BIOLOGY OF REPRODUCTION **68,** 358–362 (2003)
Published online before print 17 October 2002.
DOI 10.1095/biolreprod.102.008086

Minireview

Evolution and Nomenclature of the Zona Pellucida Gene Family

Scott C. Spargo[1] and Rory M. Hope

Laboratory of Molecular Evolution, Department of Molecular Biosciences, University of Adelaide, South Australia 5005, Australia

ABSTRACT

Three subfamilies of genes are acknowledged within the zona pellucida (ZP) gene family. At present, these subfamilies each have two names that are used interchangeably: ZPA or ZP2, ZPB or ZP1, and ZPC or ZP3. The ZPA genes encode the longest protein sequences and the ZPC genes the shortest. Recently, several sequences, which have no clear relationship to the three subfamilies, have been identified. These sequences include two paralogous ZP genes from *Xenopus laevis* and a single gene from the fish *Oryzias latipes*. We have conducted extensive phylogenetic analyses of the known ZP genes. As well as establishing the evolutionary relationships among these genes, the analyses make it clear that the dual nomenclature system is no longer feasible, because major paralogous groups are present in the ZPB (ZP1) family of genes of amniotes. We propose a unified system of nomenclature for the ZP gene family that removes the existing ambiguities.

gamete biology

INTRODUCTION

A glycoprotein coat surrounds all ovulated, vertebrate oocytes. This coat has a different name in each major vertebrate lineage: the chorion in fish, the vitelline envelope in amphibians, the perivitelline envelope in reptiles and birds, and the zona pellucida in mammals. Despite their different names, the gross structure and function of these egg coats are similar, and they will be referred to collectively as the zona pellucida, or ZP. ZP are composed of cross-linked, sulfated glycoproteins [1, 2]. Early studies in *Mus musculus* and *Rattus norvegicus* identified three glycoprotein components of the ZP using SDS-PAGE; these components were named ZP1, ZP2, and ZP3 from highest to lowest apparent molecular weight, respectively [3, 4]. Subsequent ZP glycoproteins have been named according to several criteria, including apparent molecular weight of the glycoprotein [5], protein sequence length [6], and sequence identity comparisons [7, 8].

By 1994, inconsistent naming of the ZP genes had resulted in a confused system of nomenclature, making communication about the ZP gene family difficult. To overcome this difficulty, Harris et al. [6] proposed a unified system of nomenclature in which ZP genes were named in order of the length of their encoded protein sequence, from longest to shortest. Unfortunately, under these criteria, the mouse ZP genes fall in the order ZP2, then ZP1, and then ZP3. To avoid confusion arising from swapping the names of mouse ZP1 and ZP2, a new letter-based system was introduced. Thus, ZP2 became ZPA, ZP1 became ZPB, and ZP3 became ZPC. Acceptance of the proposed new system of nomenclature has been less than universal, with the number system and the letter system being used concurrently by different groups [9, 10].

The complete cDNA sequence of the mouse ZP1 gene was determined subsequent to the sequence of the human ZPB cDNA sequence [6, 11]. It was assumed that the mouse ZP1 was orthologous to human ZPB (i.e., both genes descended from a single common ancestral gene that diverged at a speciation event), because it shared greater amino acid sequence identity with human ZPB than with human ZPA or ZPC. Even before publication of the chicken and human ZP1 gene sequences [7, 12], phylogenetic analyses suggested that mouse ZP1 and human ZPB were paralogous (i.e., descended from different genes created by a gene duplication). Hughes and Barrett [12] identified a human genomic sequence orthologous to the mouse ZP1 gene and paralogous to the human ZPB gene. Soon after, Bausek et al. [7] published the sequence of a chicken ZP1 gene. Their analysis suggested that a gene duplication, predating the divergence of fish and amphibians, gave rise to two paralogous groups of genes within the ZPB subfamily: ZP1 and ZPB.

