genetic sequence is not yet known, should also use a similar three-letter designation followed by an integer. The designation of proteins by apparent molecular weight is to be discouraged.
6. The designation of manipulated genes should follow the above standard with deleted genes being indicated by a “Δ” (i.e. deletion of the cognate gene would be indicated as Δ*cbk*) and mutated genes indicated by the addition of a superscript (i.e. a mutated cognate gene would be indicated as *cbk^{mut}*). If multiple mutants are made, these could be indicated by serial numbers (i.e. mut1, mut2, etc.) or by

designation of the specific mutation (i.e. ala to val would be *cbk^{val}* and ala to glu indicated by *cbk^{glu}*).

The above guidelines are intended to be a suggested direction to achieve greater uniformity in gene and cognate protein nomenclature. However, it is clearly understood that unique features of each of the opportunistic protists may require modification of these suggestions over time. Further refinements or modifications of these suggestions are welcome as the genetics of these interesting protistan organisms continues to evolve.

LITERATURE CITED

- Cushion MT, Smulian AG. The *Pneumocystis* genome project: update and issues. J. Eukaryot. Microbiol. 2001; (Suppl 48):182S–183S. [PubMed: 11906055]
- Fundel K, Zimmer R. Gene and protein nomenclature in public databases. BMC Bioinform. 2006; 7:372–385.
- Iwasaki T, Shirahige K, Yoshikawa H, Ogasawara N. The direct cloning of the yeast genome using the gap-filling method and the complete physical mapping of *Saccharomyces cerevisiae* chromosome VI. Gene. 1991; 109:81–87. [PubMed: 1756985]
- Kottom TJ, Limper AH. *Pneumocystis carinii* cell wall biosynthesis kinase gene *cbk1* is an environmentally responsive gene that complements cell wall defects of cbk-deficient yeast. Infect. Immun. 2004; 72:4628–4636. [PubMed: 15271923]
- Sibley LD, Pfefferkorn ER, Boothroyd JC. Proposal for a uniform genetic nomenclature in *Toxoplasma gondii*. Parasitol. Today. 1991; 7:327–328. [PubMed: 15463406]
- Stringer JR, Cushion MT. The genome of *Pneumocystis carinii*. FEMS Immunol. Med. Microbiol. 1998; 22:15–26. [PubMed: 9792057]