Available online at www.sciencedirect.com
SCIENCE DIRECT
Genomics 86 (2005) 674 – 684
GENOMICS
www.elsevier.com/locate/ygeno

The phylogenetic analysis of tetraspanins projects the evolution of cell–cell interactions from unicellular to multicellular organisms☆

Shengfeng Huang [a], Shaochun Yuan [a], Meiling Dong [a], Jing Su [a], Cuiling Yu [a], Yang Shen [a], Xiaojin Xie [a], Yanhong Yu [a], Xuesong Yu [a], Shangwu Chen [a], Shicui Zhang [b], Pierre Pontarotti [c], Anlong Xu [a,*]

[a] *State Key Laboratory of Biocontrol, Guangdong Province Key Laboratory of Therapeutic Functional Genes, The Open Laboratory for Marine Functional Genomics of State High-Tech Development Program, Department of Biochemistry, College of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*

[b] *Department of Marine Biotechnology, Ocean University of China, Qingdao 266003, People's Republic of China*

[c] *EA 3781 EGEE (Evolution, Genome, Environment), Université de Provence, Marseille Cedex 03, France*

Received 14 May 2005; accepted 17 August 2005
Available online 20 October 2005

Abstract

In animals, the tetraspanins are a large superfamily of membrane proteins that play important roles in organizing various cell–cell and matrix–cell interactions and signal pathways based on such interactions. However, their origin and evolution largely remain elusive and most of the family's members are functionally unknown or less known due to difficulties of study, such as functional redundancy. In this study, we rebuilt the family's phylogeny with sequences retrieved from online databases and our cDNA library of amphioxus. We reveal that, in addition to in metazoans, various tetraspanins are extensively expressed in protozoan amoebae, fungi, and plants. We also discuss the structural evolution of tetraspanin's major extracellular domain and the relation between tetraspanin's duplication and functional redundancy. Finally, we elucidate the coevolution of tetraspanins and eukaryotes and suggest that tetraspanins play important roles in the unicell-to-multicell transition. In short, the study of tetraspanin in a phylogenetic context helps us understand the evolution of intercellular interactions.

Keywords: Tetraspanin; Tetraspanin-like; Phylogeny; Unicellularity; Multicellularity; Coevolution

Tetraspanin is a large superfamily of transmembrane proteins that exist widely in animals. Typically, there are 33 tetraspanins in human, 36 in *Drosophila melanogaster,* and 20 in *Caenorhabditis elegans*. A typical tetraspanin features four transmembrane (TM) domains, a small extracellular (EC1) domain, and a large extracellular (EC2) domain. EC2 is the major part responsible for the specific binding of partner proteins [1]. From a crystallographic study of CD81 EC2 and an extensive 3D modeling study, a general mushroom-like structure was extrapolated for all EC2s [2,3] (Fig. 1B). Although the "mushroom" head is highly variable in length and composition in different tetraspanins, the EC2 structure is strictly defined by the "stalk" (helix A/B/E) and two 100% conserved disulfide bonds (up to four in many cases). Considering its importance and structural conservation, EC2 is used to classify different tetraspanins [3,4] (Table 1).

Tetraspanins are virtually distributed in all cell types and each cell expresses several types of tetraspanins; these molecules are involved in various cell–cell and matrix–cell interactions, including cell adhesion, migration, signal transduction, activation, proliferation, and differentiation [4–6]. Having known that each of the tetraspanins, such as CD9, CD37, CD53, CD63, and CD81, can bind with a group of protein partners via EC2 and facilitate their functional implementation, Maecker et al. [7] introduced the "molecular facilitator" concept to describe the tetraspanins' general function. Later, the concept was further defined as "molecular organizer," "tetraspanin network," and "membrane microdomain," since compelling evidence suggested that tetraspanins

☆ Sequence data from this article have been deposited with the GenBank Data Library under Accession Nos. AY955256–AY955267.

* Corresponding author. Fax: +86 20 84038377.
E-mail address: ls36@zsu.edu.cn (A. Xu).

0888-7543/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2005.08.004

coupled with specific partners were capable of interacting with each other to form a web-like supramolecular complex on membranes and hence organizing a membrane microdomain of particular function [4–6,8–12].

Despite the general mechanism, specific functions of most tetraspanins remain unknown [4]. The major obstacles to the understanding of tetraspanins are their subtle functions and functional redundancy. First, except for several highly specialized members like retina-specific peripherins [13] and bladder-specific uroplakins [14], few tetraspanins have distinct phenotypes. For instance, deletion of leukocyte-specific CD37 yields no phenotype but slightly elevates the threshold of B cell activation and up-regulates T cell proliferation on TCR signaling [15,16]. Similarly, the absence of lymphocyte-specific TSPAN32 leads only to minor up-regulated T cell proliferation [17]. Second, in another scenario, the deletion of the widespread CD9 from mice results in no extensive phenotype but disrupts the sperm–egg fusion [18–20]. However, the injection of CD81 mRNA into the CD9-absent eggs could restore the fusion rate by 50% [21]. Since CD81 has a primary structure similar to that of CD9, functional compensation based on similarity would be the perfect explanation. Fruitfly tetraspanins also demonstrate some sort of redundancy: Lbm, Tsp42Ee, and Tsp42Ej are involved in synapse formation of motor neurons and loss of Lbm delays the formation of motor neurons, and the phenotype becomes more severe when all three tetraspanins are deleted simultaneously [22]. Surprisingly, fruitflies can live normally when nine tetraspanin genes are removed simultaneously [22]. With more cases like that, people believe that functional subtlety and redundancy are common properties of tetraspanins [4–6].

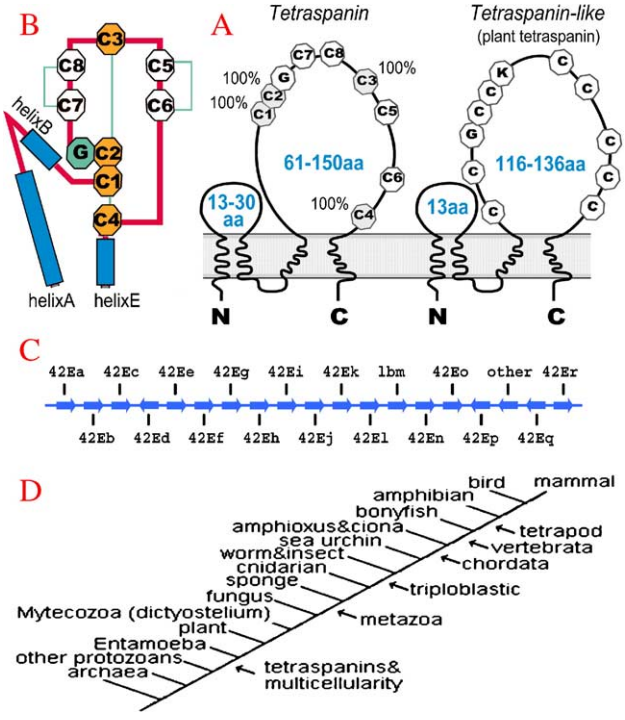

Fig. 1. Schematic diagrams of the tetraspanin structure, the tetraspanin cluster in *D. melanogaster,* and the phylogeny of eukaryotes. (A) Schematic structure of tetraspanin and tetraspanin-like protein (plant proteins), adapted from Olmos et al. ([28], reproduced by permission of the publisher). C1–C8 indicate the conserved cysteines; 100% conserved cysteines are labeled; "CCG" is the so-call tetraspanin signature. (B) The general EC2 structure of tetraspanin, adapted from Seigneuret et al. ([3], reproduced by permission of the publisher). The thin green line indicates the disulfide bonds. (C) Schematic diagram of the gene arrangement of the tetraspanin cluster in the 42E region of chromosome 2 of *D. melanogaster,* adapted from Todres et al. ([24], reproduced by permission of the publisher). Gene orientation is shown by the arrows; "other" is not a tetraspanin gene; picture is not drawn to scale. (D) Phylogenetic relationship of the eukaryotes; the last arrow indicates where tetraspanins and multicellularity emerged.