Whereas the mammalian ZP has three main glycoprotein components, the egg envelope of *Xenopus laevis* is comprised of six or more [4]. An orthologue of each of the three groups of mammalian ZP glycoprotein has been found in *X. laevis* [5, 13, 14]. Two additional *X. laevis* ZP genes, with no known mammalian orthologues, have also been characterized: ZPAX [15] and ZPA [16]. An additional gene in *Oryzias latipes*, called ZPA, shares a high level of sequence identity with the *X. laevis* ZPAX gene. Evolutionary analysis suggested that the *Xenopus* ZPA gene is an orthologue of the human and pig ZPA genes and that ZPAX evolved from a gene paralogous to the precursor of the ZPA and ZPB genes [15].

The three accepted subfamilies of genes within the ZP gene family each have two names at present: ZPA or ZP2, ZPB or ZP1, and ZPC or ZP3. The difficulties in maintain-

[1]Correspondence. FAX: 61 08 83034362;
e-mail: scott.spargo@adelaide.edu.au

Received: 10 June 2002.
First decision: 25 June 2002.
Accepted: 8 July 2002.

ISSN: 0006-3363. http://www.biolreprod.org

ing this dual nomenclature system have come to a head. It is no longer possible to use ZPB and ZP1 interchangeably to describe the same orthologous set of genes, because a subset of the genes called ZP1 is paralogous to other ZPB (ZP1) genes. In addition to these three subfamilies, two ZP genes from *X. laevis* and one ZP gene from *O. latipes* have evolutionary origins that are uncertain. Here, we show the relationships within the ZP gene family in a comprehensive phylogeny based on an analysis of cDNA sequences, and we use the phylogeny to address problems with the current system of nomenclature.

PHYLOGENETIC ANALYSES

Methodology

Sequences included in phylogenetic analyses of the ZP genes are listed in Table 1. This table includes the name for each of the genes as it appears in the GenBank database together with a suggested name based on the phylogenetic relationships established in the present study.

Nucleotide sequence alignments of the ZP genes were carried out based on protein sequence alignments using both BioEdit [17] and CLUSTAL W [18] software. Phylogenetic analyses were carried out using the following methods: 1) maximum parsimony with bootstrap; 2) UPGMA (unweighted pair group method with arithmetic mean) using maximum likelihood distances with bootstrap (PHYLIP software [19]); 3) maximum likelihood (PAUP software [20]); 4) likelihood with Markov-chain Monte Carlo (MCMC) sampling (BAMBE software [21]); and 5) quartet puzzling (Puzzle [22]).

Bootstrap analyses were conducted using 1000 pseudoreplicates (parsimony) or 100 pseudoreplicates (maximum likelihood distances) of the data set. Likelihood with MCMC analyses were performed using the likelihood model described by Tamura and Nei [23]. Codon positions were analyzed independently of each other. Nucleotide frequencies and substitution rates were estimated independently for each codon position. Branch swapping was global for both the burn-in and the cycling phase. The rate of convergence of the Markov chain to the posterior distribution was optimized as directed by the authors of the BAMBE program. The consensus tree shown (Fig. 1) was found in three separate runs from random starting points, suggested by Simon and Larget [21] as the minimum requirement for demonstrating sufficient sampling from the posterior probability distribution. Relative rates tests were performed using the program K2wuli [24], which conducts the relative rates test described by Wu and Li [25].

