Invited Review

ACTA ZOOLOGICA BULGARICA
Acta zool. bulg., 67 (2), 2015: 305-317

New Insights into the Molecular Evolution of Metazoan Peroxiredoxins

***Rigers Bakiu*[1*], *Gianfranco Santovito*[2]**

[1] Department of Aquaculture and Fisheries, Agricultural University of Tirana, Koder Kamez, 1029 Tirana, Albania; E-mail: rigers.bakiu@ubt.edu.al
[2] Department of Biology, University of Padova, 35121 Padova, Italy

Abstract: Peroxiredoxins (Prx) are enzymes present in all biological kingdoms, from bacteria to animals. The oxidised active site cysteine of Prx can be reduced by a cellular thiol, thus enabling Prx to function as a peroxidase. Peroxiredoxins have been object of an increasing interest for its pivotal role in cell defence and as conserved markers for circadian rhythms in metabolism across all three phylogenetic domains (Eukarya, Bacteria and Archaea). Metazoan cells express six Prx isoforms that are localised in various cellular compartments. Using bioinformatics tools, based on Bayesian approach, we analysed the phylogenetic relationships among metazoan Prxs, with the aim to acquire new data on the molecular evolution of these proteins. Peroxiredoxin molecular evolution analyses were performed by the application of Mr. Bayes and HyPhy software to the coding and protein sequences of deuterostomes and protostomes. The obtained results confirmed that the molecular evolution of metazoan Prx was peculiar and suggested that the positive selection may had operated for the evolution of these proteins and a purifying selection was present during this process.

Keywords: peroxiredoxin; metazoans; molecular evolution.

Introduction

An unusual antioxidant protein, now called peroxiredoxin (Prx, EC 1.11.1.15), was initially identified on the basis of its capacity to protect proteins from oxidative damage caused by reactive oxygen species (ROS), catalysing the reduction of H_2O_2, to water and alcohol in the presence of dithiotreitol (DTT). Analysis of Prx purified from yeast revealed that it did not contain conventional redox centres such as metals, heme, flavin, or selenocysteine, being very different with respect to other known antioxidant (Kim *et al.* 1988; Chae *et al.* 1994). Later, it has been found that (i) Prxs are present in all biological kingdoms, from bacteria to animals; (ii) two cysteine residues, corresponding to Cys^{47} and Cys^{170} of yeast Prx, are highly conserved among Prx family members; (iii) Prxs are homodimers arranged in a head-to-tail orientation; and (iv) Cys^{47}–SH of Prx is specifically oxidised by H_2O_2 to cysteine sulfenic acid (Cys–SOH), which is resolved by reaction with Cys^{170}–SH of the adjacent monomer. Theseis results in the formation of a disulfide link, Cys^{47}–S-S–Cys^{170} (Chae *et al.* 1994). The conserved Cys residue corresponding to Cys^{47} of yeast Prx was later referred to as the peroxidatic Cys (C_p) to reflect its sensitivity to oxidation by peroxides, and the conserved Cys residue corresponding to Cys^{170} was designated the resolving Cys (C_R) (Wood *et al.* 2003).

On the basis of the presence or absence of the C_R residue, Prxs are classified into typical 2-Cys, atypical 2-Cys, and 1-Cys Prx subfamilies (Rhee & Woo, 2011). Animal cells express six isoforms of Prxs: isoforms from 1 to 4 belong to the typical 2-Cys Prx group; Prx5 belongs to the class of 2-Cys enzymes. Isoform 6 is the only one belonging to

*Corresponding author

1-Cys Prxs (Fisher, 2011). Prx1 is mainly localised in the cytosol, nucleus and peroxisomes, but it has been found also in serum (Immenschuh *et al.* 2003). Prx2 is present in cytosol and nucleus and it has been shown to bond the cell membrane (Cha *et al.* 2000). Prx3 is located exclusively in the mitochondria (Cao *et al.* 2007). Prx4 has been found in both cytosol and endoplasmic reticulum and contains a leader peptide that is believed to be essential for protein secretion (Okado-Matsumoto *et al.* 2000). Prx5 is localised in cytosol, mitochondria and peroxisomes (Cao *et al.* 2007). Prx6 is located in cytosol, vesicles and lysosomes (Sorokina *et al.* 2011). Prx localisation is multifarious, being dependent on the cell type but also on the environmental conditions (Rhee *et al.* 2012).

Many studies suggest that Prxs are more than just simple antioxidant enzymes. For example Prx6 also acts as phospholipase A2 (Chen *et al.* 2000). Other evidences suggest that Prx oxidation also allows them to function as molecular chaperones (Jang *et al.* 2004) and regulates the cell cycle (Phalen *et al.* 2006). A very interesting proposed role is encompassed by the floodgate hypothesis (Wood *et al.* 2003; Woo *et al.* 2010), in which active Prxs normally keep H_2O_2 low (i.e. a closed floodgate) but, under signalling conditions that causes loss of function via overoxidation in a localised region of the cell, allowing H_2O_2 to build up locally (i.e, be released by an open floodgate) for signalling purposes (Hall *et al.* 2009, Hanschmann *et al.* 2013).

The abundance of Prx is high: it can account for up to 1% of all soluble cellular proteins (Wood *et al.* 2003). Furthermore, typical 2-Cys Prxs are the largest and most widely distributed subfamily (Soito *et al.* 2011). These Prxs are moonlighting proteins: at high H_2O_2 concentrations they act as holdases, whereas when the rate of ROS formation is low they are thioredoxin-dependent peroxidases (Jang *et al.* 2004). The dual functions of typical 2-Cys Prxs, modulating ROS concentrations and preventing protein aggregation, may play pivotal roles in cellular response to pathogens and external stress (Jang *et al.* 2004). The importance of Prxs is unarguable, as transgenic/knockout mouse models overexpressing or deficient in most highly expressed Prxs has demonstrated a decrease in genome stability and accelerated aging, and an overexpression of these proteins is associated with various problems related to cancers treatment (Hamilton *et al.* 2012). Recent studies show Prxs expressed in tumortumour cells play positive roles in their progression and/or metastasis in transplanted animals. Different functions of Prxs are required for their progression/metastasis *in vivo* depending on tumortumour types (Ishii *et al.* 2012).

Metazoa is one of great eukaryote kingdoms. An increasingly well resolved Proterozoic fossil record documents the presence of most of the major taxa of eukaryotes, including the rhodophytes, stramenopiles, alveolates and green plants during the late Mesoproterozoic to early Neoproterozoic (about one billion years ago). A coincident rise in acritarch diversity, combined with molecular phylogeny evidence for rapid cladogenesis, points to a major radiation of eukaryote groups at this time, sometimes referred to as the "big bang" of eukaryotic evolution (Conway Morris *et al.* 1987). The most fundamental division within the bilateral metazoans is the protostome-deuterostome branch. Protostomes include at least annelids, arthropods, molluscs and platyhelminths. They have spiral cleavage and usually mosaic development, are schizocelic, form the mouth at (or near) the site of the blastopore and mesoderm from a mesentoblast that is usually 4d. Deuterostomes include at least echinoderms and chordates, the latter group including the back-boned animals. They have radial cleavage and usually regulative development, are enterocelic, form the mouth away from the blastopore and mesoderm from endodermal cells along the archeneteron (Valentine *et al.* 1997).