Table 1
Classification of the tetraspanin EC2s

Type	Subtype	Cysteine pattern of EC2	Example
4-Cys	4a	CCG–[DN][WY]–PXXC–GC	CD53
	4b	Types between 4a and 4c	NP_586477
	4c	CCG–C–C	EhTSPs\CD9
5-Cys	5b1	CCG–[DN][WY]–PXXCXC–GC	Ce-TSP-6
6-Cys	6a	CCG–[DN][WY]–PXXCC–C–GC	CD151
	6b1	CCG–[DN][WY]–PXXCXC–C–GC	TSPAN8
	6b2	CCG–[DN][WY]–PCXC–C–GC	CD82\CD37
	6c	Types other than 6a, 6b1, 6b2	BW484458
8-Cys	–	CCG–[DN][WY]–C–C–PXXCC–C–GC	TSPAN5

Adapted and modified from Hemler [4] and Seigneuret et al. [3]. EC2s were classified according to the number of cysteines and some other critical patterns in their structures; these cysteines are supposed to form disulfide bonds.

Having uniform architecture and conserved motifs, tetraspanins are supposed to rise from a common ancestor [6,23]. However, no confident phylogenetic relation could be rebuilt to represent the evolution of tetraspanins in either human or fruitfly [6,24]. Hemler mentioned that fruitfly had only two tetraspanin orthologs in mammals or nematodes, and each phylum had its exclusive tetraspanins [4]. Recently, a new class of tetraspanin was identified from fungi [25], which was the most distant member ever reported. Despite these reports, little effort was dedicated to the phylogeny of tetraspanins and no attention was paid to the coevolution between tetraspanins and eukaryotes. In this paper, we obtained tetraspanin sequences from online resources as well as our amphioxus cDNA library and reconstructed their phylogenetic relationships. Basing on the results, we discuss the origin of tetraspanin, the structural evolution of EC2, tetraspanins' proliferation and potential functional redundancy, and the connections between the evolution of tetraspanins and that of intercellular interactions.

Results

Quest for the origin of tetraspanin

Two new tetraspanin families from fungi

Prior to this paper, *PLS1* was the only tetraspanin family reported in fungi [25]. This gene is indispensable for *Magnaporthe grisea* to invade its host plant leaf [26]. In our study, an extensive search was performed for new tetraspanins in fungal genomes and our search yielded 10 more tetraspanins (Fig. 2). Four of them belong to the *PLS1* family; another 5 of them comprise a new family (termed TSP2); the last one forms the third family (termed EcTSP). Structural analysis and exon–intron analysis proved that they were genuine tetraspanins (Supplementary Fig. 1). As to

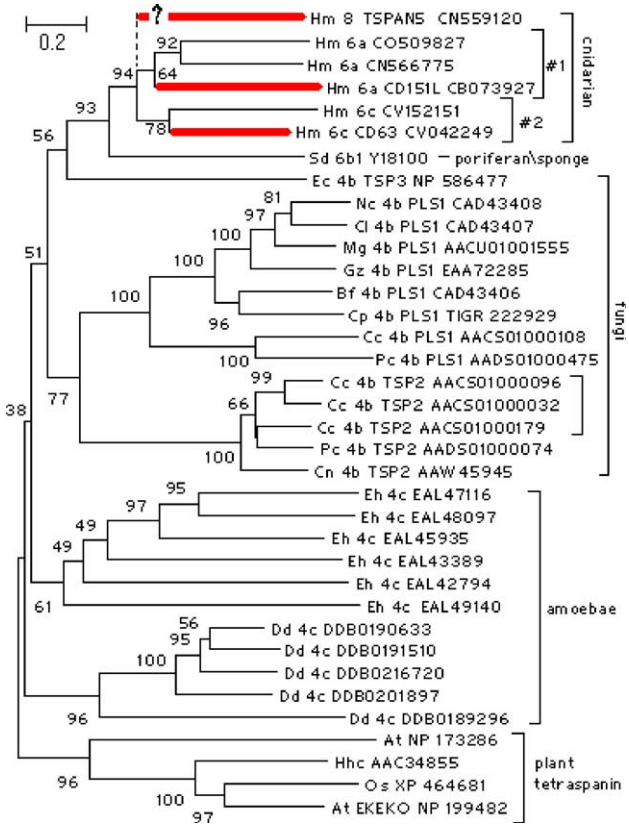

Fig. 2. Phylogenetic analysis of the early tetraspanins. Red bold line indicates genes that have vertebrate orthologs. ? shows that cnidarian TSPAN5 is not full-length; dashed line indicates the presumed position of cnidarian TSPAN5. "#1" indicates a tetraspanin subfamily of cnidarian; three more genes also belong to this family but are not shown here because of incomplete CDS. "#2" indicates another subfamily; two more genes belong to it but are not shown here because of incomplete CDS. Abbreviations used: At, plant *Arabidopsis thaliana;* Os, *Oryza sativa,* rice; Hbc, *Hemerocallis* hybrid cultivar, daylily; Eh, amoeba *Entamoeba histolytica,* parasite; Dd, amoeba *Dictyostelium discoideum;* Nc, fungus *Neurospora crassa;* Ec, fungus *Encephalitozoon cuniculi;* Cl, fungus *Colletotrichum lindemuthianum;* Mg, fungus *Magnaporthe grisea;* Gz, fungus *Gibberella zeae;* Bf, fungus *Botryotinia fuckeliana;* Cp, fungus *Coccidioides posadasii;* Cc, fungus *Coprinopsis cinerea;* Pc, fungus *Phanerochaete chrysosporium;* Cn, fungus *Cryptococcus neoformans;* Sd, sponge *Suberites domuncula;* Hm, cnidarian *Hydra magnipapillata.* (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

exon–intron configuration, one intron site is shared by both *PLS1* and TSP2, and this site is also common in metazoan tetraspanins, indicating their common origin.

Fungi consist of five major subphyla, and each has sequenced genomes available. In our search, *PLS1* was found in two subphyla, *Basidiomycota* and *Pezizomycotina,* whereas TSP2 was present only in *Basidiomycota* and EcTSP was exclusive to *Microsporidia.* Notably, three types of TSP2 were found in *Coprinopsis cinerea.* However, no tetraspanin could be recognized from 17 genomes of unicellular fungi from the other two subphyla, *Saccharomycotina* and *Schizosaccharomycetes* (7 complete genomes including yeast *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*). Since tetraspanins already existed before fungi (discussed later), it is likely that tetraspanin was lost in these unicellular fungi.

Multiple tetraspanins present in protozoan amoebae

At least six tetraspanins (termed EhTSPs) were also identified from the finished genome of *Entamoeba histolytica,* a specialized amoeba that infects human. EhTSPs are highly distant from metazoan or fungal tetraspanins (<22% identity, BLASTP) and even distant from each other (average of <25% identity, BLASTP). However, EhTSPs possess all the tetraspanin hallmarks, such as four TMs, a small EC1, a large EC2, the tetraspanin signature CCG, four 100% conserved cysteines, and the helix A/B/E in EC2 (Supplementary Fig. 1). As for the intron arrangement, four introns in three sites are present in EhTSPs, though introns are scarce in *E. histolytica* [27]. Two intron sites are conserved in metazoan tetraspanins, and one of them (X↓0XCCG), the most common site in animals, is conserved throughout *E. histolytica,* fungi, and metazoans. Hence, these two ancient introns confirmed EhTSPs as bona fide tetraspanins.