BAMBE Tree

The results of a likelihood analysis with MCMC are shown in Figure 1. A clear hierarchy of relatedness is apparent in the unrooted phylogeny of the ZP genes (Fig. 1). A suggestion for a consistent nomenclature system, based on the hierarchical structure of this phylogeny, is outlined later. At the highest level of the hierarchy lies the gene family, which includes every ZP gene. The gene family can be easily separated into four subfamilies, which are indicated in Figure 1. All the genes of a subfamily are paralogous to all the genes of other subfamilies. The genes in each of the ZPA, ZPB, and ZPC subfamilies can be traced to a single ancestral duplicate gene. The genes of the ZPX subfamily can be traced to two ancestral duplicate genes. Characterization of additional ZPX genes is required to determine if these genes are a single subfamily encompassing two paralogous lineages (like the ZPB subfamily) or two distinct subfamilies. A third level of hierarchy exists within the ZPB subfamily, which has three gene groups: ZPB, ZPB1, and ZPB2. The ZPB group is orthologous to both the ZPB1 and ZPB2 groups, whereas the two latter groups are paralogous, having evolved from different duplicates of an ancestral ZPB gene. It is unfortunate that the probabilities associated with the branching points of the ZPX genes are low. This may be attributed to the duplications that gave rise to both ZPX paralogues and the ZPB and ZPA subfamilies occurring over a short period of evolution, then being reconstructed after a comparatively long period of evolution.

Other Methods

The major branches of the maximum parsimony and the maximum likelihood distance trees had an identical topology to the tree shown in Figure 1. The bootstrap values associated with both trees were generally lower than published logical probabilities (data not shown; available on request). Branching orders of the *X. laevis* ZPX2 (*xenopus*-ax) and the ZPX1 genes were supported by maximum parsimony bootstrap values of 65% and 52%, respectively. The maximum likelihood distance tree placed the ZPX genes in their illustrated position with a bootstrap of 81% (though the relationships between the individual ZPX genes were unresolved). Quartet puzzling was unable to resolve any of the deep branches of the ZP gene tree.

EVOLUTIONARY HISTORY

At least one ZP gene evolved during the earliest stages of vertebrate evolution. This is apparent because every major vertebrate lineage has ZP genes and because the phylogenetic relationships indicate that all the major subfamilies of ZP genes evolved before the divergence of fish and amphibians. Because the products of all known ZP genes are components of the glycoprotein egg coat, the first ZP protein product likely fulfilled a related function. As a consequence of the ancient origins of the ZP gene family, identification of an appropriate outgroup for the full-length ZP gene sequences has not been possible. Without an outgroup, it is difficult to determine the temporal order of events in ZP gene evolution. Several alternate approaches were used to determine the likely order in which the subfamilies and groups of ZP genes evolved.

Almost all ZP genes have a signal peptide at their N-terminus and a *trans*-membrane domain at their C-terminus. Between these two domains, ZP genes have a conserved region originally characterized in the ZP genes: the ZP domain [26]. No appropriate outgroup exists for use with full-length ZP cDNA sequences, but other known sequences contain ZP domains [26, 27]. When the ZP domain of chicken β-tectorin was used as an outgroup for phylogenetic analyses of the ZP domains of the ZP genes, the root of the tree consistently lay between the ZPC subfamily and the remainder of the tree. Confidence levels associated with the branching order of the ZP domain trees were lower than those on the full-length tree (Fig. 1), and small changes in the branching order were observed.

Another approach to identifying the most ancient branch of the ZP gene phylogeny required the assumption that the ZP genes evolved in a clock-like manner (i.e., that the rates of nucleotide substitution are approximately equal for the genes being studied). If this assumption holds true within

TABLE 1. Characterized genes belonging to the ZP gene family.