Approximately 2.5 billion years ago, photosynthetic bacteria acquired the capacity to photodissociate water, leading to the geologically rapid accumulation of molecular oxygen during the Great Oxidation Event (GOE), when anaerobic life underwent a catastrophic decline (Edgar *et al.* 2012). Organisms that survived the transition to an aerobic environment were those that respired and/or evolved oxygen. All oxygen-utiliszing organisms acquired constitutive antioxidant defensesdefences (both small molecules and enzymes), detoxifying and scavenging the ROS that are continuously produced as a by-product of aerobic metabolism (Acworth *et al.* 1997, Halliwell & Gutteridge 1999). Among them, Prx has been object of an increasing interest for its pivotal role in cell defensedefence and as conserved marker for circadian rhythms in metabolism. Edgar and colleagues (2012) studied the cycles of Prx oxidation-reduction in Eukarya, Bacteria and Archaea and proposed that all these organisms have cellular rhythms sharing a common molecular origin. In fact, it has been proposed that the ability to survive cycles of oxidative stress may have contributed a selective advantage from the beginnings of aerobic life. Twenty-four hours cycles of Prx oxidation–reduction in all domains were observed in both Archaea, Bacteria and Eukaryota supporting the hypothesis that cellular rhythms share a common molecular origin (Edgar *et al.* 2012). However, only

a few studies have been dedicated to the evolution of Prxs (Copley *et al.* 2004; Knoops *et al.* 2007; Pérez-Sánchez *et al.* 2011). Our study aims and we decided to extend the research on Prx evolution performing phylogenetic analysis on the most studied group of eucariotic organisms (Eukarya), the animals, using Bayesian approach. Fortunately, there is sufficient information about many characterised nucleotide and amino acid sequences of metazoan Prxs. Using bioinformatics tools we analysed the phylogenetic relationships among metazoan Prxs, with the aim to acquire new data on the molecular evolution of these proteins.

Material and Methods

Amino acid and mRNA coding sequences of different Prxs isoforms from protostomes were found in the NCBI database (Table 1). These were Japanese spineless cuttlefish *Sepiella maindroni* Rochebrune, 1884, cockscomb pearl mussel *Cristaria plicata* Leach, 1815, Pacific abalone *Haliotis discus discus* Reeve, 1846, Sydney rock oyster *Saccostrea glomerata* (Gould, 1850), mud crab *Scylla paramamosain* Estampador, 1949, Chinese mitten crab *Eriocheir sinensis* H. Milne-Edwards, 1853, lugworm *Arenicola marina* (Linnaeus, 1758) and southern house mosquito *Culex quinquefasciatus* Say, 1823. The following deuterostomes were selected: rhesus macaque *Macaca mulatta* (Zimmermann, 1780), Sumatran orangutan *Pongo abelii* Lesson, 1827, human *Homo sapiens* Linnaeus, 1758, European cattle *Bos taurus* Linnaeus, 1758, wild boar *Sus scrofa* Linnaeus, 1758, brown rat *Rattus norvegicus* Linnaeus, 1758, house mouse *Mus musculus* Linnaeus, 1758, chicken *Gallus gallus domesticus* (Linnaeus, 1758), rock pigeon *Columba livia* Linnaeus, 1758, African clawed frog *Xenopus laevis* Daudin, 1802, western clawed frog *Xenopus tropicalis* (Gray, 1864), striped beakfish *Oplegnathus fasciatus* (Temminck & Schlegel, 1844), zebrafish *Danio rerio* (F. Hamilton, 1822), gilthead seabream *Sparus aurata* Linnaeus, 1758,, channel catfish *Ictalurus punctatus* (Rafinesque, 1818), rainbow trout *Oncorhynchus mykiss* (Walbaum, 1792), Atlantic salmon *Salmo salar* Linnaeus, 1758, and southern bluefin tuna *Thunnus maccoyii* (Castelnau, 1872). All respective sequences were aligned using T-Coffee multiple sequence alignment software package (Notredame *et al.* 2000).

jModelTest (Posada 2008) was used to carry out statistical selection of the best-fit models of nucleotide substitution to analyse Prx molecular evolution. Analyses were performed using 88 candidate models of nucleotide substitution and two types of information criterion (Akaike Information Criterion-AIC and Corrected Akaike Information Criterion-cAIC). Models of nucleotide substitution allow for the calculation of probabilities of change between nucleotides along the branches of a phylogenetic tree. The use of a particular substitution model may change the outcome of the phylogenetic analysis (Lemmon and Moriarty, 2004). Statistical model selection has become an essential step for the estimation of phylogenies from DNA sequence alignments. ProtTest3 was used for the selection of the best-fit model of analysed protein evolution (Darriba *et al.* 2011). The program ProtTest is one of the most popular tools for selecting models of amino acid replacement, a routine step in phylogenetic analysis. One hundred twenty-two candidate models of amino acid replacement and three types of criteria (Akaike Information Criterion-AIC, Corrected Akaike Information Criterion-cAIC and Bayesian Information Criterion-BIC) were used in these statistical analyses. The Prx cDNA and amino acid sequences phylogenetic trees were built using the Bayesian inference (BI) method implemented in Mr. Bayes 3.2 (Ronquist *et al.* 2012). Four independent runs, each one with four simultaneous Markov Chain Monte Carlo (MCMC) chains, were performed for 1,000,000 generations sampled every 1000 generations. FigTree software (http://tree.bio.ed.ac.uk/software/figtree) was used to display the annotated phylogenetic trees. In order to detect the presence of positive or negative selection on Prx molecular evolution we used statistical methods implemented in HyPhy package (Kosakovsky Pond *et al.* 2005); SLAC, FEL and REL methods are able to detect the presence of possible positive selection and Mixed Effects Model of Evolution (MEME) program is able to detect even sites under episodic diversifying selection (Murrell *et al.* 2012).

Results and discussion

Metazoan Prxs mRNA sequences were aligned using T-Coffee software in combined libraries of local and multiple alignments, which are known to induce high accuracy and performance in sequence alignments. The obtained alignment was 1222 residues (nucleotides and gaps) long and the mean score value was 59, indicating that the alignment was reliable (Notredame *et al.* 2000). jModelTest 0.1.1 software determined that GTR+G model was the best-fit model of Prxs cDNA sequence evolution with a gamma shape value (four rate categories) of 0.72 using AIC and cAIC statistical criteria (-lnL= 16365.654). Phylogenetic relationships between all these different Prxs were determined using the most powerful

Table 1. Prx sequences used for Bayesian phylogeny, their NCBI accession numbers and the respective references

Species Isoform	References	Nucleotide accession number	Protein accession number
Arenicola marina Prx6	Loumaye *et al.*, 2008	DQ059567	AAY96294
Bos taurus Prx6	Salmeri *et al.*, 2012	NM_174643	NP_777068
Columba livia Prx6	Gao, 2012	JQ364950	AFD04441
Cristaria plicata Prx6	Pei *et al.*, 2010	HQ199304	ADN06076
Culex quinquefasciatus Prx6	Atkinson *et al.*, 2007	XM_001861490	XP_001861525
Danio rerio Prx1	Cox *et al.*, 2014	NM_001013471	NP_001013489
Danio rerio Prx2	Liu *et al.*, 2013	NM_001002468	NP_001002468
Danio rerio Prx3	Lu *et al.*, 2014	NM_001013460	NP_001013478
Danio rerio Prx4	Mukaigasa *et al.*, 2012	NM_001089425	NP_001082894
Danio rerio Prx5	Lu *et al.*, 2014	NM_001024406	NP_001019577
Danio rerio Prx6	Mukaigasa *et al.*, 2012	NM_200805	NP_957099
Eriocheir sinensis Prx6	Mu *et al.*, 2009	EU626070	ACF35639
Gallus gallus Prx6	Caldwell *et al.*, 2005	NM_001039329	NP_001034418
Haliotis discus discus Prx6	Nikapitiya *et al.*, 2006	EF103356	ABO26614
Homo sapiens Prx6	Strausberg *et al.*, 2002	BC053550	AAH53550
Ictalurus punctatus Prx6	Yeh and Klesius, 2007	DQ779284	ABG77029
*Macaca mulatta*Prx6	Predicted	NM_001266085	NP_001253014
Mus musculus Prx6	Pacifici *et al.*, 2014	NM_007453	NP_031479
Oncorhynchus mykiss Prx6	Predicted	NM_001165132	NP_001158604
Oplegnathus fasciatus Prx6	De Zoysa *et al.*, 2012	GQ903768	ADJ21808
Pongo abelii Prx6	Predicted	NM_001132889	NP_001126361
Rattus norvegicus Prx6	Paula *et al.*, 2013	NM_053576	NP_446028
Saccostrea glomerata Prx6	Green *et al.*, 2009	FJ626708	ACQ73550
Salmo salar Prx1	Loo *et al.*, 2012	NM_001140823	NP_001134295
Salmo salar Prx5	Andreassen and Hoyheim, 2013	BT149950	AGH92554
Scylla paramamosain Prx6	Fu *et al.*, 2008	FJ429110	ACJ53746
Sepiella maindroni Prx6	Song *et al.*, 2010	HQ662844	AEI52300
Sparus aurata Prx1	Perez-Sanchez *et al.*, 2011	GQ252679	ADI78064
Sparus aurata Prx2	Perez-Sanchez *et al.*, 2011	GQ252680	ADI78065
Sparus aurata Prx3	Perez-Sanchez *et al.*, 2011	GQ252681	ADI78066
Sparus aurata Prx4	Perez-Sanchez *et al.*, 2011	GQ252682	ADI78067
Sparus aurata Prx5	Perez-Sanchez *et al.*, 2011	GQ252683	ADI78068
Sparus aurata Prx6	Perez-Sanchez *et al.*, 2011	GQ252684	ADI78069
Sus scrofa Prx6	Liu *et al.*, 2011	NM_214408	NP_999573
Thunnus maccoyii Prx2	Sutton *et al.*, 2010	EU093980	ABW88997
Xenopus (*Silurana*) *tropicalis* Prx6	Klein *et al.*, 2002	NM_001011325	NP_001011325
Xenopus laevis Prx6	Shafer *et al.*, 2011	NM_001089200	NP_001082669