However, since *E. histolytica* is a parasite, we could not exclude the possibility that these molecules are just mimics of the hosts' tetraspanins. So we resorted to a free-living amoeba, *Dictyostelium discoideum,* and we discovered five more tetraspanins from it. Interestingly, these sequences have no so-called tetraspanin signature CCG, but structural analysis and intron analysis justified their identity (Supplementary Fig. 1). We also applied our search scheme to other protozoans, but no more tetraspanins were found, even in the complete genomes of *Leishmania major* and malaria *Plasmodium falciparum.* Therefore, either our search method is not sensitive enough or tetraspanin is not present in these protozoans at all.

Tetraspanin-like proteins in plants share common origin with tetraspanins

Our search scheme yielded a large quantity of sequences from plants, but they all belong to the tetraspanin-like family reported previously [28]. This family has more than 17 members in *Arabidopsis thaliana.* It is also present in rice, maize, daylily, and even mosses. A member of this family, *EKEKO,* will disturb leaf cell differentiation and lead to developmental abnormalities if mutated [28]. Though almost no similarity alignment could be made with tetraspanins, structurally these proteins resemble tetraspanins: four TMs, a small EC1, a large EC2, and a set of cysteine motifs in EC2 (Fig. 1A). One conspicuous difference is that this family does not have the tetraspanin signature CCG in EC2 but a GCCK instead [28]. Another difference is that this family has a much longer segment in between TM3 and the GCCK motif (about 30 aa longer than that of metazoan tetraspanins). Despite these discrepancies, this family is closer to tetraspanin than to other TM4 proteins, such as claudins, occludins, and L6 proteins [28].

Our analysis further supported that the tetraspanin-like family was a distant branch of tetraspanin (Supplementary Fig. 1). In the *A. thaliana* genome, two new members of this family (NP_173286 and NP_683494) were found. Instead of GCCK, they have the YCCA motif, which is closer to the tetraspanin signature CCG, and they have a much shorter segment in between TM3 and YCCA motif (20 aa shorter). On

the other hand, the tetraspanin signature CCG is not actually 100% conserved, for not only were CCY, CCK, and CCC found in amoeba tetraspanins but also CCA was discovered in two animal tetraspanins, XM_311541 and CD526816, from mosquito and Pacific oyster, respectively. Moreover, structure prediction showed that tetraspanin-like proteins have equivalents to tetraspanin's helix A/B/E in EC2. Finally, the most compelling support is that the only intron site of this family (X↓0XGCCK or X↓0XYCCA) corresponds to the most ancient and frequent intron site of tetraspanins. Altogether, plant tetraspanin-like and animal tetraspanin should have a common origin and we proposed that plant tetraspanin-like proteins were actually plant tetraspanins.

Tetraspanins in porifera and cnidaria

The metazoan (animal) ancestor was split into three major lineages, sponges (porifera, the real multicellular animal), diploblasts (cnidarians, having two geminal layers), and triploblasts (worms, insects, and mammals, having three geminal layers), of which the di- and triploblasts form a monophylogenetic group (Fig. 1D). So far, only one tetraspanin (6b1 type) was identified in sponges, but we believe there will be more since sponges are capable of complicated cell–cell interactions. This molecule was reported to participate in tight junction formation together with the scaffold protein MAGI [29,30]. The diploblastic cnidarian is another early diverged metazoan and at least 12 distinct tetraspanin sequences were found in our search. These sequences belong to three subfamilies, 6a type, 6c type, and a sequence of 8-Cys type. Tetraspanins from sponges and cnidarians are typical metazoan tetraspanins, because they are highly similar to vertebrate tetraspanins and have six or eight cysteines in EC2, as is exclusive to animal tetraspanins.

Phylogenetic approach to the origin of tetraspanins

Based on the alignment in Supplementary Fig. 1, we rebuilt the phylogenetic relationship of the early tetraspanins (Fig. 2). Using this and other information, several conclusions could be drawn. First, the history of tetraspanins can be dated back at least to the last common ancestor of amoebae, plants, and metazoans/fungi (Fig. 1D). Second, multiple tetraspanins are present in amoebae, plants, fungi, and metazoans, and each phylum develops its exclusive tetraspanin repertoire. Third, orthologs of vertebrate CD63 (6-Cys type) and TSPAN5 (8-Cys type) may already exist in the last ancestor of the metazoan lineage. Fourth, the cnidarian CD151-like molecule shares 40% identity with *Xenopus* CD151, but is clustered with vertebrate CD63 in phylogenetic analysis, suggesting that CD151 may have derived from CD63 before the speciation of cnidarians and functionally diverged later. Fifth, cnidarian TSPAN5 is a partial sequence, but structurally and phylogenetically it is orthologous to the molecules of the 8-Cys lineage in vertebrate. In addition, the sponge tetraspanin was recognized as CD63 previously [31], but our phylogenetic analysis did not support such assignment. BLASTP showed that this gene was closer to cnidarian CD151-like (26% identity) than to cnidarian CD63, but this protein has the YTTV endocytic signal at the end of the C-terminal, a characteristic of CD63.

Worms' tetraspanins are highly divergent

C. elegans possesses 20 tetraspanins and schistosomes have more than 25. However, worms' tetraspanins not only are distant from those of other phyla but also are highly divergent from each other. Consequently, phylogenetic analysis is difficult and less informative. However, phylogenetic analysis (Supplementary Fig. 2) still provided some interesting information: (1) worms' tetraspanins can be divided into several distinct groups, including CD63-like, CD151-like, and 8-Cys tetraspanins; a Sj25/TE736 family, and a group that is weakly similar to vertebrate CD9/CD82/TSPAN4/TSPAN8/TSPAN18; (2) two tetraspanins, Sm23 and CeTsp-12, are apparently orthologous to vertebrate CD63 and TSPAN5, respectively, and another two are highly similar to TSPAN8 and CD151 (CD082381 and CD078924). Note that though schistosomes are parasites, Sm23 has the YENV endocytic signal (a CD63 characteristics) at the end of the C-terminal, so it could not be a mimic of the vertebrate CD63.

Too many tetraspanins for Drosophila melanogaster?

The phylogenetic analysis of insect tetraspanins is shown in Fig. 3. Seventeen orthologous families (No. 1–No. 17 in Fig. 3) are supposed to share between at least two insect orders or between insects and vertebrates. As in worms, insect tetraspanins could be divided into four major groups: CD63-like, CD151-like, 8-Cys tetraspanins, and a group weakly similar to CD9/CD82/TSPAN4/TSPAN8/TSPAN18. Notably, 5 families have corresponding orthologs in vertebrates, including CD63, CD151, TSPAN5, TSPAN7, and TSPAN31 (Nos. 2, 5, 8, 13, 16 in Fig. 3); 3 families (Nos. 3, 4, 17 in Fig. 3) are highly divergent from other insect and noninsect tetraspanins, suggesting that they have specialized roles.

As shown in Fig. 3, *D. melanogaster* has at least 36 tetraspanins, but only 15 of them could be found in other insects, suggesting that *D. melanogaster* has up to 21 exclusive tetraspanins. Actually other insects like honeybee and mosquito also have their exclusive sets of tetraspanins (not shown on the tree), but they seem not to have so many tetraspanins. *D. melanogaster* has only about 14,000 genes but has at least 36 tetraspanins. This number is more than that of *C. elegans,* which has 19,000 genes but only 20 tetraspanins, and is even more than that of vertebrates, whereas vertebrates have about two to three times the gene number of *D. melanogaster.* Thus, where did these fruitfly tetraspanins come from?