Species name	Common name	Tree label	Accession no.	GenBank name	Suggested name
Callithrix jacchus	Marmoset	marmoset-1	Y10822	ZP1	ZPB2
		marmoset-2	Y10767	ZP2	ZPA
Canis familiaris	Dog	dog-c	U05780	ZPC	ZPC
		dog-a	U05779	ZPA	ZPA
Carassius auratus	Goldfish	goldfish-3	Z48974	ZP3	ZPC
Coturnix japonica	Japanese quail	quail-1	AB061520	ZP1	ZPB1
Cyprinus carpio	Common carp	carp-3a/3c	L41637/L41639	ZP3A/3C	ZPCa[a]
		carp-3b	L41638	ZP3B	ZPCb[a]
		carp-2a	Z72491	ZP2a	ZPBb[b]
		carp-2b	Z72492	ZP2b	ZPBa[b]
		carp-2c	Z72493	ZP2c	ZPBc[b]
Dunio rerio	Zebrafish	zebrafish-3	AF095457	ZP3	ZPC
		zebrafish-2	AF331968	ZP2	ZPB
Felis catus	Cat	cat-a	U05776	ZPA	ZPA
		cat-b	U05777	ZPB	ZPB2
		cat-c	U05778	ZPC	ZPC
Gallus gallus	Chicken	chicken-1	AJ28967	ZP1	ZPB1
		chicken-b	AB025428	ZPB	ZPB2
		chicken-c	AB031033	ZPC	ZPC
Homo sapiens	Human	human-a	M90366	ZPA	ZPA
		human-b	U05781	ZPB	ZPB2
		human-c	A18567	ZPC	ZPC
		human-1	AC004126	ZP1	ZPB1
Macaca radiata	Macaque	macaque-1	Y10381	ZP1	ZPB2
		macaque-2	Y10690	ZP2	ZPA
		macaque-3	X82639	ZP3	ZPC
Mesocricetus auratus	Golden hamster	hamster-3	M63629	ZP3	ZPC
Microtus brandti	Brandt's vole	vole-c	AF304487	ZPC	ZPC
Mus musculus	House mouse	mouse-1	MN_009580	ZP1	ZPB1
		mouse-2	NM_011775	ZP2	ZPA
		mouse-3	NM_011776	ZP3	ZPC
Notomys alexis	Hopping mouse	nalex-c	AY078054	ZPC	ZPC
Oncorhynchus mykiss	Rainbow trout	trout-α	AF231706	VEP alpha	ZPBa[c]
		trout-β	AF231707	VEP beta	ZPBb[c]
		trout-γ	AF231708	VEP gamma	ZPC
Oryctolagus cunniculus	Rabbit	rabbit	M58160	ZP	ZPB2
Oryzias latipes[d]	Japanese medaka	medaka-a	AF128807	ZPA	ZPX1
		medaka-b	AF128808	ZPB	ZPBa
		medaka-c1	AF128809	ZPC protein 1	ZPCe
		medaka-c3	AF128811	ZPC protein 3	ZPCd
		medaka-c4	AF128812	ZPC protein 4	ZPCb
		medaka-c5	AF128813	ZPC protein 5	ZPCa
		medaka-cgH	D89609	choriogenin H	ZPBb
		medaka-lsf	D38630	L-SF	ZPCc
Pseudomys australis	Plains rat	paust-c	AY078055	ZPC	ZPC
Pseudopleuronectus americanus	Winter flounder	flounder	U03674	ZP	ZPB
Rattus norvegicus	Rat	rat-1	AB000928	ZP1	ZPB1
		rat-2	NM_031150	ZP2	ZPA
		rat-3	D78482	ZP3	ZPC
Salmo salar	Salmon	salmon-esp	AJ000665	ZPB	ZPB
Sus scrofa	Pig	pig-1/2	S74651/D45064	ZP1/2	ZPA[e]
		pig-3	D45065	ZP3	ZPC
Trichosurus vulpecula	Brush-tailed possum	possum-a	AF263014	ZPA	ZPA
		possum-b	AF263013	ZPB	ZPB
		possum-3	AF079524	ZP3	ZPC?
Xenopus laevis	African clawed frog	*X. laevis*-69	AF038151	69 kDa/ZPA[f]	ZPA
		X. laevis-37	U44950	37 kDa	ZPB
		X. laevis-c	U44952	ZPC	ZPC
		X. laevis-ax	AF225906	ZPAX	ZPX1
		X. laevis-a	U44949	ZPA[f]	ZPX2