statistical method of BI. We compute a majority rule tree for all the trees sampled during the MCMC. We decided to use BI, because this method is much faster than the Maximum Likelihood (ML) for big datasets (Douady *et al.* 2003). The best phylogeny generated by the BI method is depicted in 1.

Amino acid sequences of Prxs were aligned and a 322 residue-long alignment was obtained. The obtained alignment was better than that previously achieved for nucleotide sequences, because its score (86) was significantly better. ProtTest3 was used for the evolution best-fit model determination. WAG+G resulted the best model, with a gamma shape value (four rate categories) of 1.0 using all statistical criteria: AIC, cAIC and BIC (-lnL= -7241.91). In Fig. 2 is shown the best phylogeny generated by the application of BI method into the Prxs amino acid sequences.

In both cladrograms, three main clusters were present: one including all typical Prx 2-Cys (isoforms 1, 2, 3 and 4), Prx5 isoforms (atypical 2-Cys Prxs) and the third grouping Prx6 isoforms (1-Cys). This distribution confirmed the previously obtained results using non-Bayesians methods (Pérez-Sánchez *et al.* 2011). In particular, the phylogenetic relationships among the typical 2-Cys isoforms were com-

Table S1. Positively selected codons identified by using computational techniques (SLAC-Single Likelihood Ancestor Counting, FEL-Fixed Effects Likelihood, REL-Random Effects Likelihood and MEME-Mixed Effects Model of Evolution) and their calculated statistics: *p-value* – the number of false positive tests; *q-value* – number of false positive significant tests; *dN* – number o nonsynonymous substitutions for site; *dS* or α – number of synonymous substitutions for site; β^- lineage-specific and β^+ unrestricted – non-synonymous substitutions for site; *posterior probabilities* and *empirical Bayes factors* (ratio of posterior and prior odds of having ω =dN/dS > 1 at a given site) – two measures for determining whether a site is under positive selection

Codon	SLAC			FEL					REL					MEME							
	dN-dS	Normalized dN-dS	p-value	dS	dN	dN/dS	Normalized dN-dS	p-value	E[dS]	E[dN]	Normalized E[dN-dS]	Posterior Probability	Bayes Factor	α	β-	Pr[β=β-]	β+	Pr[β=β+]	p-value	q-value	
16	9.15713	1.13119	0.001973	0.1482	2.9797	20.113	0.351151	0.00219	0.2445	2.0552	1.81062	1	13193000	0.1248	0.1248	6.00E-09	3.0294	1	0.003	0.03823	
106	8.7161	1.0767	0.003961	0.2436	2.5821	10.6	0.290003	0.00422	0.1889	1.7522	1.56329	1	30126900	0.2133	0.2133	6.00E-09	2.4262	1	0.006	0.05457	
124	6.11164	0.754975	0.009654	5.00E-09	2.1403	4.00E+08	0.265428	2.35E-03	0.093	1.47E+00	1.37872	1	6560730	0	0	6.00E-09	2.0368	1	0.0037	0.04247	
129	6.93142	0.856242	0.006305	0.1019	2.4136	23.687	0.286676	0.00119	0.101	1.6027	1.50176	1	220018000	0.083	0.0637	6.00E-09	2.4091	1	0.0016	0.03891	
139	10.5823	1.30723	0.007901	5.00E-09	3.8971	8.00E+08	0.483288	1.90E-04	1.1835	2.79E+00	2.60861	1.00E+00	64824.5	0	0	1.66E-01	4.8352	0.83372	0.0003	0.013	
149	6.14108	0.758611	0.007809	0	1.6935	Infinite	0.21001	0.00135	0.0813	1.4364	1.35513	1	1422990	0	0	1.02E-01	1.7028	0.89837	0.0028	0.03903	
159	8.29008	1.02408	0.003479	5.00E-09	2.7132	5.00E+08	0.336468	6.80E-05	0.0748	1.81E+00	1.73989	1	2.041E+10	0	0	6.00E-09	2.6154	1	0.0002	0.01258	
162	9.0526	1.11827	0.000602	0	2.6756	Infinite	0.331806	4.10E-05	0.0705	1.7637	1.69322	1	1.10E+11	0	0	2.74E-01	3.9012	0.72555	3.00E-05	0.00808	
185	7.34477	0.907304	0.000802	0	1.9101	Infinite	0.236881	6.90E-05	0.0665	1.45E+00	1.37862	1	934304000	0	0	6.00E-09	1.8403	1	0.0001	0.00941	
210	12.7151	1.57071	0.000899	5.00E-09	1.8043	4.00E+08	0.22375	3.52E-03	0.0976	1.52E+00	1.42271	1.00E+00	27229.4	0	0	5.84E-01	5.1242	0.41649	0.0037	0.04105	
220	9.47341	1.17026	0.00137	0.099	3.2297	32.625	0.388245	1.90E-04	0.1145	2.29E+00	2.17128	1	1.169E+09	0.1115	0.1115	6.00E-09	3.1483	1	0.0003	0.01301	
227	9.79703	1.21023	0.003694	0	3.0694	Infinite	0.38065	3.70E-05	0.0825	2.12E+00	2.03312	1	777602000	0	0	6.00E-09	3.1601	1	8.00E-05	0.00781	
236	9.28409	1.14687	0.002922	0	2.3314	Infinite	0.289126	4.79E-03	0.2155	1.59E+00	1.37357	1	462273								
260	6.0459	0.746854	0.008575	0	1.5505	Infinite	0.192282	1.10E-03	0.0703	1.43E+00	1.35918	1	1.74E+07	0	0	6.00E-09	1.5164	1	0.0025	0.03878	
261	7.57435	0.935665	0.004778	0.2091	2.2616	10.814	0.254535	2.28E-03	0.1444	1.53E+00	1.38826	1	6.78E+07	0.1948	0.1948	6.00E-09	2.2162	1	0.0033	0.04037	
263	6.10328	0.753941	0.008069	5.00E-09	1.6126	3.00E+08	0.199977	1.32E-03	0.0723	1.43E+00	1.36025	1	3.26E+07	0	0	1.15E-02	1.5918	0.98855	0.0017	0.03677	
269	14.1287	1.74533	0.002773	5.00E-09	4.1148	8.00E+08	0.510288	4.30E-04	0.3335	2.81E+00	2.47379	1.00E+00	2.25E+03	0	0	6.00E-09	5.26E+00	1	0.0016	0.04499	
308	10.7535	1.32839	0.009462	5.00E-09	1.9976	4.00E+08	0.247727	7.05E-03	0.1185	1.58E+00	1.45676	1	2.45E+05								

