



ELSEVIER

Contents lists available at ScienceDirect

Free Radical Biology and Medicine

journal homepage: www.elsevier.com/locate/freeradbiomed

Bioinformatics analyses provide insight into distant homology of the Keap1–Nrf2 pathway



Ranko Gacesa^a, Walter C. Dunlap^a, Paul F. Long^{a,b,*}

^a Institute of Pharmaceutical Science, King's College London, 150 Stamford Street, London SE1 9NH, UK

^b Department of Chemistry, King's College London, 150 Stamford Street, London SE1 9NH, UK

ARTICLE INFO

Article history:

Received 30 January 2015

Received in revised form

3 June 2015

Accepted 6 June 2015

Available online 25 June 2015

Keywords:

Adaptation of microbiota
Distant Homology Search Pipeline
Mycosporine-like amino acids
Mycosporines
Protein modeling
Sordariomycetes
Taxonomy Landscape Mapper
Virtual screening

ABSTRACT

An essential requirement for the evolution of early eukaryotic life was the development of effective means to protect against metabolic oxidative stress and exposure to environmental toxicants. In present-day mammals, the master transcription factor Nrf2 regulates basal level homeostasis and inducible expression of numerous detoxifying and antioxidant genes. To examine early evolution of the Keap1–Nrf2 pathway, we present bioinformatics analyses of distant homology of mammalian Keap1 and Nrf2 proteins across the Kingdoms of Life. Software written for this analysis is made freely available on-line. Furthermore, utilizing protein modeling and virtual screening methods, we demonstrate potential for Nrf2 activation by competitive inhibition of its binding to Keap1, specifically by UV-protective fungal mycosporines and marine mycosporine-like amino acids (MAAs). We contend that coevolution of Nrf2-activating secondary metabolites by fungi and other extant microbiota may provide prospective compound leads for the design of new therapeutics to target activation of the human Keap1–Nrf2 pathway for treating degenerative diseases of ageing.

© 2015 Published by Elsevier Inc.

1. Introduction

The emergence of oxygenic photosynthesis, evolved first by proto-cyanobacteria approximately 3.4 billion years ago, gave rise to the Earth's oxygen atmosphere, rendering subsequent progression to eukaryotic and metazoan life possible [1]. Such an oxidative environment, however, posed a significant challenge to early life forms, requiring effective means of oxidative cytoprotection. In mammals, the Kelch-like ECH-associated protein 1 (Keap1) forms a complex with the nuclear factor erythroid 2-related factor 2 (Nrf2). The Keap1–Nrf2 complex dissociates in response to reactive oxygen species (ROS), releasing Nrf2 that binds to the nuclear antioxidant response element (ARE) to coordinate transcription of multiple antioxidant, detoxifying, and cell survival genes [2,3]. Belonging to the 'cap-n-collar' family of transcription factors that have a distinct basic leucine–zipper motif [4], the

domain elements of Nrf2 are highly conserved across many diverse species, with orthologs having been detected in *Caenorhabditis elegans* (SKN-1) [5], *Drosophila melanogaster* (Nrf2-like) [6], and yeast (YAP-1) [7]. A prokaryotic homolog of Nrf2 (possibly OxyR or SoxR) [8,9] has also been suggested to protect UV-tolerant bacteria by augmenting coenzyme Q reduction via activation of cellular NAD(P)H: quinone oxidoreductase (NQOR) [10,11]. We contend that early adaptive features of the Keap1–Nrf2 pathway conserved in extant microbiota may serve as a novel pharmacomimetic model for the discovery of new therapeutic activators of the human oxidative stress response that may retard the progression of age-related degenerative disease, stimulate the innate immune response, and suppress carcinogenesis [12–15]. Accordingly, a new bioinformatics conduit to search and map distant homology has been developed and, in addition, Bayesian inference methods have been used to construct phylogenetic trees of Keap1–

Abbreviations: ADT, AutoDock Tools; ARE, antioxidant response element; BLAST, Basic Local Alignment Search Tool; DHSP, Distant Homology Search Pipeline; NTMT, NCBI Taxonomy Mapping Tool; HMM, Hidden Markov Model; HMMER3, HMM-based utility version 3; Keap1, Kelch-like ECH-associated protein 1; KEGG, Kyoto Encyclopedia of Genes and Genomes; MAAs, mycosporine-like amino acids; MCMC, Metropolis–Hastings Markov Chain Monte Carlo; NQOR, NAD(P)H: quinone oxidoreductase; NCBI, National Center for Biology Information; NR, nonredundant; Nrf2, nuclear factor erythroid 2-related factor 2 protein; OxyR, protein transcriptional autoregulator of bacterial hydrogen-inducible genes; PDB, Protein Data Bank; psi-BLAST, Position Specific Iterative Basic Local Alignment Search Tool; ROS, reactive oxygen species; SKN-1, *Caenorhabditis elegans* protein homolog of Nrf2; SoxR, protein transcriptional autoregulator of bacterial superoxide-inducible genes; TLM, Taxonomy Landscape Mapper; UCFS, University of California, San Francisco; UFF, universal force field; YAP1, yeast AP-1 protein homolog of Nrf2

* Corresponding author at: Institute of Pharmaceutical Science, King's College London, 150 Stamford Street, London SE1 9NH United Kingdom. Fax: +44 207 484 4842.