[a] *Cyprinus carpio* (carp) ZP3A and ZP3C (current nomenclature) are most closely related to each other in the phylogeny and share 96% sequence identity. They are probably alleles of the same gene and are putatively called carp ZPCb. Carp has a second, larger ZPC gene now called carp ZPCa.
[b] *Cyprinus carpio* (carp) ZP2A, ZP2B, and ZP2C (current nomenclature) are more closely related to each other than the other fish ZPB genes. It is not certain whether theses sequences are separate genes, alleles of the same gene, or a combination of both. They are temporarily assigned the name ZPB and lettered in order of protein sequence length; ZPBa (2B), ZPBb (2A) and ZPBc (2C).
[c] The *Oncorhynchus mykiss* (rainbow trout) ZP genes, currently called vitelline envelope proteins α and β, share 72% identity in their cDNA. Based on this low level of sequence identity, these genes are putatively classed as paralogues and temporarily assigned the names ZPBa (alpha) and ZPBb (beta).
[d] The *Oryzias latipes* (Japanese medaka) ZPB genes, currently ZPB and choriogenin-H, are clearly shown to be paralogues in the phylogeny. ZPB is now ZPBa, and choriogenin-H is now ZPBb. The ZPC genes, currently five full-length sequences, are lettered from largest to smallest protein products (ZPCa to ZPCe).
[e] *Sus scrofa* (pig) ZP1 and ZP2 sequences share 99% nucleotide sequence identity over 2203 base pairs. This suggests, in the absence of intron sequence data, that these sequences are alleles of the same gene, *S. scrofa* ZPA.
[f] *Xenopus laevis* sequences xla69 and xlax1 are both listed as xlZPA in GenBank, though they are different genes.