Table S2. Negatively selected codons identified by using computational techniques (SLAC-Single Likelihood Ancestor Counting, FEL-Fixed Effects Likelihood and REL-Random Effects Likelihood) and their calculated statistics: *p-value* – the number of false positive tests; *dN* – number of nonsynonymous substitutions for site; *dS* – number of synonymous substitutions for site; *posterior probabilities* and *empirical Bayes factors* (ratio of posterior and prior odds of having ω =dN/dS < 1 at a given site) – two measures for determining whether a site is under negative selection

Codon	SLAC			FEL					REL				
	dN-dS	Normalized dN-dS	p-value	dS	dN	dN/dS	Normalized dN-dS	p-value	E[dS]	E[dN]	Normalized E[dN-dS]	Posterior Probability	Bayes Factor
20	-11.358	-1.40252	0.00026	6.6837	0.46277	0.07	-0.771473	4.21E-06	5.465	0.28	-5.18972	0.998794	11019.9
163	-4.8376	-0.597586	1.1976	1.1976	0	0	-0.148523	0.00095	1.098	0.21	-0.892806	0.999825	75796.9

Table S3. Positively selected codons identified using the data subset including Prx6 coding sequences from *X. tropicalis*, *C. plicata*, *A. marina*, *S. paramamosain*, *E. sinensis*, *H. discus discus*, *S. maindroni* and *S. glomerata*. Computational techniques (FEL-Fixed Effects Likelihood, REL-Random Effects Likelihood and MEME-Mixed Effects Model of Evolution) were applied and calculated statistics are reported

Codon	FEL					REL					MEME						
	dS	dN	dN/dS	Normalized dN-dS	p-value	E[dS]	E[dN]	Normalized E[dN-dS]	Posterior Probability	Bayes Factor	α	β-	Pr[β=β-]	β+	Pr[β=β+]	p-value	q-value
96	5.00E-09	1.84887	3.70E+08	0.955191	2.76E-02	0.019351	5.31E-01	0.511949	9.90E-01	103.083	5.00E-09	4.89E-10	3.55E-01	2.68E+00	6.45E-01	3.52E-02	4.93E-01
108	5.00E-09	1.7315	3.46E+08	0.89455	2.45E-02	0.016347	5.15E-01	0.498974	9.93E-01	140.889	0	0	6.00E-09	1.90262	1	0.0209	0.5198
112	5.00E-09	1.72799	3.46E+08	0.892736	2.80E-02	0.016732	5.15E-01	0.4983	9.93E-01	134.566	0	0	6.00E-09	1.69764	1	0.0315	0.5422
126	5.00E-09	1.20115	2.40E+08	0.620555	4.45E-02	0.013299	4.58E-01	0.444394	9.96E-01	231.242							
127	5.00E-09	1.2942	2.59E+08	0.668627	4.87E-02	0.017199	4.58E-01	0.440754	9.92E-01	127.76	5.00E-09	4.99E-09	6.00E-09	1.37E+00	1	4.65E-02	4.96E-01
143	5.00E-09	1.71512	3.43E+08	0.886091	2.82E-02	0.016986	5.11E-01	0.493914	9.92E-01	130.579	5.00E-09	4.99E-09	6.00E-09	1.67E+00	1	3.17E-02	5.08E-01
147	5.00E-09	1.33782	2.68E+08	0.691165	3.77E-02	0.012953	4.65E-01	0.451779	9.96E-01	240.347							
199	5.00E-09	2.40312	4.81E+08	1.24154	1.28E-02	0.01968	6.72E-01	0.652494	9.90E-01	99.9964	5.00E-09	4.99E-09	6.00E-09	2.37E+00	1	1.44E-02	8.06E-01
210	5.00E-09	1.10103	2.20E+08	0.56883	4.94E-02	0.012936	4.49E-01	0.436025	9.96E-01	253.813							
220	5.00E-09	1.16401	2.33E+08	0.601366	4.81E-02	0.013549	4.54E-01	0.440821	9.95E-01	221.996							

Table S4. Negatively selected codons identified by using the data subset including Prx6 coding sequences from *X. tropicalis*, *C. plicata*, *A. marina*, *S. paramamosain*, *E. sinensis*, *H. discus discus*, *S. maindroni* and *S. glomerata*. Computational techniques (SLAC-Single Likelihood Ancestor Counting, FEL-Fixed Effects Likelihood and REL-Random Effects Likelihood) were applied and calculated statistics are reported

Codon	SLAC			FEL					REL				
	dN-dS	Normalized dN-dS	p-value	dS	dN	dN/dS	Normalized dN-dS	p-value	E[dS]	E[dN]	Normalized E[dN-dS]	Posterior Probability	Bayes Factor
10	-5.33772	-2.75765	0.002558	6.70022	5.00E-09	0	-3.46156	0.000451	2.21125	0.026502	-2.18475	0.999993	135611
11	-2.18987	-1.13136	0.037037	3.95938	5.00E-09	0	-2.04555	0.000758	1.67697	0.027609	-1.64936	0.999996	252936
12	-4.51862	-2.33447	0.004216	412.604	5.00E-09	0	-213.165	4.19E-05	3.39193	0.019916	-3.37202	0.999999	1.22E+06
13	-4.4566	-2.30243	0.004394	5.39369	5.00E-09	0	-2.78656	0.001043	2.01664	0.027414	-1.98922	0.999917	11972.5
15	-3.64979	-1.8856	0.004115	14.7723	5.00E-09	0	-7.63185	3.13E-05	3.36955	0.036982	-3.33257	1	4.51E+09
20	-2.91983	-1.50848	0.01249	1241.46	5.00E-09	0	-641.38	0.000163	3.72818	0.051906	-3.67627	1	1.31E+07
25	-3.37632	-1.74432	0.020774	5.06041	5.00E-09	0	-2.61438	1.27E-03	1.94891	0.021085	-1.92782	0.99981	5.23E+03
35	-2.95361	-1.52594	0.030038	16.2418	2.76E-01	0.017	-8.24837	0.001214	3.49025	0.333878	-3.15637	1	7.58E+06
36	-3.59685	-1.85826	0.018305	2.96374	5.00E-09	0	-1.53117	5.83E-03	1.74999	0.025746	-1.72424	0.999245	1.32E+03
38	-4.35539	-2.25015	0.004708	1522.03	5.00E-09	0	-786.334	0.000169	2.95471	0.021624	-2.93308	0.999997	2.85E+05
39	-2.91983	-1.50848	0.012346	3.8513	5.00E-09	0	-1.98971	6.03E-04	1.7192	0.027609	-1.69159	1	5.04E+06
41	-3.61222	-1.8662	0.018149	3.68796	5.00E-09	0	-1.90533	0.00464	1.86015	0.024408	-1.83574	0.999389	1.63E+03
43	-2.91983	-1.50848	0.012346	4.42266	5.00E-09	0	-2.2849	3.34E-04	1.89099	0.024909	-1.86608	1	8.58E+06
44	-2.18987	-1.13136	0.037037	4.50411	5.00E-09	0	-2.32698	0.00092	1.82604	0.027592	-1.79844	0.999994	1.63E+05
46	-3.61563	-1.86796	0.024259	4.05905	5.00E-09	0	-2.09704	6.35E-03	1.88395	0.035383	-1.84857	0.999457	1.83E+03
48	-3.64979	-1.8856	0.004115	13.1928	5.00E-09	0	-6.81587	4.66E-05	2.76819	0.02491	-2.74328	1	9.84E+08
49	-2.71714	-1.40377	0.037508	3.36881	5.00E-09	0	-1.74044	4.80E-03	1.70428	0.028729	-1.67556	0.99941	1.69E+03
50	-2.76069	-1.42627	0.016891	13.4678	5.00E-09	0	-6.95794	7.18E-04	2.29902	0.054331	-2.24469	1	1.09E+08
59	-3.82529	-1.97627	0.023387	31.8233	2.00E-01	0.006	-16.3374	1.07E-03	3.51887	0.204275	-3.31459	0.999946	1.85E+04
63	-2.24434	-1.1595	0.034967	1.82006	5.00E-09	0	-0.940304	5.98E-03	1.16275	0.032041	-1.13071	0.999855	6.85E+03
69	-2.91983	-1.50848	0.012346	14.0584	5.00E-09	0	-7.26305	2.24E-04	2.58859	0.036989	-2.5516	1	1.88E+07
71	-2.18987	-1.13136	0.039348	2.39547	5.00E-09	0	-1.23758	5.15E-03	1.94708	0.042866	-1.90421	0.999975	4.00E+04
72	-2.89208	-1.49415	0.033926	2.41539	5.00E-09	0	-1.24787	9.69E-03	1.63412	0.035435	-1.59868	0.999146	1.16E+03
75	-2.91983	-1.50848	0.012346	5.17873	5.00E-09	0	-2.67551	6.36E-04	1.83222	0.052531	-1.77969	1	1.04E+07
78	-3.63352	-1.8772	0.017937	260.578	5.00E-09	0	-134.623	2.32E-03	2.30211	0.020371	-2.28174	0.999523	2.08E+03
178	-3.49587	-1.80609	0.031831	5.9078	2.90E-01	0.049	-2.90222	6.59E-03	2.13655	0.343866	-1.79268	0.999988	8.18E+04

patible with the hypothesis exposed by Copley and *colleagues* (2004) which suggested an evolutionary model based on the appearance of 3 groups of Prxs (1-2, 3 and 4) through the loss of N-terminal extension in Prx3s and the acquisition of C-terminal extension in Prx4s.