E-mail address: paul.long@kcl.ac.uk (P.F. Long).

<http://dx.doi.org/10.1016/j.freeradbiomed.2015.06.015>

0891-5849/© 2015 Published by Elsevier Inc.

Nrf2 evolution across major eukaryotic taxa. A protein model and virtual screen were also established to predict likely activation of the Keap1–Nrf2 pathway utilizing a library of structurally diverse natural products [16,17].

2. Materials and methods

2.1. Data retrieval

Custom databases of archaeal, bacterial, and fungal proteins were constructed from the National Center for Biology Information (NCBI) nonredundant (NR) database [18] and the NCBI Taxonomy database [19]. Sequences of human Keap1 and Nrf2 proteins, together with known homologs and predicted orthologs, were acquired from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [20] and are displayed in [Supplementary Data File 1](#). A novel distant homology search pipeline (called DHSP) was developed to increase the sensitivity and precision of distant homology searches by utilizing multiple Hidden Markov models (HMMs). The pipeline, described in [Supplementary Data File 2](#) and freely available at <https://github.com/rgacesa/DHSP>, performs a psi-BLAST [21] sequence alignment search against the NR database to detect close homology to generate HMM models. DHSP uses HMMER3, a HMM-based sequence alignment tool [22], for high sensitivity distant homology detection. To minimize false positive hits, DHSP performs searches using multiple HMM models employing the Smith–Waterman algorithm [23] to align potential distant homologs against known sequences, thereby filtering out those sequences that fail to align. A tool for mapping distant homologs against the NCBI taxonomic database was additionally developed. This tool, called Taxonomy Landscape Mapper (TLM), displays results in a user friendly visual format as described in [Supplementary Data File 3](#). TLM is made freely available for use at <https://github.com/rgacesa/TLM>.

2.2. Phylogenetic reconstruction of Keap1 and Nrf2 homology

Multiple alignments of Keap1 and Nrf2 homologs detected by DHSP and TLM were constructed utilizing ClustalW2 [24]. Phylogenetic reconstruction of Keap1 and Nrf2 proteins were assembled using the MrBayes 3.2 Bayesian inference analysis tool [16] with a mixed model (aamodelpr=mixed) for automatic estimation of the amino acid matrix during a Metropolis–Hastings Markov Chain Monte Carlo (MCMC) simulation, which was run for 2,000,000 generations using 25 chains. Post-run examination of parameters indicated convergence to the JTT amino acid substitution model (with posterior probability of 100%). Among-site rate variation was set to the gamma model with 4 categories. Other parameters of MrBayes run were left at default values. Probability and tree summaries were inspected manually to confirm simulation convergence, and all simulations resulting in “good convergence” were accepted if the average standard deviation of split frequencies was < 0.01 with a convergence value (Potential Scale Reduction Factor) approaching 1.0 [25]. Phylogenetic trees were inspected and edited using Archaeopteryx [26] and MEGA 6.1 [27] software tools.

2.3. Virtual screen for competitive inhibitors of Keap1–Nrf2 binding

Published data [28] extracted from the Protein Data Bank (PDB) [29] entitled “Crystal Structure of the Kelch–Neh2 Complex” (PDB-2FLU, <http://www.rcsb.org/pdb/explore.do?structureId=2flu>) was used to construct a virtual compound screening model to predict the release of Nrf2 by competitive inhibition of Keap1–Nrf2 binding. A model of the human Keap1–Nrf2 interaction was

prepared for UCSF DOCK 6.0 [30] and AutoDock Vina [31] virtual screening algorithms employing standard protocols described in [Supplementary Data File 4](#). For ligand assessments, including mycosporine-like amino acid (MAA) predictions, the Avogadro chemical utility platform [32] was used to create and optimize compound binding models by molecular dynamics simulation using universal force field (UFF) parameters [33] and the steepest descent algorithm [34]. To assess MAA–Keap1 docking results, a comparison set of approximately 1100 Brazilian natural products was assembled from two ZINC catalogs [35–37] for predictive contrast. All ligands were prepared for docking using the AutoDock Tools (ADT) script “prepare_ligand4.py” (for AutoDock Vina) and the UCSF Chimera visualization toolkit [38] (for DOCK6). Both AutoDock Vina and DOCK6 were configured for high docking precision with DOCK6 selected for a flexible docking protocol with 2000 orientations per ligand and 400 iterations for energy minimization, and the AutoDock Vina space search exhaustiveness was set to 20. The UCSF Chimera ViewDock utility was used to manually examine the docking results. Ligands were assessed first by the number of potential hydrogen bonds available for binding within the protein binding pocket, and all ligands forming less than 3 hydrogen bonds were rejected. Those selected were evaluated manually, and all ligands without potential binding to critical positions of the Keap1–Nrf2 docking pocket were discarded. Final evaluation of “viable” docking ligands was performed by assessing the combined docking scores calculated as $2 \times$ Vina docking energy + DOCK6 binding energy + DOCK6 docking score. Each variable in the equation was scaled by subtracting the mean value and dividing by the standard deviation.