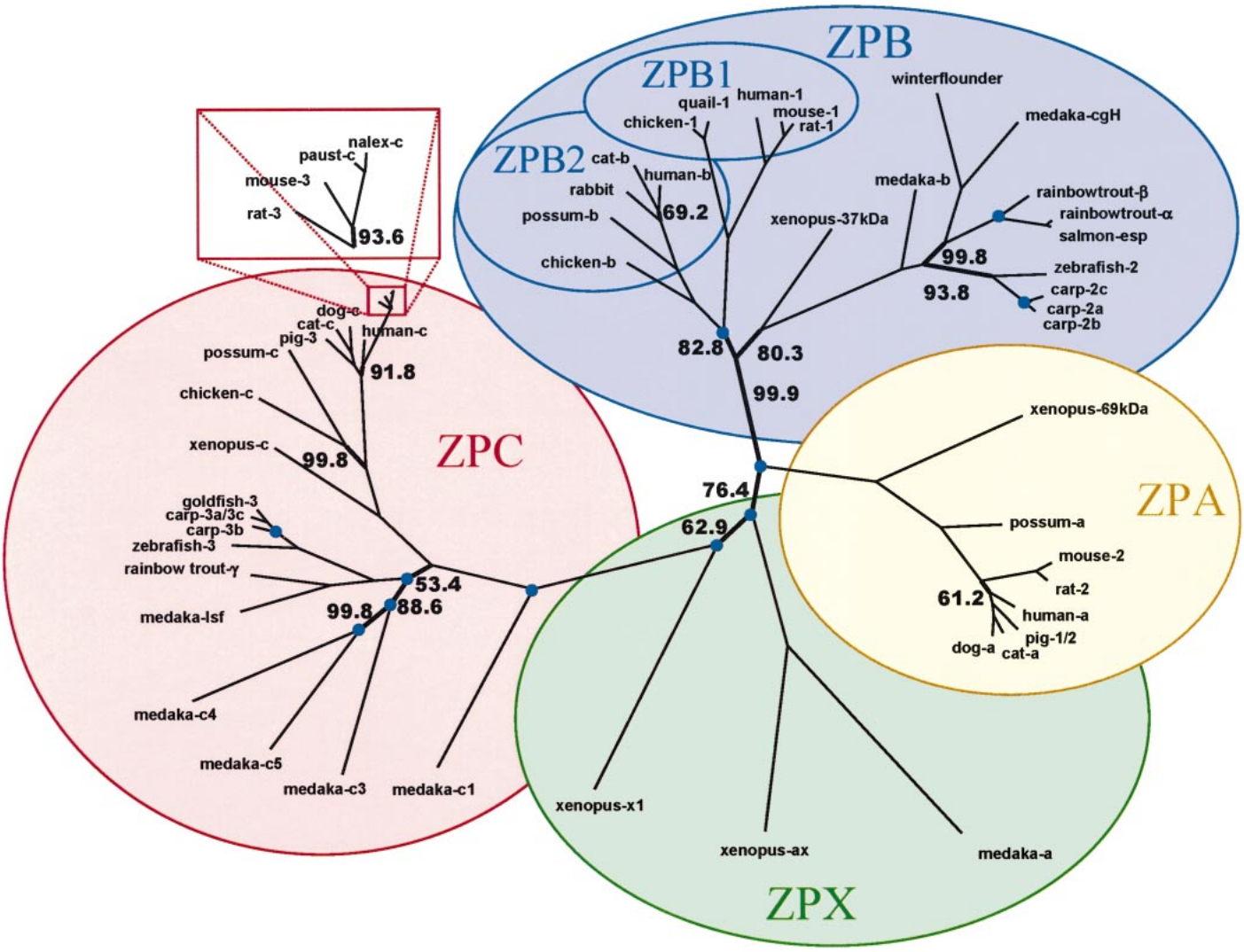

FIG. 1. Phylogenetic tree of the ZP gene family derived using likelihood with MCMC (BAMBE [21]). Several of the genes in Table 1 are not present here because of computational limitations. Relative branch lengths indicate rates of evolution along a particular branch. Subfamilies of ZP genes are labeled (ZPA, ZPB, ZPC, and ZPX), and groups are also labeled (ZPB1 and ZPB2). Logical probabilities associated with branching orders and sequences likely to be alleles of the same gene are indicated. Where probabilities are not shown, the value is 100%. Internal branching nodes, which can only be explained by a duplication event, are marked with an open circle. In the absence of contradictory evidence, other branching nodes are assumed to represent speciation events.

statistical limits and a phylogeny is derived under this assumption, then the deepest branch of this phylogeny should approximate the position of the root of the tree. Relative-rates tests using the method of Wu and Li [25], which requires an outgroup, showed that the assumption of a molecular clock held true across the ZP domain of almost all of the ZP genes (data not shown). Rodent ZPAs, *X. laevis* ZPC, and the ZPB1 genes were the exceptions. Subsequent UPGMA and BAMBE analyses, for which the molecular clock was assumed, of both full-length and ZP domain sequences supported a phylogenetic root between the ZPC subfamily and the remainder of the tree.

These findings provide some evidence that the first event in ZP evolution was a gene duplication event, which gave rise to the ancestral ZPC gene and to the precursor of the ZPA, ZPB, and ZPX subfamilies. This precursor duplicated at least three times over a short period of evolutionary history, giving rise to the ancestral ZPX genes and the ancestral ZPA and ZPB genes. These events all occurred early during vertebrate evolution, before evolution of the first amphibians. Subsequently, duplication events have occurred in several lineages, with the most notable occurring early during evolution of the amniotes and giving rise to the ZPB1 and ZPB2 groups within the ZPB subfamily.