The deuterostomes Prxs were equally positioned and the respective branches were supported by the highest posterior probability values. The only exception was the Prx6 of *X. tropicalis*, which in both trees emerged far from the sequence of *X. laevis*. These results were peculiar. They might suggest the appearance of Prx6 gene at different times, through independent duplication events of the ancestor, even after the speciation event that led to the differentiation of *Chordata* phylum. Based on the currently available data, it was not possible to confirm that the hypothetical multiple appearance of Prx6 could be a distinguishing feature of Amphibians or might indeed have happened also in other classes. Whatever the case, this duplication event could not be recent but should have occured at least 416-360 million years ago, and might be typical of Prx6. In fact, among the various isoforms, only Prx6 are known to be represented by multiple sub-isoforms, such as in *Drosophila melanogaster* (DPX-6005 and DPX-2540, Radyuk *et al.* 2001) and *Ciona intestinalis* (unpublished personal data).

The distribution of invertebrate sequences was less consistent. All protostome Prx6 sequences were grouped except for the protein of *C. quinquefasciatus* that emerged as a single branch (Fig. 1). This position was the same in the phylogenetic analysis performed on amino acid sequences, but in this case it resulted to be the sister group of invertebrate Prxs (Fig. 2). Instead, the sequence of *A. marina* seems to be more related to vertebrate Prxs. Probably, the discordance emerged from the comparison of the two phylogenetic topologies may be linked to a possible differences in substitution rates, which are usually caused by positive and/or negative selection.

Evolutionary biologists typically have invoked two types of selective forces shaping the evolution of species. One is the purifying selection, which favours the conservation of existing phenotypes. The other is the positive selection (also known as Darwinian selection), which promotes the emergence of new phenotypes. Positive selection can leave a set of telltale signatures in the genes, such as the rapid divergence of functional sites between species (diversifying selection) and the depression of polymorphism within species (Kreitman *et al.* 2000, Yang & Bielawski 2000, Bamshad & Wooding 2003). The imprint of natural selection (positive selection) on protein coding genes is often difficult to identify because selection is frequently transient or episodic, i.e. it affects only a subset of lineages. To verify the selection type in Prx evolution we used existing computational techniques (SLAC, FEL and REL) implemented in the HyPhy package. These techniques are designed to identify sites subject to pervasive selection (a large proportion of positively selected sites) but may fail to recognise sites where selection is episodic. For this reason we used MEME method that is able to identify instances of both episodic and pervasive positive selection at individual site level. The obtained results (Table S1) indicated that nearly 5% of the 308 codons were positively selected and Table S2less than 1% of them were negatively selected (Table S2). These results suggested that Prx genes were more susceptible to positive selection than to negative one.

It is known that if all existing gene sequences had not been screened for recombination, selection analyses of alignments with recombinants in them using a single tree could generate misleading results (Kosakovsky Pond *et al.* 2005). Thus, we used GARD program (Kosakovsky Pond *et al.* 2006) to identify possible breakpoints in the Prx gene sequences. One breakpoint has been found (-lnL = 31879), but Kishino Hasegawa (KH) tests indicated it as statistically non-significant breakpoint in each level of significance. However, the calculated mean substitution value was 2.89 substitutions per site, revealing that there was a huge amount of divergence among all the analysed sequences. In this case all the possible positive or negative selection tests could be non-reliable. In order to verify the results we carried out the selection tests for a subset of sequences that were not so divergent. We included in this data subset only the Prx sequences that were responsible for the emerged discordance between the two phylogenetic topologies. The Prx6 coding sequences of *X. tropicalis*, *C. plicata*, *A. marina, S. paramamosain, E. sinensis, H. discus discus, S. maindroni* and *S. glomerata* were aligned using T-Coffee software. We obtained a good score (92). The calculated mean substitutions value was 0.77 substitutions per site, which meant that the divergence among the subset sequences was low. Thus, we carried out the previously used selection tests (Tables S3 & S4). Additionally, for this data subset GARD application found one breakpoint (-lnL = 8983.58), but Kishino Hasegawa (KH) tests assigned it as statistically non-significant breakpoint in each level of significance. No positive selection results were obtained by using the SLAC method (Table S3).

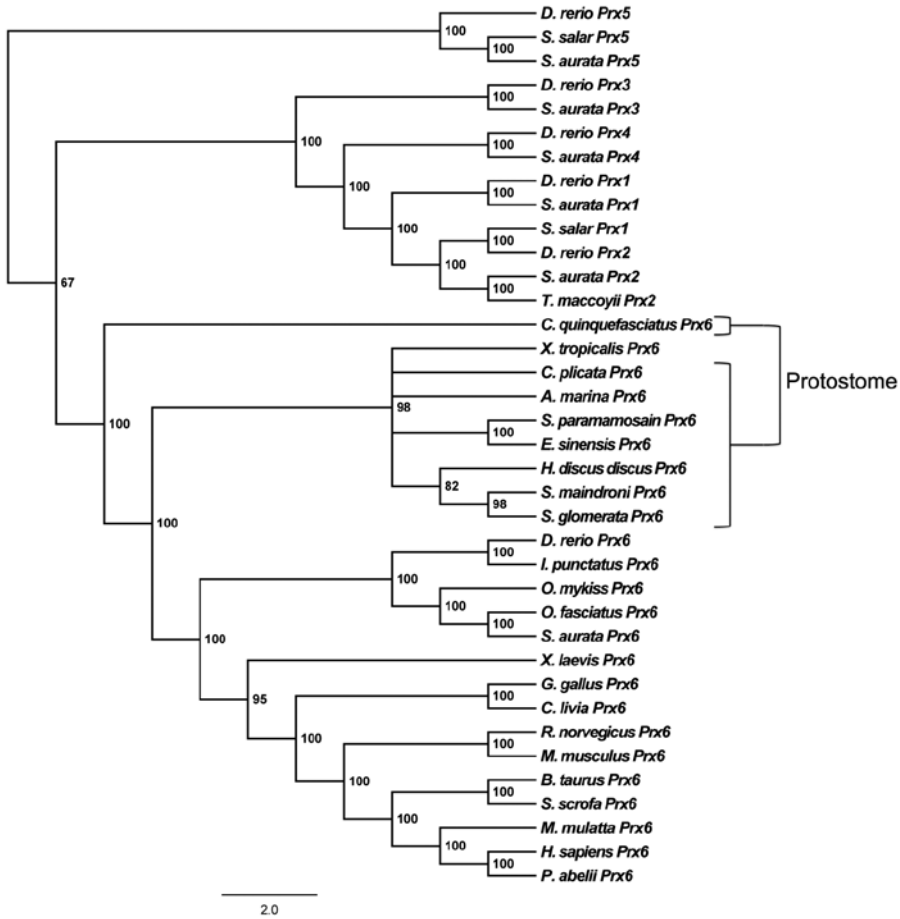

Fig. 1. Phylogenetic relationships among metazoan Prxs using coding sequences and BI (Bayesian Inference) method (arithmetic mean = -29057.97; harmonic mean = -29083.98). Posterior probability values higher than 50% are indicated on each node. The scale for branch length (2.0 substitution/site) is shown below the tree