As previously reported [24], 18 fruitfly-exclusive (or specific) tetraspanins come from the famous tetraspanin cluster (DmTSP42E). This gene cluster resides in a <30-kb interval at the 42E region of *D. melanogaster* and contains half of the 36 tetraspanins (Fig. 1C). Since neither ortholog nor similar gene cluster was found in the draft genomes of *Anopheles gambiae* (a malaria-transmitting mosquito), *Bombyx mori* (silk worm), or *Apis mellifera* (honeybee), this cluster is truly specific to *D.*

melanogaster and probably originated through rapid and repeated tandem duplications. Consistent with this origin, gene deletion indicated that half the tetraspanins of this cluster are functionally insignificant ([22]; discussed later). On the other hand, the presumed orthologs for this cluster could be found in the unfinished genomes of two other fruitflies (*D. yakuba* and *D. pseudoobscura*) and in a distant relative of the fruitfly, *Glossina morsitans;* therefore this cluster probably developed within the fly order. Should this be the case, the gene prototype for this cluster was assumed to be a CD63-like tetraspanin (boxed and linked in Fig. 3). Considering the fact that fruitfly suffered from severe gene loss, whereas its tetraspanin families are unexpectedly large, it is tempting to speculate that it was some sort of gene compensation.

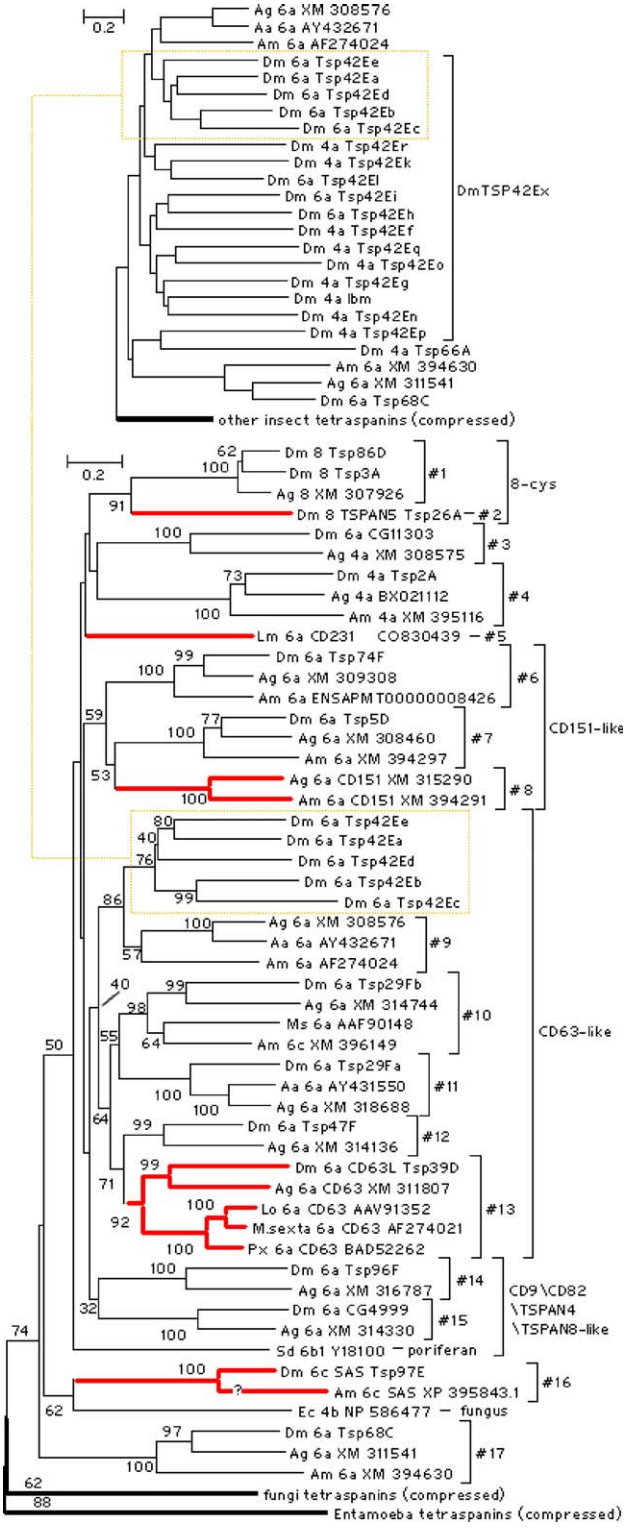

Fig. 3. Phylogenetic analysis of the insect tetraspanins. ? indicates the gene is not full length. Red bold line indicates the gene that has orthologs in vertebrate. #1–#17 indicate 17 families that are supposed to be shared by insects. Families of #3/#4/#17 are highly divergent from both other insect tetraspanins and vertebrate tetraspanins. The upper tree shows the phylogenetic topology of the 18 tetraspanins from the 42E region of *D. melanogaster;* rectangles in yellow indicate the probable prototypes for this cluster. Abbreviations used: Ec, fungus *Encephalitozoon cuniculi;* Sd, sponge *Suberites domuncula;* Dm, *D. melanogaster;* Ag, *Anopheles gambiae,* mosquito; Am, *Apis mellifera,* honeybee; Lo, *Lonomia obliqua;* Px, *Plutella xylostella,* diamondback moth; Ms, *Manduca sexta,* tobacco hornworm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Phylogenetic analysis of vertebrate tetraspanins

In our study 33 tetraspanins were found in human, 32 in mouse, 27 in chicken, 31 in *Xenopus tropicalis* (or *Xenopus laevis,* both clawed frogs), and 47 in bonyfish (*Danio rerio* or *Takifugu rubripes*). In addition, sequences from *Strongylocentrotus purpuratus* (sea urchin, 3 sequences), *Ciona intestinalis* (sea squirt, 28 sequences), and *Brachiostoma belcheri* (Chinese amphioxus, 12 cDNAs) were also taken into account considering their special evolutionary positions right at the root of the vertebrates (Fig. 1D). The phylogenetic tree of the vertebrate tetraspanins is shown in Supplementary Fig. 3. Based on the tree, vertebrate tetraspanins could be classified into 36 orthologous families and 23 of them could be definitely traced back to the invertebrates (Table 2). And notably, all 8-cysteine tetraspanins formed a monophyly, suggesting that they diverged early and evolved independently.

There are 13 orthologous families that have no invertebrate orthologs. Among them, CD9/CD81/TSPAN2, CD82/CD37, and TSPAN8 are the most mysterious, since they seem to emerge suddenly in vertebrates and bear special EC2 structures (4c, 6b2, and 6b1 type, respectively). If we eliminated the interference from the highly divergent members (TSPAN12 and TSPAN32), the specialized members (uroplakins and RDSs), and the members that diverged early and evolved independently (8-Cys tetraspanins and TSPAN31/TSPAN13), the phylogenetic analysis could show that TSPAN4/TSPAN9/CD53, TSPAN18/TSPAN1, CD9/CD81/TSPAN2, CD82/CD37, and TSPAN8 share a common origin (Supplementary Fig. 4). And their common origin is also supported by other evidence: first, as mentioned before, worms and insects each have a group of tetraspanins that are weakly similar to CD9/CD82/TSPAN4/TSPAN8/TSPAN18; although the similarities are too low to establish any orthology, they do imply where the ancestor of CD9/CD82/TSPAN4/TSPAN8/TSPAN18 came from; second, TSPAN4 and TSPAN18 have presumed orthologs in *S. purpuratus* and *Ci. intestinalis,* respectively; and third, a large TSPAN8-like family is present in *Ci. intestinalis*