Other important observations and conclusions can be inferred from the tree shown in Figure 1. First, a consistent grouping, supported by all the phylogenetic methods used, of the *Gallus gallus* (chicken) and *Trichosurus vulpecula* (brush-tailed possum) ZPC genes to the exclusion of other ZPC genes is observed. A possible explanation for this relationship is that the chicken and possum ZPC genes are paralogous to the eutherian ZPC genes. Second, the gene duplication that gave rise to the ZPA genes occurred before the divergence of fish and amphibians; therefore, at some point during evolution, fish had a ZPA gene. Fish have either lost the ZPA gene or have an as-yet-unidentified ZPA gene or pseudogene. The alternate possibilities, that the fish ZP genes evolved from a ZPA/ZPB precursor or that the ZPX1 genes are actually ZPA orthologues, are phylogenetically unlikely. Third, the ZPX genes originated from du-

plications before the divergence of fish and amphibians. It is anticipated that genes currently present in fish and amphibians were also present in early amniotes. The absence of any known ZPX orthologues in extant amniotes suggests the loss of ZPX precursors early during amniote evolution.

ZP GENE NOMENCLATURE

A list of ZP genes is given in Table 1. This list includes the current name assigned to each ZP gene in GenBank followed by a name based on a suggested, standard system of nomenclature. We propose that the basis for this system be the division of the ZP genes into subfamilies based on the evolutionary relationships within the ZP gene family, that these subfamilies should be named in order of their relative size, and that the protein sequence length provides the most reliable and consistent measure of size for this purpose. To minimize the changes involved in adopting a standard nomenclature system, we have modified the closest existing nomenclature system [6]. This system, known as the Harris nomenclature system, groups the ZP genes into three phylogenetic subfamilies, which are then assigned letters alphabetically in order of coding sequence length. These three subfamilies are ZPA, ZPB, and ZPC, with uppercase letters used to indicate subfamilies. The Harris nomenclature does not address the possibility of paralogues within a subfamily. We suggest that group-level paralogues in the same subfamily should be numbered in order of coding sequence length, with reference to *Homo sapiens* when ambiguity exists (in birds and mammals, ZPB/ZP1 genes become ZPB1 and ZPB2 genes). Paralogues below the group level of the phylogeny should be differentiated using an additional lowercase letter, also in order of coding sequence length (in rainbow trout, vitelline envelope proteins α and β become ZPBa and ZPBb, respectively). In all cases, the gene name should follow the hierarchy of the phylogeny: subfamily first (uppercase letter), then group (number, when appropriate), and finally, a paralogue label (lowercase letter, when appropriate).

We have made an exception in the nomenclature system for the two *X. laevis* genes and the *O. latipes* gene that have no known amniote orthologues. To ensure that the *X. laevis* and *O. latipes* ZPA, ZPB, and ZPC genes have the same name as their orthologues, these three sequences are named ZPX1 and ZPX2 (Table 1).

This is only an initial attempt to create a logical system of nomenclature. It is designed to air the current problems and to promote discussion of possible solutions. We propose that those individuals with interest in the ZP genes and proteins establish a working group to examine the problems surrounding the present systems of nomenclature and make recommendations for a new system that would find wide acceptance.

ACKNOWLEDGMENT

We would like to thank Associate Professor W. Breed, Department of Anatomical Sciences, University of Adelaide, for his helpful comments and suggestions throughout the course of this project.