3. Results

3.1. Data mining of microbial protein databases

Databases of archaea, bacteria, fungi, and plant proteins were analyzed for distant homology to human Keap1 and Nrf2 proteins using the newly developed software tools we named the Distant Homology Search Pipeline (DHSP) and the Taxonomy Landscape Mapper (TLM). All databases except for archaea were found to contain high numbers of Keap1 homologs ([Supplementary Data File 5](#)). In contrast, close homologs to human Nrf2 domain Neh1–Neh6 sequences were detected only in the database of fungal proteins ([Supplementary Data Files 6 and 7](#)), primarily in Ascomycetes belonging to the Class *Sordariomycetes*, many of which are insect and plant pathogens ([Table 1](#)).

[Table 1](#) shows species selected based on the prediction score of detected homologs to Keap1 and Nrf2 proteins, all having the presence of genes for mycosporine-like amino acid (MAA) biosynthesis ([Supplementary Data File 8](#)). Nrf2 prediction scores are

Table 1
Keap1 and Nrf2 protein scoring of sequence homology in fungal genomes

Species (common name)	Prediction score	
	Nrf2	Keap1
<i>Cordyceps militaris</i> (Scarlet caterpillar club fungus)	8	46
<i>Beauveria bassiana</i> (White muscardine fungus)	7	43
<i>Fusarium graminearum</i> (Wheat head blight fungus)	10	66
<i>Colletotrichum gloeosporioides</i> (Postharvest fruit rot fungus)	8	79
<i>Magnaporthe oryzae</i> (Rice blast fungus)	15	79
<i>Fusarium pseudograminearum</i> (Wheat crown rot fungus)	10	60
<i>Verticillium dahlia</i> (Verticillium wilt fungus)	10	40
<i>Colletotrichum higginsianum</i> (Crucifer anthracnose fungus)	14	54
<i>Metarhizium acridum</i> (Green muscardine fungus)	13	38
<i>Colletotrichum graminicola</i> (Maze anthracnose fungus)	15	73