Table 2
Distribution of vertebrate tetraspanins and their orthologs in other species

	cnidarian	worms	insect	seaurchin	Ciona & amphioxus	bonyfish	amphibian	bird	mammal	notes (function, expresion profile,etc)
CD63	√	√	√	√	√	√	√	√	√	binds integrins, widespread
CD151	?1	2?	√	√	√	√	√	√	√	binds integrins, widespread
TSPAN11						√			√	unknown
TSPAN5	?3	√	√	√	√	√	√	√	√	unknown, widespread
TSPAN17						√	√	√	√	unknown
TSPAN14						√	√	√	√	unknown
TSPAN7			√		√	√	√	√	√	mental retardation, widespread
TSPAN7-like (BX324164)*						√	√	√		unknown
TSPAN6						√	√	√	√	unknown
TSPAN31			√		√	√	√	√	√	unknown
TSPAN13						√	√	√	√	unknown
TSPAN4				√	√	√	√	√	√	unkown, widespread
TSPAN9						√	√	√	√	unknown, various tissues
CD53						√	√	√	√	adaptive immunity specific
TSPAN15					√	√	√	√	√	unknown, various tissues
TSPAN15-like (BU232147)*						√		√		unknown
TSPAN33					√	√	√	√	√	unknown
TSPAN18					√	√	√	√	√	nerve system
TSPAN1						√	√	√	√	unknown
TSPAN8	?4				?5	√	√	√	√	unknown
TSPAN3					√	√	√	√	√	unknown, widespread
TSPAN3-like (NM_001002748)*						√	√			unknown
TSPAN12					√	√	√	√	√	unknown, various tissue
CD82						√	√	√	√	immunity, widespread
CD37							√		√	adaptive immunity
CD9						√	√	√	√	adaptive immunity, brain, widspread
CD81						√	√	√	√	adaptive immunity, brain, widspread
TSPAN2						√	√	√	√	brain function,various tissues
CD9-like (NM_200552)*						√				unknown, mistaked as CD9 previous
TSPAN10						√		√	√	unknown, eye
URK1A						√	√	√	√	bladder specific
URK1B						√	√	√	√	bladder specific
RDS						√	√	√	√	retina specific
RDS35						√	√			retina specific
RDS2						√	√	√		retina specific
ROM1									√	retina specific
TSPAN32								√	√	leukocytes specific
TSPAN19									√	unknown, in human only
TSPAN16									√	unknown, in human only

All 33 human tetraspanins and all vertebrate tetraspanins that are present in at least two representative species are included. Genes of common origin are grouped together and indicated by alternate greens, "√" indicates the gene is found in the species. A black box indicates the gene is supposed to be present in the species although it is not found yet; a gray box indicates a co-orthologous relation; for instance, vertebrate CD151 and TSPAN11 are co-orthologs to the insect CD151, (*) Genes that have no other names available. (?1) Cnidarian CD151-like has support from BLAST (up to 40% identity to *Xenopus* CD151) but little support from phylogenetic analysis (?2) Schistosome CD151-like is not full length, though similar to vertebrate CD151 (35% identity); its identity is dubious (?3).

and amphioxus (Supplementary Fig. 5) and the best invertebrate homolog to CD9 is from this family. Taken together, CD82/CD37 and TSPAN8 should share a common origin with TSPAN4 and TSPAN18; with the analysis on the structural evolution of EC2 (discussed later), the ancestor of the CD9/CD81/TSPAN2 is supposed to derive from an ancient TSPAN8-like tetraspanin before the vertebrata radiation.

The phylogenetic tree also defined 12 paralogous families, which account for 32 vertebrate tetraspanins (Table 2). These paralogous families should have arisen from the duplications that occurred between the chordata radiation and the vertebrata radiation according to the phylogenetic analysis (Supplementary Fig. 3). Moreover, in bonyfish many tetraspanins are present in two copies (Supplementary Fig. 3), indicating another round of duplications. We then examined those tetraspanin loci in the human genome as well as in fish genomes and found that at least 10 paralogous families, including TSPAN5/TSPAN17/TSPAN14, TSPAN9/TSPAN4/CD53, CD9/CD81/TSPAN2, RDS/RDS35/CRDS2/ROM1, TSPAN15/TSPAN15-like, URK1A/URK1B, TSPAN7/TSPAN16, TSPAN13/TSPAN31,

CD37/CD82, and CD151/TSPAN11, are definitely produced by en bloc duplications, because their members are separated in the so-called paralogous genomic regions (data not shown). Taken together, the major expansion of the vertebrate tetraspanin repertoire fits well into the massive duplication hypothesis, in which one or two rounds of whole-genome duplication gave rise to the tetrapods and another round of massive duplication created the ray-finned fishes [32,33].

Discussion

Primary and secondary evolvements of the basic EC2 structure

Tetraspanin EC2 is the major part of the protein responsible for binding partner proteins. Assuming that a certain 3D structure recognizes a certain partner repertoire, then for various potential partners variations of EC2 are inevitable. Theoretically, variation of EC2 would be associated with three kinds of evolution events: (1) coevolution of major partners; (2) minor modification of partner repertoire, namely, gains or losses of some partners; and (3) duplication and subsequent alteration of the partner repertoire. So it is self-evident that the structural changes in EC2 reflect the changes in partner spectrum. To study the major structural changes of EC2, EC2s were classified according to their basic structural patterns (Table 1).

Each EC2 type should emerge in the sequence 4c → 4b → 4a → (5b1) → 6b1 → 6a/6c → 8-Cys, as shown by the phylogenetic trees in Fig. 2 and Supplementary Fig. 2. The 4b type (e.g., EcTSP) is the intermediate type between the ancient 4c type and the 4a type (e.g., TE736/Sj25 family). The 6b1 type should have derived from the 4a type and is the ancestor of both 6a/6c types and 8-Cys type, since structurally 6b1 type is easy to shift to 6a or 6c type by loss or gain of residues in between C3 and C5. And the 8-Cys type is supposed to derive from the 6a type. In addition, the presence of the 5b1 type (CeTSP-4/5/6) confirms the possible conversion between 4a and 6b1 type; and CeTSP12, an ortholog to TSPAN5 (8-Cys-type), suggests the conversion between 6a and 8-Cys type. We called these changes the primary evolvements. The primary evolvements probably reflect the need for more variation in EC2, because more variation could handle more partners. After all, 4-Cys EC2 is not stable enough to accommodate many residues, whereas more stabilizers, such as disulfide bonds, can contain more residues.

Phylogenetic analysis showed that 8-Cys tetraspanins originated very early and evolved independently (Supplementary Fig. 3). However, phylogenetic analysis failed to support the monophyly for other types of tetraspanins, indicating the existence of secondary evolvements.

Loss of extra cysteines (C5–C8) is the usual secondary change, such as 8-Cys → 6a (CeTSP12) and the more frequent 6a → 4a change (DmTSP42E family, TSPAN18s of sea squirt, CD53, etc.). The most typical example is CD53. In human and zebrafish, CD53 is a 4a type, whereas in frog and *Onchorynchus mykiss* (bonyfish) it is a 6a type, suggesting that CD53 could afford losing extra cysteines without affecting its orthology, namely, its major functions. Point mutations have demonstrated that different parts of CD9 EC2 are responsible for different partners [34,35]. Analogously, the lost part in human and zebrafish CD53 EC2 is probably responsible for minor partners or minor functions.

Another case of secondary change involved TSPAN8 (6b1), CD82/CD37 (6b2), and CD81/CD9/TSPAN2 (4c). As mentioned above, CD9/CD81/TSPAN2, CD82/CD37, TSPAN8, TSPAN4/TSPAN9/CD53, and TSPAN18/TSPAN1 form a monophyly and have distant homologs in worms, insects, sea squirt, and amphioxus. Superficially, members of this group are active in structural changing, in other words, they have the propensity to lose or gain cysteines, for example: CeTSP4/5/6 (5b1), CeTSP9 (4c), CD082381 (Sm-TSPAN8-like, 4a), TSPAN18 (4a and 6a), and CD53 (4a and 6a). Therefore, these structures' origin could be postulated: an ancestral gene of 6a type gave rise to TSPAN4, TSPAN18, and CD53, in which CD53 and TSPAN18 inherited the "losing cysteine habit"; CD82 (6b2) and TSPAN8 (6b1) were also descendants of this ancestral gene, but they may have at some point lost an extra cysteine (C5) and then gained a new one in a different position and become the 6b2 and 6b1 type (6a → 5a → 6b2/6b1). As for CD9, it was probably derived from an ancient TSPAN8-like member by losing two extra cysteines (6b1 → 4c), because there are some TSPAN8-like members in sea squirt and amphioxus that not only are highly similar to CD9 but also have CD9 features such as having no DW, PXSC, or GC motifs.