We compared the results presented in Table S4 with the crystal structure of Prx1 from *S. mansoni* (SmPrx1, PDB code 3ztlD, solved by X-ray, resolution 3.00Å) characterised by Saccoccia and colleagues (2012). In the active site of SmPrx1 low-molecular-weight dimer, residues from 47 to 50, including the C_P (Cys^{48}), form the first turn of the α2 helix. The compact hydrophobic pocket is formed by Tyr^{40}, Pro^{41}, Ala^{42}, Thr^{45} and Pro^{49} that, together with Arg^{124}, surround the C_P (Saccoccia *et al.* 2012). Pro^{49} is highly conserved in typical 2-Cys Prxs (Hofmann *et al.* 2002) and reduces the propensity of the first turn of the α2 helix to retain its secondary structure. Our results indicated that only codons 41, 48 and 49 were negatively selected (Table S4), confirming that some (but not all) of the important for peroxidase activity amino acids were conserved during the evolution of Prxs. While these data confirmed the soundness of our results, on the other hand it brought out some questions on the absence of conservation for the other amino acids that were important for catalytic activity. In this subset of sequences from invertebrates, the number of negatively selected codons (about 12% of the 224 codons) was higher than the positively selected ones (approximately 5%, Tables S3 & S4).

All these results confirmed that the molecular evolution of metazoan Prx was peculiar and suggested that the positive selection may have operated into the evolution of these proteins and a purifying selection has been present during this process. Probably, the natural selection was responsible for the deviation from the evolution pathway of Prx genes characteristic of the protostome isoforms.

Further analyses are needed to verify if the positively selected codons, identified by bioinformatic approach, really codify essential amino acids. A possible verification could be done by site-specific mutagenesis of those nucleotides that could structurally and functionally affect the enzymatic activity of these proteins.

Purifying selection is important for the evolution of a gene family, because it can help the belonging genes to maintain their optimal function.

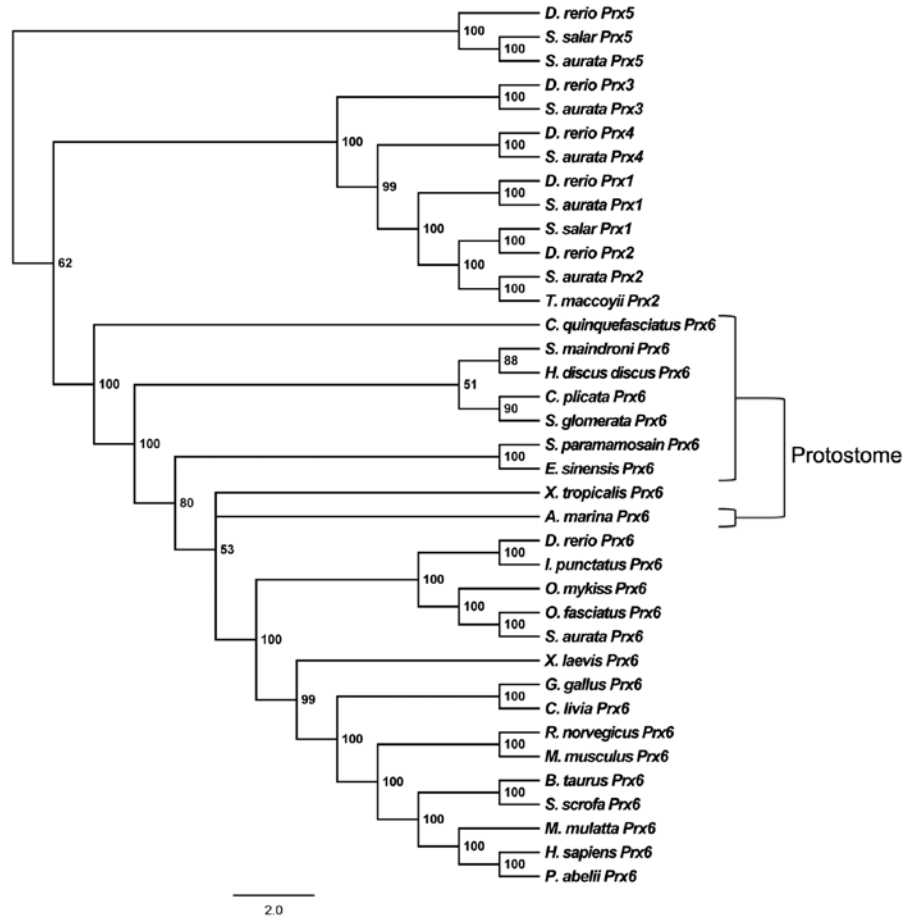

Fig. 2. Phylogenetic relationships among metazoan Prxs using amino acid sequences and BI (Bayesian Inference) method (arithmetic mean = -12412.89; harmonic mean = -12437.68). Posterior probability values higher than 50% are indicated on each node. The scale for branch length (2.0 substitution/site) is shown below the tree

However, positive selection is an important source of evolutionary innovation and is a major force underlying the adaptation of species to a new environment (Kosiol *et al.* 2008). Many proteins have been found under positive selection, such as those involved in immunity (MHC, immunoglobulin VH, class 1 chitinas), proteins or pheromones involved in reproduction (abalone sperm lysin, sea urchin bindin, proteins in mammals) and proteins that acquire new functions after gene duplication (Yang & Bielawski 2000). We suggest to include Prxs among those proteins, as their diversification allowed the cells of all organisms to acquire a powerful defence system against the risk of oxidative stress.

References

Acworth I. N., Mc Cabe D. R. and Maher T. J. 1997. The analysis of fre radicals, their reaction products, and antioxidants. In: Baskin S.I., Salem H editors, Oxidants, Antioxidants, and Free Radicals. Washington DC: Taylor and Francis, 23-77.

Andreassen R. and Hoyheim B. 2013. Full length cDNAs in salmo salar pre-smolt and tissue specific libraries. Unpublished.

Atkinson P. W., Hemingway J., Christensen B. M., Higgs S., Kodira C., Hannick L., Megy K. O'Leary S., Pearson M., Haas B. J., Mauceli E., Wortman J. R., Lee N. H., Guigo R., Stanke M., Alvarado L., Amedeo P., Antoine C. H., Arensburger P., Bidwell S. L., Crawford M., Camaro F., Devon K., Engels R., Hammond M., Howarth C., Koehrsen M., Lawson D., Montgomery P., Nene V., Nusbaum C., Puiu D., Romero-Severson J., Severson D. W., Shumway M., Sisk P., Stolte C., Zeng Q., Eisenstadt E., Fraser-Liggett C., Strausberg R., Galagan J., Birren B. and Collins F. H. 2007. The Broad Institute Genome Sequencing Platform Annotation of *Culex pipiens quinquefasciatus.* Unpublished.

Bamshad M. & Wooding S.P. 2003. Signatures of natural selection in the human genome. – *Nature Reviews Genetics,* **4**: 99-111.

Caldwell R. B., Kierzek A. M., Arakawa H., Bezzubov Y., Zaim J., Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M., Plachy J., Carninci P., Hayashizaki Y. and Buerstedde J. M. 2005. Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis. – *Genome Biology*, **6**: R6.

Cao Z., Lindsay J. G. and Isaacs N. W. 2007. Mitochondrial peroxiredoxins. – *Subcellular Biochemistry*, **44**: 295-315.

Cha M. K., Yun C. H. and Kim I. H. 2000. Interaction of human thiol-specific antioxidant protein 1 with erythrocyte plasma membrane. – *Biochemistry*, **39**: 6944-6950.

Chae H. Z., Chung S. J. and Rhee S. G. 1994. Thioredoxin-dependent peroxide reductase from yeast. – *Journal of Biological Chemistry*, **269**: 27670-27678.