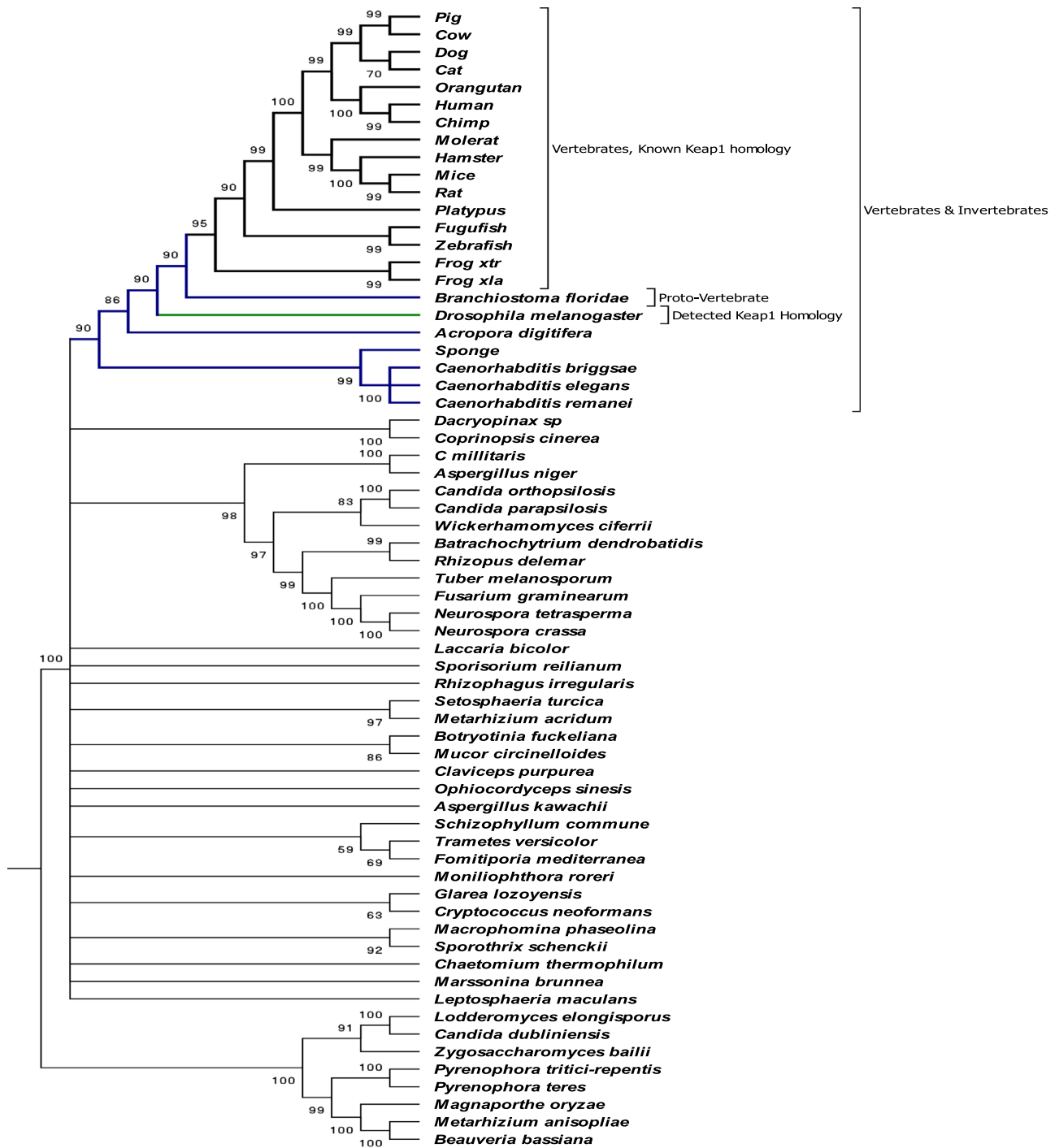


Fig. 1. Bayesian phylogenetic reconstruction of Keap1 evolution. The phylogenetic tree was constructed from 16 vertebrate Nrf2 homologs, 7 invertebrate homologs, and 42 fungal homologs. Those experimentally confirmed are marked as “detected” homologs and sequences predicted as homologous in prior accounts are marked as “presumed” homologs. Posterior probabilities (in percent) are displayed and splits below 50% are collapsed.

the sum of numbers of detected homologs for Nrf2 domain Neh1–Neh6 conserved sequences. The Keap1 prediction scores are the sum of Keap1 conserved kelch1–kelch6 and BTB domain sequences.

3.2. Phylogenetic reconstruction of Keap1–Nrf2 homologies

In order to examine evolution of the Keap1–Nrf2 pathway, we performed phylogenetic reconstruction using predicted Keap1 and Nrf2 fungal homologs, as well as a selection of known homologs from key animal species. These animals comprise invertebrate and

vertebrate species commonly used as model organisms in biology that include three species from the genus *Caenorhabditis* and the fruit fly *Drosophila melanogaster*[5,6]. The sponge *Amphimedon queenslandica* and coral *Acropora digitifera* were chosen as examples of early metazoan taxa (phyla Porifera and Cnidaria). Vertebrate sequences were chosen from common model organisms (Rat, Mouse, Zebrafish, and frog *Xenopus laevis*) and their close relatives. *Platypus* was chosen as an example of early mammals and the lancelet *Branchiostoma floridae* was selected to represent an ancient vertebrate animal. All amino acid sequences are listed in [Supplementary Data File 9](#). Bayesian reconstructions of