All the analyses above favor a modularized structure for EC2: in this structure helix A/B/E and four 100% conserved cysteines provide a robust "chassis," and the gain or loss of extra disulfide bonds and other changes in amino acid composition affect only the higher structures on the chassis. Disulfide bridges and other motifs further partition EC2 into subdomains, and different subdomains not only respond to different proteins, but also prevent changes within one subdomain from affecting other subdomains.

Proliferation and functional redundancy of tetraspanins

The function of tetraspanin largely relies on specific recognition; hence, for various potential partners multiple tetraspanins are inevitable. So it is not beyond expectation that the animal developed such a large family of tetraspanins. Having a uniform architecture, tetraspanins are apparently free of exon shuffling, intramolecular duplications, and drastic architectural changes; therefore evolution of this family is simply a process of duplication and mutation accumulation. If the mutation accumulation lags behind the duplication or fails to endow the duplicate with a new function, functional redundancy becomes inevitable. There are two levels of redundancy for tetraspanins: (1) functions of tetraspanins are unimportant and dispensable, such as Lbm/DmTsp42Ee/DmTsp42Ej, and (2) functions of tetraspanins are substitutable for one another due to sequence similarity, such as CD9/CD81.

From amoebae to amphioxus, organisms expanded their tetraspanin families through independent duplications. Most of these duplications focused on a few "hot spots" (in a hot spot duplication occurred at least twice and produced at least three paralogs), for instance, tetraspanins in amoebae, CcTSP2, the

TE736/Sj25 family in schistosomes, the CD63-like and CD151-like families in insects, the DmTSP42E cluster in *D. melanogaster,* and a large TSPAN8-like family in amphioxus (8 of all 12 amphioxus tetraspanins from our cDNA library fall into this family). These "inbred" proliferations unavoidably brought about the functional redundancy. The extreme example is the DmTSP42E cluster. In this cluster, not only *lbm,* Tsp42Ee, and Tsp42Ej exhibited functional compensation for each other in motor neuron development, but also the deletion of half the tetraspanins in the cluster did not noticeably disturb the fruitfly's life cycle [22].

In vertebrates at least 10 paralogous tetraspanin lineages (accounting for 2/3 vertebrate tetraspanins) were proved to generate by whole-genome duplications. Since these lineages are divergent from each other, the potential functional redundancy caused by similarity should be limited within each lineage. Compared with the situation in independent duplications, in massive duplications the coevolution of new proteins often rapidly leads to new partnerships, new complexes, and even new cell types. In this sense, functional redundancy between duplicates that arose in massive duplications would be better contained. Take CD9/CD81 as an example: albeit their functional compensation has been proved, differentiated expression profiles still make a great difference in the sperm–egg fusion (see the introduction and [36]). In addition, it has to be noted that unlike in nonvertebrates, few tetraspanins in vertebrates were produced by independent duplications after the massive duplications; for example, only three human tetraspanins may have come from independent duplications (TSSC6, TSPAN16, and TSPAN19). Considering that independent duplications are also frequent in vertebrates, such as Ig's, MHC molecules, galectins, and olfactory receptors, it is interesting to see that the proliferation of tetraspanins slows down in vertebrates.

Both nonvertebrates and vertebrates tend to produce and keep many tetraspanins, disregarding whether their functions are redundant or not. But due to different duplication modes, functional redundancy is more common in nonvertebrates in theory, or in other words, independent duplications are less efficient in producing novel-functioning tetraspanins compared with the massive duplications. However, despite the similar tetraspanins, those divergent members or lineages should have specific functions, such as TSP68C and Ce-TSP-15. If phenotypes of those divergent members are subtle, then other reasons should be considered.

Functions of the early tetraspanins

It is not simply a coincidence that multiple tetraspanins exist in two distant amoebae. Amoebae feature a cell membrane of complex dynamic morphology that is involved in cytokinesis, phagocytosis, motility, chemotaxis, and signal transduction [37]. Since tetraspanins have critical roles in defining membrane microdomains, we could easily associate tetraspanins with the amoeba's ever-changing membrane. Compared with the parasite *E. histolytica,* whose evolution could be affected by the parasitic life, *D. discoideum* is a well-characterized free-living amoeba. Because *D. discoideum* can normally live as separate unicells but interacts with others to form a multicellular structure when stimulated by adverse conditions like starvation, it is a model organism for the study of basic multicellular behaviors, such as developmental pattern formation, cell sorting, and cell-type differentiation [37]. Therefore, this amoeba provides the opportunity to unravel the basic tetraspanin secrets both in a unicellular context and in a multicellular context. Considering its upcoming genome and the ease with which it can be cultured and manipulated, *D. discoideum* shall be an excellent model for tetraspanin study.

In the fungal world, tetraspanin *PLS1* is indispensable for *M. grisea* and *Botrytis cinerea* to invade their host plant's leaves [26,38]. If *PLS1* is disrupted, the appressorium (the invasive cell) grows normally but fails to form the subcellular invasive apparatus, a "peg"-like hypha, to penetrate into the plant cell [26]. Therefore, *PLS1* is apparently required for the morphogenesis of the subcellular hypha, reminiscent of the pseudopod of amoeba and the filopodia of migrating human cells [12,39]. In the migrating MDA-MB-231 cell, tetraspanins and integrins form a complex at the tip of the filopodia [12]. As for *Encephalitozoon cuniculi,* a highly specialized unicellular parasite of mammals, it forms a tube-like invasive apparatus that can be extruded to inject the sporoplasm into the target cell, so it could be assumed that the EcTSP is retained for this invasion. In addition, since a large part of fungi are simple multicellular organisms, fungal tetraspanins may participate in multicellular behaviors, such as the formation of mycelia and fruiting bodies. Note that although these tetraspanins are all from parasitic fungi, neither does the *PLS1* resemble the plant tetraspanins nor the EcTSP simulate the mammal tetraspanins, indicating that their roles in invasion do not depend on mimicries of the hosts.

In plants, tetraspanin-like proteins are abundant and definitely underwent independent evolution. However, so far only one has been characterized, the *EKEKO* from *A. thaliana.* Disruption of *EKEKO* leads to round, tumor-like, multinucleated leaf cells congregating in the wrong places [28], reminiscent of the roles of animal tetraspanins in cell proliferation, migration, and transduction of differentiation signals. Because of the common origin with animal tetraspanins, plant tetraspanins should function by the same mechanism. However, given that plants have a strong cell wall and that plant tetraspanins have great structural difference, it is interesting to ask how these proteins mediate intercellular interactions and whether they are involved in the plasmodesmata formation and gamete fusion.

Tetraspanins coevolved with cell–cell (matrix–cell) interactions

Consistent with the data from physiology, biochemistry, molecular biology behavior, and developmental biology, the phylogenetic analysis of elongation factor-1a placed the amoeba *D. discoideum* among the metazoa/fungi and plantae groups [40]. Since this amoeba is able to live as unicells as well as to form true multicellular bodies, it lives in the twilight zone between unicellularity and multicellularity. In this sense,

tetraspanins of amoeba are expected to play important roles in the unicell-to-multicell transition. On the other hand, though multiple tetraspanins are found in the fungal world, they are actually lost in those unicellular fungi (except the specialized *E. cuniculi*), suggesting the correlation between tetraspanin and multicellularity. Indeed, it has been proven that tetraspanins are thriving in amoebae, multicellular fungi, metazoans, and even plants. Based on these facts, an evolutionary story may be proposed: tetraspanins came into being in the ancient unicellular eukaryotes with the dynamic membrane morphology, but soon were co-opted for cell–cell interactions and thus helped kick off the multicellular era.