Chen J. W., Dodia C., Feinstein S. I., Jain M. K. and Fisher A. B. 2000. 1-Cys peroxiredoxin, a bifunctional enzyme with glutathione peroxidase and phospholipase A2 activities. – *Journal of Biological Chemistry*, **275**: 28421-28427.

Conway Morris S., Peel J. S., Higgins A. K., Soper N. J. and Davis N. C. 1987. A burgess shale-like fauna from the lower cambrian of north greenland. – *Nature,* **345**: 802-805.

Copley S. D., Novak W. R. P. and Babbitt P. C. 2004. Divergence of Function in the Thioredoxin Fold Suprafamily: Evidence for Evolution of Peroxiredoxins from a Thioredoxin-like Ancestor. – *Biochemistry*, **43**: 13981-13995.

Cox A. G., Saunders D. C., Kelsey P. B. Jr, Conway A. A., Tesmenitsky Y., Croce K. J., North T. E. and Goessling W. 2014. S-nitrosothiol signaling regulates liver development and improves outcome following toxic liver injury. – *Cell Reports*, **6**: 56-69.

Darriba D., Taboada G.L., Doallo R. and Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. – *Bioinformatics*, **27**: 1164-1165.

De Zoysa M., Ryu J. H., Chung H. C., Kim C. H., Nikapitiya C., Oh C., Kim H., Saranya Revathy K., Whang I. and Lee J. 2012. Molecular characterization, immune responses and DNA protection activity of rock bream (*Oplegnathus fasciatus*), peroxiredoxin 6 (Prx6). – *Fish Shellfish Immunology*, **33**: 28-35.

Douady C. J., Delsuc F., Boucher Y., Doolittle W. F. and Douzery E.J.P. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. – *Molecular Biology Evolution*, **20**: 248-254.

Edgar R. S., Green E. W., Zhao Y., van Ooijen G., Olmedo M., Qin X., Xu Y., Pan M., Valekunja U. K., Feeney K. A., Maywood E. S., Hastings M. H., Baliga N. S., Merrow M., Millar A.J., Johnson C. H., Kyriacou C. P., O'Neill J. S. and Reddy A. B. 2012. Peroxiredoxins are conserved markers of circadian rhythms. – *Nature*, **485**: 459-464.

Fisher A. B. 2011. Peroxiredoxin 6: a bifunctional enzyme with glutathione peroxidase and phospholipase A2 activities. – *Antioxidants and Redox Signaling*, **15**: 831-844.

Fu M. J., Zou Z. H., Zhang Z. P. and Wang Y. L. 2008. Molecular characterization and expression profile of glutathione peroxidase in the developing ovary and testis of *Scylla paramamosain*. Unpublished.

Gao P. F. 2012. Molecular Cloning and cDNA Sequence Analysis of Peroxiredoxin of Pigeon. Unpublished.

Green T. J., Dixon T. J., Devic E., Adlard R. D. and Barnes A. C. 2009. Differential expression of genes encoding anti-oxidant enzymes in Sydney rock oysters, *Saccostrea glomerata* (Gould) selected for disease resistance. – *Shellfish Immuno*logy, **26**: 799-810.

Hall A., Karplus P. A. and Poole L. B. 2009. Typical 2-Cys peroxiredoxins – structures, mechanisms and functions. – *FEBS Journal,* **276**: 2469–2477.

Halliwell B. and Gutteridge J. M. C. 1999. Free Radicals in Biology and Medicine, 3rd ed. Oxford: Oxford Science Publications.

Hamilton R.T., Walsh M.E. and Van Remmen H. 2012. Mouse Models of Oxidative Stress Indicate a Role for Modulating Healthy Aging. – *Journal of Clinical and Experimental Pathology*, **S4**: 1–14.

Hanschmann E-A, Godoy J.R., Berndt C., Hudemann C. and Lillig C.H. 2013. Thioredoxins, glutaredoxins, and peroxiredoxins--molecular mechanisms and health significance: from cofactors to antioxidants to redox signaling. – *Antioxidants and Redox Signaling,* **19**:1539-605.

Hofmann B., Hecht H.J. and Flohé L. 2002. Peroxiredoxins. – *Biological Chemistry*, **383**: 347–364.

Immenschuh S., Baumgart-Vogt E., Tan M., Iwahara S., Ramadori G. and Fahimi H.D. 2003. Differential cellular and subcellular localization of heme-binding protein 23/peroxiredoxin I and heme oxygenase-1 in rat liver. – *Journal of Histochemistry and Cytochemistry*. **51**: 1621–1631.

Ishii T., Warabi E. and Yanagawa T. 2012. Novel roles of peroxiredoxins in inflammation, cancer and innate immunity. – *Journal of Clinical Biochemistry and Nutrition*, **50**: 91-105.

Jang H.H., Lee K.O., Chi Y.H., Jung B.G., Park S.K., Park J.H., Lee J.R., Lee S.S., Moon J.C., Yun J.W., Choi Y.O., Kim W.Y., Kang J.S., Cheong G-W, Yun D-J, Rhee S.G., Cho M.J. and Lee S.Y. 2004. Two enzymes in one; two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. – *Cell*, **117**: 625–635.

Kim K.W., Kim I.H., Lee K.Y., Rhee S.G. and Stadtman E.R. 1988. The isolation and purification of a specific protector protein which inhibits enzyme inactivation by a thiol/Fe(III)/O_2 mixed-function oxidation system. – *Journal of Biological Chemistry*, **263**: 4704–4711.

Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W. and Richardson P. 2002. Genetic and genomic tools for *Xenopus* research: The NIH Xenopus initiative. – *Developmental Dynamics,* **225**: 384-391.

Knoops B., Loumaye E. and Van Der Eecken V. 2007. Evolution of the peroxiredoxins. – *Subcellular Biochemistry*, **44**: 27–40.

Kosakovsky Pond S.L. and Frost S.D.W. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. – *Molecular Biology and Evolution*, **22**: 1208–1222.

Kosakovsky Pond S.L., Frost S.D.W. and Muse S.V. 2005. HyPhy: hypothesis testing using phylogenies. – *Bioinformatics*, **21**: 676–679.

Kosakovsky Pond S.L., Posada D., Gravenor M.B., Woelk C.H. and Frost S.D.W. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. – *Molecular Biology and Evolution*, **23**: 1891–1901.

Kosiol C., Vinar T., da Fonseca R.R., Hubisz M.J., Bustamante C.D., Nielsen R. and Siepel A. (2008). Patterns of positive selection in six mammalian genomes. – *PLOS Genetics,* **4**: e1000144.

Kreitman M. 2000. Methods to detect selection in populations with applications to the human. -*Annual Review of Genomics and Human Genetics*, **1**: 539–559.

Lemmon A.R. and Moriarty E.C. 2004. The importance of proper model assumption in Bayesian phylogenetics. – *Systematic Biology*, **53**: 265-277.

Liu C., Xu H., Lam S.H. and Gong Z. 2013. Selection of reliable

biomarkers from PCR array analyses using relative distance computational model: methodology and proof-of-concept study. – *PLoS ONE*, **8**: E83954.

Liu Y., Wu W.J., Zuo B., Ren Z.Q. and Xiong Y.Z. 2011. Polymorphism in coding region of pig PRDX6 gene and its genetic effects analysis. – *Yi Chuan,* **33**: 743-748.

Loo G.H., Sutton D.L. and Schuller K.A. 2012. Cloning and functional characterisation of a peroxiredoxin 1 (NKEF A) cDNA from Atlantic salmon (*Salmo salar*) and its expression in fish infected with *Neoparamoeba perurans*. – *Fish Shellfish Immunology*, **32**: 1074-1082.

Loumaye E., Andersen A.C., Clippe A., Degand H., Dubuisson M., Zal F., Morsomme P., Rees J.F. and Knoops B. 2008. Cloning and characterization of Arenicola marina peroxiredoxin 6, an annelid two-cysteine peroxiredoxin highly homologous to mammalian one-cysteine peroxiredoxins. – *Free Radical Biology and Medicine*, **45**: 482-493.

Lu A., Hu X., Wang Y., Shen X., Li X., Zhu A., Tian J., Ming Q. and Feng Z. 2014. iTRAQ analysis of gill proteins from the zebrafish (*Danio rerio*) infected with *Aeromonas hydrophila*. – *Fish Shellfish Immunology*, **36**: 229-239.