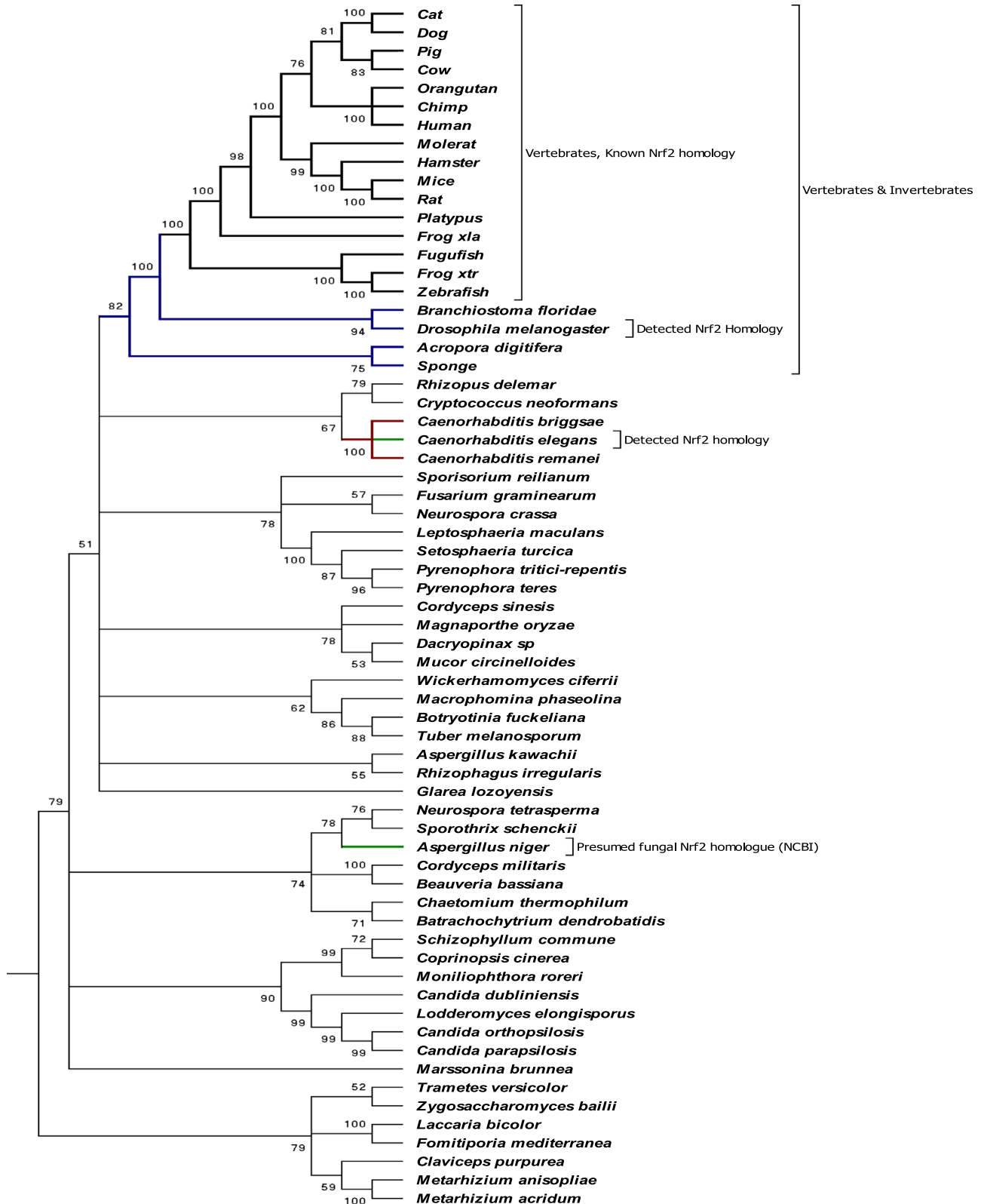


Fig. 2. Bayesian phylogenetic reconstruction of Nrf2 evolution. The phylogenetic tree was constructed from 16 vertebrate Nrf2 homologs, 7 invertebrate homologs, and 42 fungal homologs. Those experimentally confirmed are marked as “detected” homologs and those predicted from the literature are marked as “presumed” homologs. Posterior probabilities (in percent) are displayed and splits below 50% are collapsed.

both Keap1 phylogeny (Fig. 1) and Nrf2 phylogeny (Fig. 2) were consistent with conventional species evolution, with the expected grouping of major vertebrate taxa and showing a clear split between the vertebrates and the invertebrates. Fungal Keap1 and

Nrf2 sequences, however, were both highly divergent and could be grouped into three clades. Surprisingly, phylogenetic reconstruction of Nrf2 homologs placed all three *Caenorhabditis* proteins within the fungal groupings (Fig. 2), instead of residing with other

Table 2
Virtual screening results

Compound	Score	Structure assignment	Biological function
ZINC49048037	0.75	AGN-PC-07CJ71	Acetylcholinesterase inhibitor [41]
ZINC15120547	-2.66	Crassinervic acid	Antifungal [42]
ZINC00622123	0.77	Griseofulvin	Antifungal [43]
ZINC13411177	0.02	Similar to Strictifolione	Antifungal [44]
ZINC14447808	1.76	AGN-PC-077JEH	Antifungal [45]
ZINC40973915	-9.01	Similar to Ixoside	Antioxidant [46]
ZINC31157290	-2.60	Secoxyloganin	Antioxidant [47]
ZINC05998957	-2.17	Lirioresinol A	Antioxidant [48]
ZINC15119278	-1.04	Similar to Yatein	Antioxidant [49]
ZINC00898006	-0.15	Rubrofusarin	Antioxidant [50]
ZINC02563652	-0.04	Alloisimperatorin	Antioxidant [51]
ZINC01580260	0.23	Cleomiscosin A	Antioxidant [52]
ZINC69482380	1.63	Similar to Maclurin	Antioxidant [53]
ZINC06037073	-0.97	Similar to Emodin	Cytotoxic, anticancer [54]
ZINC84154280	-2.54	Geranyloxy- <i>p</i> -benzoic acid	Farnesoid X receptor agonist [55]
ZINC26490614	-2.69	Procyanidin B2	Nrf2 activator [56]
ZINC30726399	-9.93	Betanidin	Nrf2 activator [57]
ZINC69482045	-6.62	Similar to Ursolic acid	Nrf2 activator [58]
ZINC69481913	-6.40	Similar to Ursolic acid	Nrf2 activator [58]
ZINC17263588	-6.17	Chlorogenic acid	Nrf2 activator [59]
ZINC84154032	-5.75	Similar to Morroniside	Nrf2 activator [60]
ZINC84153764	-4.32	Similar to Morroniside	Nrf2 activator [60]
ZINC04102166	-4.28	Geniposidic acid	Nrf2 activator [61]
ZINC01714287	-3.40	Piperine	Nrf2 activator [62]
ZINC03870412	-3.06	Epigallocatechin gallate (EGCG)	Nrf2 activator [63]
ZINC00073693	-2.12	Pinocembrin	Nrf2 activator [64]
ZINC12428433	-1.84	Butein	Nrf2 activator [65]
ZINC71316232	-1.69	Similar to Chlorogenic acid	Nrf2 activator [59]
ZINC01531693	-1.57	Similar to Piperine	Nrf2 activator [62]
ZINC03872070	-1.52	Chrysin	Nrf2 activator [66]
ZINC00897734	-1.50	Similar to Quercetin	Nrf2 activator [67]
ZINC00156701	-1.41	Naringenin	Nrf2 activator [68]
ZINC00113309	1.69	Fraxetin	Nrf2 activator [69]
ZINC01561070	-0.11	Similar to Quercetin	Nrf2 activator [67]
ZINC14728348	0.14	Similar to Quercetin	Nrf2 activator [67]
ZINC05733652	-1.36	Diosmetin	Potential Nrf2 activator, antioxidant [70]
ZINC3382113	-1.73	Similar to Phlorizin	Potential Nrf2 activator [71]
ZINC69482290	-3.37	Similar to Glucoerucin	Potential Nrf2 activator [72]
ZINC05733537	-0.86	Ermanin, similar to Quercetin (Nrf2 activator)	Potential Nrf2 activator [67]
ZINC84153966	-3.86	Similar to Acetoside	Potential Nrf2 activator [73]
ZINC13108875	-2.42	Similar to Burchellin	Potential pesticide [74]
mycosporine glycine-valine	0.24	Mycosporine-like amino acid	UV-protectant, antioxidant [75]
mycosporine glycine	4.21	Mycosporine-like amino acid	UV-protectant, antioxidant [75]
MAA, Porphyra 334	0.58	Mycosporine-like amino acid	UV-protectant, antioxidant [75]
ZINC15252691	-5.23	Gaudichaudianic acid	Trypanicide [76]