In metazoans, evolution can be viewed as a force driving further differentiation of cell types and further complication of the corresponding cell–cell interactions. The phylogeny of tetraspanin is consistent with this process (Table 2). In vertebrates, new cell types and new systems came up with new tetraspanins, such as TSPAN32/CD37/CD53 for leukocytes, CD9/CD81/TSPAN2 for brain and adaptive immune system, peritherin/RDS for retina, and uroplakins for bladder. Interestingly, the emergence of these new tetraspanins could be attributed to the massive duplications that shaped the modern vertebrates. As to the invertebrates, each phylum has exclusive tetraspanins, too. These phylum-specific tetraspanins, especially those highly divergent members and lineages, may correspond to phylum-specific cell–cell interactions. For instance, Ce-TSP-15 is critical for the epidermal integrity of nematodes [41] and TSP68C has a role in regulating the proliferation of insect hemocytes [42]. Actually, half of the nematode tetraspanins and two other insect-specific families (Nos. 3, CG11303, and 4, TSP2A; Fig. 3) are highly divergent and likely assume specific roles.

In theory, as an "organizer" of the functional microdomain, tetraspanin has the potential to make the impossible process possible or make the possible more efficient and adaptable. Thus unlike other proteins of limited implication, like a catalyst tetraspanin has potential functions for any possible process on membrane. That may be why tetraspanins are found in various cell–cell interactions and why organisms produced and kept many tetraspanins whether they were redundant or not. Unlike immunoglobulins, integrins, and other protein families whose structures vary in domain composition, tetraspanins maintain identical architectures of four TMs and two ECs and preserve proper conservation throughout evolution, which makes their evolution easy to follow. Furthermore, the functional mechanism based on specific recognition makes tetraspanin the faithful marker of the corresponding partners, functional complexes, and intercellular interaction events. Taken together, studying tetraspanins in a phylogenetic context should provide an approach to following the evolution of intercellular interactions in an integral manner.

Conclusion

In this paper, we attempted to reconstruct tetraspanins' phylogeny and evolution scenario to provide a framework for further study. We discussed the evolution of animal tetraspanins in great detail. but not that of plant tetraspanins. We discussed the evolution of EC2's basic structure, while we could not understand those structural features in amoeba and plant tetraspanins, nor did we discuss other domains of the tetraspanins. We speculated that tetraspanins are required for the amoeba's dynamic morphology and we also presumed that tetraspanins are necessary for the formation of multicellular structures, yet both hypotheses require extensive experimental validation. Nevertheless, based on our extensive analysis we believe that tetraspanins have been coevolving with the eukaryotes' intercellular interactions from the beginning. Compared with other bioprocesses, intercellular interactions are evolving on the fast lane, so to study tetraspanins in a phylogenetic context should help track down the evolution of cell–cell and matrix–cell interactions.

Materials and methods

Cloning tetraspanins from Chinese amphioxus

Six cDNA libraries (ovary, neurula, gastrula, larva, and intestine) of Chinese amphioxus (*Branchiostoma belcheri tsingdaunese*) were constructed and sequenced as described [43]. All tetraspanin cDNA clones were identified by RPSBLAST and subjected to further sequencing. All sequences were deposited with GenBank (Accession Nos. AY955256–AY955267).

Data preparation from representative organisms

Genomic data were downloaded from Ensembl (http://www.ensembl.org/) or JGI (http://www.jgi.doe.gov/), and nucleotide data were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/). ESTs were assembled into consensuses with STACKPACK v2.0 before use.

Identification and verification of tetraspanin sequences

Stand-alone RPS-BLAST (Model pfam00335) was used to identify tetraspanins from downloaded data. For some fungi, plants, and protozoans, search was performed on the NCBI BLAST server; various sequences were used as bait, including consensus from PSSM Model pfam00335. A PSSM model generated from PSI-BLAST may be used if possible. For the amoeba *D. discoideum,* search was performed on http://dictybase.org/db/cgi-bin/blast.pl. For some fungal genomic results, cDNAs were predicted by FGENGSH (http://www.softberry.com/). Tetraspanins were verified and annotated according to the results of the search against GenBank with NET-BLAST.

Alignment and phylogenetic study

Topological information was predicted by the TMHMM2.0 server (http://www.cbs.dtu.dk/-services/TMHMM-2.0/) and used to ensure that four TMs and two ECs were aligned correctly. BLAST2SEQ was used to produce pair-wise alignments. ClustalW 1.83 was used in multiple alignments. Alignments were edited and refined with GeneDoc software. The phylogenetic tree was built with Mega v2.1. The neighbor-joining method was used to calculate the trees, with 1000 bootstrap tests and handling gaps with pair-wise deletion.

In the phylogenetic analysis, if not mentioned explicitly, only those sequences that had four TMs and two ECs were used. To improve the alignment, long and highly variable N-/C-termini were deleted; in other words, only the "tetraspanin core" (4 TM + 2 EC) was used for alignment and tree calculation. In the tree-building process, highly divergent or unstable sequences were cut away one by one recursively. In the end, if possible, these cut-away sequences were mapped onto the final tree to show their presumable positions, and the given bootstrap values for these cut-away sequences were from the overall tree building that included all sequences.

Special source of some sequences

Six tetraspanins of the amoeba *D. discoideum* were obtained from DictyBase (http://dictybase.org/). Some fungal tetraspanins were predicted proteins of genomic sequences, such as *Phanerochaete chrysosporium* PLS1, which was predicted from genomic sequence AACU01001555; one fungal tetraspanin, PLS1 of *Coccidioides posadasii,* was predicted from an unfinished fragment of genome (ID: TIGR_222929). Two sequences were from the zebrafish genome (http://www.ensembl.org/Danio_rerio/, ENSDART00000035079 and ENSDART00000002217). Four sequences were from the *T. rubripes* (pufferfish) genome (http://www.ensembl.org/Fugu_rubripes/, SINFRUT00000140037, SINFRUT00000156433, SINFRUT00000133051, and SINFRUT00000133548). Five sequences were from the *X. tropicalis* (western clawed frog) genome (http://genome.jgi-psf.org/Xentr3/Xentr3.home.html, 8463, 5320, 7507, 6893, and 4585). Two sequences were from the *Ci. intestinalis* (sea squirt) genome (http://genome.jgi-psf.org/ciona4/ciona4.home.html, ci0100147944 & ci0100141936).

Acknowledgments

We thank Jiantao Huang for exquisite artworks. This work was supported by grant from State High-Tech Development Project (863) of Ministry of Science and Technology of China (2005AA626011, 2004AA621030, 2003AA626010), Key Project (0107) of Ministry of Education, Project (30300264) of National Natural Science Foundation, and Key Projects of Commission of Science and Technology of Guangdong Province and Guangzhou City.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ygeno.2005.08.004.