REFERENCES

1. Wassarman PM. Mammalian fertilization: molecular aspects of gamete adhesion, exocytosis and fusion. Cell 1999; 96:175–183.
2. Parillo F, Fagioli O, Dall'Aglio C, Verini-Supplizi A. Lectin histochemical detection of sulfoglycans in the zona pellucida of mammalian antral oocytes. Acta Histochem 2000; 102:193–202.
3. Repin VS, Akimova IM. A study of the protein composition of the zona pellucida of mammalian oocytes and zygotes by a method of microelectrophoresis in polyacrylamide gel. Biokhimiia 1976; 41:50–57.
4. Bleil JD, Wassarman PM. Structure and function of the zona pellucida: identification and characterization of the proteins of the mouse oocyte's zona pellucida. Dev Biol 1980; 76:185–202.
5. Tian J, Gong H, Lennarz WJ. *Xenopus laevis* sperm receptor gp69/64 glycoprotein is a homologue of the mammalian sperm receptor ZP2. Proc Natl Acad Sci U S A 1999; 96:829–834.
6. Harris JD, Hibler DW, Fontenot GK, Hsu KT, Yurewicz EC, Sacco AG. Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB, and ZPC gene families. DNA Seq 1994; 4:361–393.
7. Bausek N, Waclawek M, Schneider WJ, Wohlrab F. The major chicken egg envelope protein ZP1 is different from ZPB and is synthesized in the liver. J Biol Chem 2000; 275:28866–28872.
8. Wang H, Gong Z. Characterization of two zebrafish cDNA clones encoding egg envelope proteins ZP2 and ZP3. Biochim Biophys Acta 1999; 1446:156–160.
9. Kubo H, Kawano T, Tsubuki S, Kotani M, Kawasaki H, Kawashima S. Egg envelope glycoprotein gp37 as a *Xenopus* homologue of mammalian ZP1, based on cDNA cloning. Dev Growth Differ 2000; 42: 419–427.
10. Vo LH, Hedrick JL. Independent and hetero-oligomeric-dependent sperm binding to egg envelope glycoprotein ZPC in *Xenopus laevis*. Biol Reprod 2000; 62:766–774.
11. Epifano O, Liang LF, Dean J. Mouse Zp1 encodes a zona pellucida protein homologous to egg envelope proteins in mammals and fish. J Biol Chem 1995; 270:27254–27258.
12. Hughes DC, Barrat LR. Identification of the true human orthologue of the mouse *Zp1* gene: evidence for greater complexity in the mammalian zona pellucida? Biochim Biophys Acta 1999; 1447:303–306.
13. Kubo H, Kawano T, Tsubuki S, Kawashima S, Katagiri C, Suzuki A. A major glycoprotein of *Xenopus* egg vitelline envelope, gp41, is a frog homologue of mammalian ZP3. Dev Growth Differ 1997; 39: 405–417.
14. Yang JC, Hedrick JL. cDNA cloning and sequence analysis of the *Xenopus laevis* egg envelope glycoprotein gp43. Dev Growth Differ 1997; 39:457–467.
15. Lindsay LL, Wallace MA, Hedrick JL. A hatching enzyme substrate in the *Xenopus laevis* egg envelope is a high molecular weight ZPA homologue. Dev Growth Differ 2001; 43:305–313.
16. Lindsay LL, Yang JC, Hedrick JL. Identification and characterisation of a unique *Xenopus laevis* egg envelope component, ZPD. Dev Growth Differ 2002; 44:205–212.
17. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999; 41:95–98.
18. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22:4673–4680.
19. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Seattle: Department of Genetics, University of Washington; 1993: Computer Program.
20. Swofford DL. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland, MA: Sinauer Associates; 1998: Computer Program.
21. Simon D, Larget B. Bayesian analysis in molecular biology and evolution (BAMBE), version 2.03 beta. Pittsburgh, PA: Department of Mathematics and Computer Science, Duquesne University; 2000: Computer program.
22. Strimmer K, von Haeseler A. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol Biol Evol 1996; 13:964–969.
23. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 1993; 10:512–526.
24. Jermiin LS. K2wuli version 1.0. Canberra: John Curtin School of Medical Research, Australian National University; 1996: Computer program.
25. Wu CI, Li WH. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci U S A 1985; 82:1741–1745.
26. Bork P, Sander C. A large domain common to sperm receptors (Zp2 and Zp3) and TGF-β type III receptor. FEBS Lett 1992; 300:237–240.
27. Killick R, Legan PK, Malenczak C, Richardson GP. Molecular cloning of chick β-tectorin, an extracellular matrix molecule of the inner ear. J Cell Biol 1995; 129:535–547.