invertebrates as would be expected.

3.3. Protein modeling and virtual screening of Nrf2 activation

Fungal genomes typically encode enzymes that express the biosynthesis of mycosporines and the related mycosporine-like amino acid (MAA) family of UV-protective and antioxidant metabolites [39,40]. In order to assess whether MAAs have potential to initiate a protective response through the Keap1–Nrf2 pathway, we performed a virtual screen of approximately 1100 diverse natural products including 20 MAAs. Of the ligands tested, 75 met the criteria for potential inhibitors of the Keap1–Nrf2 interaction.

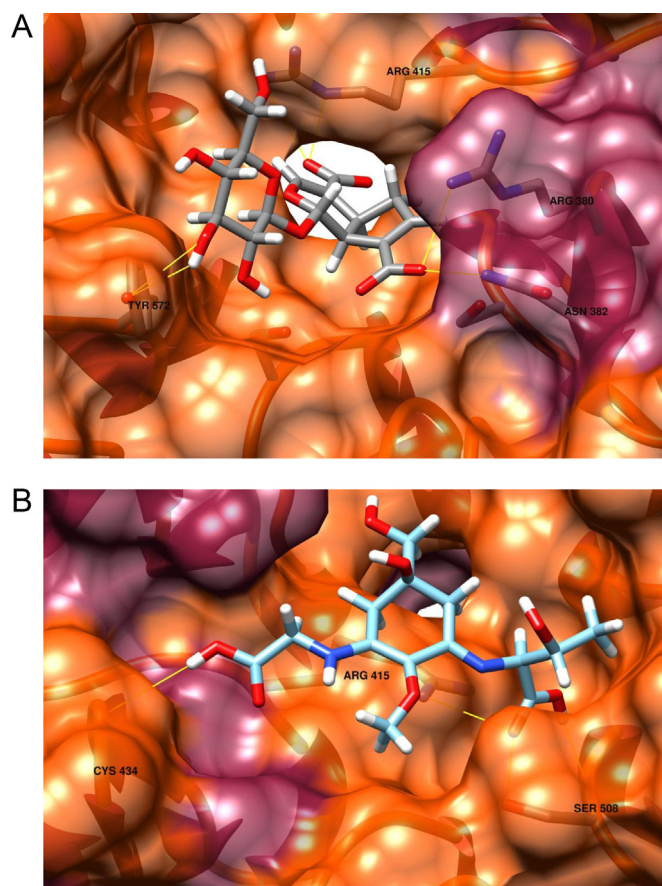


Fig. 3. Betanidin (A) and porphyra-334 (B) substrate docking models. The protein binding models depict the cross section of only the molecular surface of the human Keap1 kelch-like repeats β -propeller docking pocket. Predicted hydrogen bonds between ligand and Keap1 are depicted in yellow and amino acid residues involved in the formation of bonds are labeled.

These criteria were determined by the docking position of the ligand within the Keap1–Nrf2 interaction pocket, the potential to form hydrogen bonds with Keap1, and importantly the docking score. Out of the 75 compounds (Table 2), 25 are known to be Nrf2 activators, while another 11 compounds are known antioxidants but not reported previously to activate Nrf2. These 11 compounds included 3 MAAs (mycosporine–glycine–valine, mycosporine–glycine and porphyra-334). Examples for the binding of betanidin and porphyra-334 within the docking region of Keap1 are shown in Fig. 3.