References

[1] C.S. Stipp, T.V. Kolesnikova, M.E. Hemler, Functional domains in tetraspanin proteins, Trends Biochem. Sci. 28 (2003) 106–112.
[2] K. Kitadokoro, et al., CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs, EMBO J. 20 (2001) 12–18.
[3] M. Seigneuret, A. Delaguillaumie, C. Lagaudriere-Gesbert, H. Conjeaud, Structure of the tetraspanin main extracellular domain: a partially conserved fold with a structurally variable domain insertion, J. Biol. Chem. 276 (2001) 40055–40064.
[4] M.E. Hemler, Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain, Annu. Rev. Cell Dev. Biol. 19 (2003) 397–422.
[5] M.E. Hemler, Specific tetraspanin functions, J. Cell Biol. 155 (2001) 1103–1107.
[6] C. Boucheix, E. Rubinstein, Tetraspanins, Cell. Mol. Life Sci. 58 (2001) 1189–1205.
[7] H.T. Maecker, S.C. Todd, S. Levy, The tetraspanin superfamily: molecular facilitators, FASEB J. 11 (1997) 428–442.
[8] M.D. Wright, G.W. Moseley, A.B. van Spriel, Tetraspanin microdomains in immune cell signalling and malignant disease, Tissue Antigens 64 (2004) 533–542.
[9] M. Yunta, P.A. Lazo, Tetraspanin proteins as organisers of membrane microdomains and signalling complexes, Cell Signalling 15 (2003) 559–564.
[10] J.M. Tarrant, L. Robb, A.B. van Spriel, M.D. Wright, Tetraspanins: molecular organisers of the leukocyte surface, Trends Immunol. 24 (2003) 610–617.
[11] F. Le Naour, et al., Tetraspanins connect several types of Ig proteins: IgM is a novel component of the tetraspanin web on B-lymphoid cells, Cancer Immunol. Immunother. 53 (2004) 148–152.
[12] F. Berditchevski, E. Odintsova, Characterization of integrin–tetraspanin adhesion complexes: role of tetraspanins in integrin signaling, J. Cell Biol. 146 (1999) 477–492.
[13] C.J. Loewen, O.L. Moritz, R.S. Molday, Molecular characterization of peripherin-2 and rom-1 mutants responsible for digenic retinitis pigmentosa, J. Biol. Chem. 276 (2001) 22388–22396.
[14] F.X. Liang, et al., Organization of uroplakin subunits: transmembrane topology, pair formation and plaque composition, Biochem. J. 355 (2001) 13–18.
[15] K.P. Knobeloch, et al., Targeted inactivation of the tetraspanin CD37 impairs T-cell-dependent B-cell response under suboptimal costimulatory conditions, Mol. Cell. Biol. 20 (2000) 5363–5369.
[16] A.B. van Spriel, et al., A regulatory role for CD37 in T cell proliferation, J. Immunol. 172 (2004) 2953–2961.
[17] J.M. Tarrant, et al., The absence of Tssc6, a member of the tetraspanin superfamily, does not affect lymphoid development but enhances in vitro T-cell proliferative responses, Mol. Cell. Biol. 22 (2002) 5006–5018.
[18] K. Kaji, et al., The gamete fusion process is defective in eggs of Cd9-deficient mice, Nat. Genet. 24 (2000) 279–282.
[19] F. Le Naour, E. Rubinstein, C. Jasmin, M. Prenant, C. Boucheix, Severely reduced female fertility in CD9-deficient mice, Science 287 (2000) 319–321.
[20] K. Miyado, et al., Requirement of CD9 on the egg plasma membrane for fertilization, Science 287 (2000) 321–324.
[21] K. Kaji, S. Oda, S. Miyazaki, A. Kudo, Infertility of CD9-deficient mouse eggs is reversed by mouse CD9, human CD9, or mouse CD81: polyadenylated mRNA injection developed for molecular analysis of sperm–egg fusion, Dev. Biol. 247 (2002) 327–334.
[22] L.G. Fradkin, J.T. Kamphorst, A. DiAntonio, C.S. Goodman, J.N. Noordermeer, Genomewide analysis of the Drosophila tetraspanins reveals a subset with similar function in the formation of the embryonic synapse, Proc. Natl. Acad. Sci. USA 99 (2002) 13663–13668.
[23] V. Horejsi, C. Vlcek, Novel structurally distinct family of leucocyte surface glycoproteins including CD9, CD37, CD53 and CD63, FEBS Lett. 288 (1991) 1–4.
[24] E. Todres, J.B. Nardi, H.M. Robertson, The tetraspanin superfamily in insects, Insect Mol. Biol. 9 (2000) 581–590.
[25] M. Gourgues, et al., A new class of tetraspanins in fungi, Biochem. Biophys. Res. Commun. 297 (2002) 1197–1204.
[26] P.H. Clergeot, et al., PLS1, a gene encoding a tetraspanin-like protein, is required for penetration of rice leaf by the fungal pathogen Magnaporthe grisea, Proc. Natl. Acad. Sci. USA 98 (2001) 6963–6968.
[27] U. Wilihoeft, E. Campos-Gongora, S. Touzni, I. Bruchhaus, E. Tannich, Introns of Entamoeba histolytica and Entamoeba dispar, Protist 152 (2001) 149–156.
[28] E. Olmos, B. Reiss, K. Dekker, The ekeko mutant demonstrates a role for tetraspanin-like protein in plant development, Biochem. Biophys. Res. Commun. 310 (2003) 1054–1061.
[29] W.E. Muller, H. Schwertner, I.M. Muller, Porifera a reference phylum for evolution and bioprospecting: the power of marine genomics, Keio J. Med. 53 (2004) 159–165.
[30] T. Adell, et al., Evolution of metazoan cell junction proteins: the scaffold protein MAGI and the transmembrane receptor tetraspanin in the demosponge Suberites domuncula, J. Mol. Evol. 59 (2004) 41–50.
[31] W.E. Muller, et al., Initiation of an aquaculture of sponges for the sustainable production of bioactive metabolites in open systems: example, Geodia cydonium, Mar. Biotechnol. (NY) 1 (1999) 569–579.
[32] L. Abi-Rached, A. Gilles, T. Shiina, P. Pontarotti, H. Inoko, Evidence of en bloc duplication in vertebrate genomes, Nat. Genet. 31 (2002) 100–105.
[33] S. Ohno, Gene duplication and the uniqueness of vertebrate genomes circa 1970–1999, Semin. Cell Dev. Biol. 10 (1999) 517–522.
[34] H. Hasuwa, et al., CD9 amino acids critical for upregulation of diphtheria toxin binding, Biochem. Biophys. Res. Commun. 289 (2001) 782–790.

[35] G.Z. Zhu, et al., Residues SFQ (173–175) in the large extracellular loop of CD9 are required for gamete fusion, Development 129 (2002) 1995–2002.
[36] K. Kaji, S. Oda, S. Miyazaki, A. Kudo, Infertility of CD9-deficient mouse eggs is reversed by mouse CD9, human CD9, or mouse CD81: polyadenylated mRNA injection developed for molecular analysis of sperm–egg fusion, Dev. Biol. 247 (2002) 327–334.
[37] A. Kuspa, R. Sucgang, G. Shaulsky, The promise of a protist: the Dictyostelium genome project, Funct. Integr. Genom. 1 (2001) 279–293.
[38] M. Gourgues, A. Brunet-Simon, M.H. Lebrun, C. Levis, The tetraspanin BcPls1 is required for appressorium-mediated penetration of Botrytis cinerea into host plant leaves, Mol. Microbiol. 51 (2004) 619–629.
[39] L. Eichinger, S.S. Lee, M. Schleicher, Dictyostelium as model system for studies of the actin cytoskeleton by molecular genetics, Microsc. Res. Tech. 47 (1999) 124–134.
[40] S.L. Baldauf, W.F. Doolittle, Origin and evolution of the slime molds (Mycetozoa), Proc. Natl. Acad. Sci. USA 94 (1997) 12007–12012.
[41] H. Moribe, et al., Tetraspanin protein (TSP-15) is required for epidermal integrity in Caenorhabditis elegans, J. Cell Sci. 117 (2004) 5209–5220.
[42] S.A. Sinenko, B. Mathey-Prevot, Increased expression of Drosophila tetraspanin, Tsp68C, suppresses the abnormal proliferation of ytr-deficient and Ras/Raf-activated hemocytes, Oncogene 23 (2004) 9120–9128.
[43] C.Y. Mou, et al., EST analysis of mRNAs expressed in neurula of Chinese amphioxus, Biochem. Biophys. Res. Commun. 299 (2002) 74–84.