Marchini J.F., Brown K.K., Stamler J.S., Colagiovanni D.B., Rosenthal G.J.,

Mu C., Zhao J., Wang L., Song L., Zhang H., Li C., Qiu L. and Gai Y. 2009. Molecular cloning and characterization of peroxiredoxin 6 from Chinese mitten crab *Eriocheir sinensis*. – *Fish Shellfish Immunol*ogy, **26**:821-827.

Mukaigasa K., Nguyen L.T., Li L., Nakajima H., Yamamoto M. and Kobayashi M. 2012. Genetic evidence of an evolutionarily conserved role for Nrf2 in the protection against oxidative stress. – *Molecular and Cellular Biology*, **32**: 4455-4461.

Murrell B., Wertheim J.O., Moola S., Weighill T., Scheffler K. and Kosakovsky Pond S.L. 2012. Detecting individual sites subject to episodic diversifying selection. – *PLOS Genetics*, **8**: e1002764.

Nikapitiya C., Kang H.-S. and Lee J. 2006. Novel gene discovery from *Haliotis discus discus* by normalized cDNA library analysis. Unpublished.

Notredame C., Higgins D.G. and Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. – *Journal of Molecular Biology*, **302**: 205–217.

Okado-Matsumoto A., Matsumoto A., Fujii J. and Taniguchi N. 2000. Peroxiredoxin IV is a secretable protein with heparin-binding properties under reduced conditions. – *Journal of Biochemistry*, **127**: 493–501.

Pacifici F., Arriga R., Sorice G.P., Capuani B., Scioli M.G., Pastore D., Donadel G., Bellia A., Caratelli S., Coppola A., Ferrelli F., Federici M., Sconocchia G., Tesauro M., Sbraccia P., Della-Morte D., Giaccari A., Orlandi A. and Lauro D. 2014. Peroxiredoxin 6, a novel player in the pathogenesis of diabetes. – *Diabetes*, **63**:3210-3220.

Paula F.M., Ferreira S.M., Boschero A.C. and Souza K.L. 2013. Modulation of the peroxiredoxin system by cytokines in insulin-producing RINm5F cells: down-regulation of PRDX6 increases susceptibility of beta cells to oxidative stress. – *Molecular and Cellular Endocrinology*, **374**: 56-64.

Pei P., Wen C. and Hu B. 2010. Molecular cloning, characterization and expression analysis of peroxiredoxin 6 from Cristaria plicata and the antioxidant activity of its recombinant protein. Unpublished.

Perez-Sanchez J., Bermejo-Nogales A., Calduch-Giner J.A., Kaushik S. and Sitja-Bobadilla A. 2011. Molecular characterization and expression analysis of six peroxiredoxin paralogous genes in gilthead sea bream (*Sparus aurata*): Insights from fish exposed to dietary, pathogen and confinement stressors. – *Fish Shellfish Immunology*, **31**: 294-302.

Pérez-Sánchez J., Bermejo-Nogales A., Calduch-Giner J.A., Kaushik S. and Sitjà-Bobadilla A. 2011. Molecular characterization and expression analysis of six peroxiredoxin paralogous genes in gilthead sea bream (*Sparus aurata*): insights from fish exposed to dietary, pathogen and confinement stressors. – *Fish Shellfish Immunology*, **31**: 294–302.

Phalen T.J., Weirather K., Deming P.B., Anathy V., Howe A.K., van der Vliet A., Jönsson T.J., Poole L.B. & Heintz N.H. 2006. Oxidation state governs structural transitions in peroxiredoxin II that correlate with cell cycle arrest and recovery. – *Journal of Cell Biology*, **175**: 779–789.

Posada D. 2008. jModelTest: phylogenetic model averaging. – *Molecular Biology and Evolution*, **25**: 1253–1256.

Radyuk S.N., Sohal R.S. and Orr W.C. 2001. Thioredoxin peroxidases can foster cytoprotection or cell death in response to different stressors: over- and under-expression of thioredoxin peroxidase in *Drosophila* cells. – *Biochemistry Journal*, **371**: 743–752.

Rhee S.G. and Woo H.A. 2011. Multiple functions of peroxiredoxins: peroxidases, sensors and regulators of the intracellular messenger H_2O_2, and protein chaperones. – *Antioxidants and Redox Signaling*, **15**: 781–794.

Rhee S.G., Woo H.A., Kil I.S. and Bae S.H. 2012. Peroxiredoxin functions as a peroxidase and a regulator and sensor of local peroxides. – *Journal of Biological Chemistry*, **287**: 4403–4410.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A. and Huelsenbeck J.P. 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. – *Systematic Biology*, **61**: 539–542.

Saccoccia F., Di Micco P., Boumis G., Brunori M., Koutris I., Miele A.E., Morea V., Sriratana P., Williams D.L., Bellelli A. and Angelucci F. 2012. Moonlighting by different stressors: crystal structure of the chaperone species of a 2-Cys peroxiredoxin. – *Cell,* **20**: 429–439.

Salmeri M., Motta C., Mastrojeni S., Amodeo A., Anfuso C.D., Giurdanella G., Morello A., Alberghina M., Toscano M.A. and Lupo G. 2012. Involvement of PKCalpha-MAPK/ERK-phospholipase A(2) pathway in the *Escherichia coli* invasion of brain microvascular endothelial cells. – *Neuroscience Letters*, **511**: 33-37.

Shafer M.E., Willson J.A. and Damjanovski S. 2011. Expression analysis of the peroxiredoxin gene family during early development in *Xenopus laevis*. – *Gene Expression Patterns*, **11**: 511-516.

Soito L., Williamson C., Knutson S.T., Fetrow J.S., Poole L.B. and Nelson K.J. 2011. PREX: PeroxiRedoxin classification indEX, a database of subfamily assignments across the diverse peroxiredoxin family. – *Nucleic Acids Research*, **39**: 332–337.

Song W., Wang C., Shao C., Mu C., Zhang Y. and Zhan P. 2010. Molecular cloning and characterization of a novel peroxiredoxin 6 protein from *Sepiella maindroni*. Unpublished.

Sorokina E.M., Feinstein S.I., Zhou S. and Fisher A.B. 2011. Intracellular targeting of peroxiredoxin 6 to lysosomal organelles requires MAPK activity and binding to 14-3-3ε. – *American Journal of Physiology-Cell Physiology*,**300**: 1430–1441.

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo, M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., Mc Kernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A. 2002. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. – *Proceedings of the National Academy of Sciences U.S.A.*, **99**: 16899-16903.

Sutton D.L., Loo G.H., Menz R.I. and Schuller K.A. 2010 Cloning and functional characterization of a typical 2-Cys peroxiredoxin from southern bluefin tuna (*Thunnus maccoyii*). – *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **156**: 97-106.

Valentine J.W. 1997. Cleavage patterns and the topology of metazoan tree of life. – *Proceedings of the National Academy of Sciences U. S. A.,* **94**: 8001–8005.

Woo H.A., Yim S.H., Shin D.H., Kang D., Yu D.Y. and Rhee S.G. 2010. Inactivation of peroxiredoxin I by phosphorylation allows localized H_2O_2 accumulation for cell signaling. -*Cell*, **140**: 517–528.

Wood Z.A., Poole L.B. and Karplus P.A. 2003. Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. – *Science*, **300**: 650–653.

Wood Z.A., Schröder E., Robin Harris J. and Poole L.B. 2003. Structure, mechanism and regulation of peroxiredoxins. – *Trends in Biochemical Sciences*, **28**: 32–40.

Yang Z. and Bielawski J.P. 2000. Statistical methods for detecting molecular adaptation. – *Trends in Ecology and Evolution*, **15**: 496–503.

Yeh H.-Y. and Klesius P.H. 2007. cDNA cloning, characterization, and expression analysis of channel catfish (*Ictalurus punctatus* Rafinesque, 1818) Peroxiredoxin 6 gene. – *Fish Physiology and Biochemistry* **33**: 233-239.

Received: 18.03.2015
Accepted: 13.08.2015