Listed compounds in Table 2 are deemed “viable” according to their docking score, number of potential hydrogen bonds within the Keap1–Nrf2 binding pocket, and by manual inspection of the docking profile. The table is sorted according to the biological function of docked ligands. Entries are given only for compounds with assigned structures; a complete list with unknown structure assignments is given in Supplementary Data File 10

4. Discussion

The Keap1–Nrf2 pathway is a major regulator of antioxidant protection in mammalian cells, and is responsible for the transcription of over 200 cytoprotective genes encoded by the nuclear antioxidant response element (ARE). Given that the Keap1–Nrf2 pathway is so important for cytoprotection in mammals, it might be expected that homology is evolutionary preserved from simple progenitors. Such is consistent with a homologous Keap1–Nrf2

pathway confirmed in *Drosophila melanogaster* [6] and that identified in *Caenorhabditis elegans* [5]. Our distant homology search has revealed, for the first time, that Keap1 and Nrf2 homologs are present in fungal taxa (Table 1) and absent in bacteria, archaea, and plants (Supplementary Date File 4). The presence of both Keap1 and Nrf2 homologs in fungi, and that fungi are evolutionary closer to animals [77] than all other taxa examined, encouraged further investigation. Phylogenetic reconstruction of Keap1 homology in key species from vertebrate, invertebrate, and fungal taxa (Fig. 1) demonstrated that Keap1 evolution fits as expected within the “Tree of Life” [78]. However, phylogenetic reconstruction of Nrf2 (Fig. 2) showed an unusual discrepancy in the placement of the genus *Caenorhabditis*. Genomic data mining affirms that *C. elegans* does encode proteins highly similar to Keap1 (Supplementary Data File 11), consistent with its assigned position in the Keap1 phylogenetic tree (Fig. 1). These Keap1-like proteins have not been implicated in inhibition the Nrf2 homolog of *C. elegans*, specifically the protein designated SKN-1 [5,79], and closer phylogenetic examination of this functional placement is thus warranted. Instead, SKN-1 activity is regulated by interaction with protein WDR-23 in a manner strikingly similar to Nrf2–Keap1–Cul3 ubiquitination and degradation [80]. Interestingly, while WDR-23 has no significant sequence similarity to Keap1 (see Supplementary Date File 11), the 3-D structure is remarkably similar to that of Keap1, with both proteins containing a beta-propeller superstructure. Unlike Keap1, however, the WDR-23 beta-propeller is formed from WD-40 protein motif repeats (see <http://www.rcsb.org/pdb/explore/explore.do?structureId=4LG9>) rather than kelch-like repeats (see <http://www.rcsb.org/pdb/explore.do?structureId=2flu>). Importantly, contemporary research has established that Nrf2 activity in *C. elegans* is controlled additionally by the human beta-transducin repeat-containing protein (β -TrCP) [81,82], which contains WD-40 protein motifs arranged into a beta-propeller superstructure [83]. Comparison between *C. elegans* WDR-23 and the human β -TrCP protein (Supplementary Data File 11) reveals significant similarity between these two WD-40 beta propeller proteins, indicating remarkably strong evolutionary continuity of function for Nrf2 control shared between worms and higher animals.

According to our predictions of Nrf2 homologs in various fungi (Fig. 1 and Table 1) and previous findings of the Nrf2 homolog YAP-1 in yeast [7], the Keap1–Nrf2 pathway has most certainly evolved after the eukarya separated from the prokarya, but prior to the fungal–metazoan split. Therefore, beta-propeller inhibitor proteins of Nrf2 may have evolved in early animals more than once by convergent evolution, which might explain the observed differences of Nrf2 activation in *C. elegans* and that of higher animals. Prior bioinformatics analysis of the sea anemone *Nematostella vectensis* has also detected homologs of Nrf2, Keap1 and two small binding Maf proteins required for Nrf2–ARE gene promotion [60]. Our analyses of coral (*Acropora digitifera*) and sponge (*Amphimedon queenslandica*) genomes also revealed more than one small Maf homolog [data not shown], in addition to encoding both Keap1 and Nrf2 homologs. These results indicate that species within the genus *Caenorhabditis* may have lost SKN-1 regulation by a Keap1-like protein (possibly R12E2.1), but have retained the β -TrCP-like inhibitor WDR-23. This loss may be due to a lack of evolutionary pressure associated with the soil-dwelling, often hypoxic, lifestyle of these animals [84]. Further investigation by inclusion of additional deep-branching taxa is required; nevertheless, our phylogenetic analyses show clearly that the Keap1–Nrf2 pathway predates the fungal–metazoan divergence. Developing an evolutionary clock to determine if there is a correlation between the evolutionary origins of Nrf2 and the emergence of atmospheric oxygen on Earth is an avenue worthy of further consideration.

A virtual screening assay was performed to assess the potential for fungal metabolites to function as competitive inhibitors of Keap1–Nrf2 binding. Although there is no published data for the disruption of Keap1–Nrf2 binding by ixoside metabolites, our predictions (Table 2) match the high scoring antioxidants betanidin, chlorogenic acid, and compounds similar to ursolic acid, which are known activators of Nrf2 [58,85,86]. We found also that the MAAs, mycosporine–glycine, mycosporine–glycine–valine, and porphyra-334, may serve as viable docking ligands based on their docking score, docking profile, and potential to form critical hydrogen bonds within the Keap1–Nrf2 docking pocket. All three MAAs, often expressed in high cellular concentrations, are widely accepted to be UV-inducible sunscreen protectants [39] and mycosporine–glycine and porphyra-334 are reported also to have antioxidant properties [87,88]. While docking scores of MAAs are typically higher than many of the other viable ligands, implying potentially lower binding affinity, these compounds are predicted to form several critical hydrogen bonds with Keap1 binding pocket and have passed manual inspection of binding poses. Additional research will determine if these compounds may cause disruption of Keap1–Nrf2 binding to activate the transcription of nuclear ARE cytoprotective genes. Finding microorganisms with Keap1–Nrf2 homology offer an early evolutionary model for the adaptive signaling of the Keap1–Nrf2 pathway, as well as providing an endogenous source of stress-inducible metabolites having potential to activate the nuclear ARE for therapeutic consideration.

5. Conclusions

Data mining of microbial protein databases has revealed distant homology to Keap1 and Nrf2 proteins in fungi, especially amongst taxa of Phylum: Ascomycota / Class: Sordariomycetes. Phylogenetic reconstruction of Keap1–Nrf2 shows that the pathway evolved prior to the fungal–metazoan divergence. Unexpectedly, the Nrf2 evolutionary tree shows mismatch for genus *Caenorhabditis* within the expected taxonomic model, potentially from sequence degeneration of Nrf2 or lack of evolutionary pressure possibly due to the soil-dwelling lifestyle of these worms. Lastly, virtual screening for competitive inhibition of Keap1–Nrf2 binding predicts the potential for Nrf2 activation by UV-protective mycosporine-like amino acids.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

Financial support for this work has come from the British Medical Research Council of the United Kingdom (MRC Grant G82144A to R.G. and P.F.L.) in partnership with Prof. Daslav Hranueli, Dr. Jurica Zucko, and SemGen Ltd (Zagreb, Croatia). The authors thank Prof. Stephen Sturzenbaum, King’s College London, for suggestions regarding the physiology of *Caenorhabditis elegans*.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.freeradbiomed.2015.06.015](https://doi.org/10.1016/j.freeradbiomed.2015.06.015).

References

- [1] R. Hooper, Revealing the dawn of photosynthesis, *New Sci.* 14 (2006).
- [2] T.T.W. Kensler, N. Wakabayashi, S. Biswal, Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway, *Annu. Rev. Pharmacol. Toxicol.* 47 (2007) 89–116.
- [3] Q. Ma, Role of Nrf2 in oxidative stress and toxicity, *Annu. Rev. Pharmacol. Toxicol.* 53 (2013) 401–426.
- [4] H. Motohashi, M. Yamamoto, Nrf2-Keap1 defines a physiologically important stress response mechanism, *Trends Mol. Med.* 10 (2004) 549–557.
- [5] J.H. An, K. Vranas, M. Lucke, H. Inoue, N. Hisamoto, K. Matsumoto, et al., Regulation of the *Caenorhabditis elegans* oxidative stress defense protein SKN-1 by glycogen synthase kinase-3, *Proc. Natl. Acad. Sci. USA* 102 (2005) 16275–16280.
- [6] G.P. Sykiotis, D. Bohmann, Keap1/Nrf2 signaling regulates oxidative stress tolerance and lifespan in *Drosophila*, *Dev. Cell* 14 (2008) 76–85.
- [7] J. Lee, C. Godon, D. Spector, J. Garin, M.B. Toledano, G. Lagniel, et al., Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast, *J. Biol. Chem.* 274 (1999) 16040–16046.
- [8] M. Zheng, G. Storz, Redox sensing by prokaryotic transcription factors, *Biochem. Pharmacol.* 59 (2000) 1–6.
- [9] S.O. Kim, K. Merchant, R. Nudelman, W.F. Beyer, T. Keng, J. DeAngelo, et al., OxyR: A molecular code for redox-related signaling, *Cell* 109 (2002) 383–396.
- [10] W.C. Dunlap, A. Fujisawa, Y. Yamamoto, UV radiation increases the reduced coenzyme Q ratio in marine bacteria, *Redox Rep.* 7 (2002) 3–6.
- [11] W.C. Dunlap, A. Fujisawa, Y. Yamamoto, M. Inoue, Tropical UV-tolerant bacteria may provide a pharmacomimetic model for anti-ageing research and cancer prevention, *Mar. Biotechnol.* 6 (2004) S223–S230.
- [12] M.J. Calkins, D.A. Johnson, J.A. Townsend, M.R. Vargas, J.A. Dowell, T. P. Williamson, et al., The Nrf2/ARE pathway as a potential therapeutic target in neurodegenerative disease, *Antioxid. Redox Signal.* 11 (2009) 497–508.
- [13] S.J. Chapple, R.C.M. Siow, G.E. Mann, Crosstalk between Nrf2 and the proteasome: therapeutic potential of Nrf2 inducers in vascular disease and aging, *Int. J. Biochem. Cell Biol.* 44 (2012) 1315–1320.
- [14] B. Gao, A. Doan, B.M. Hybertson, The clinical potential of influencing Nrf2 signaling in degenerative and immunological disorders, *Clin. Pharmacol.* 6 (2014) 19–34.
- [15] H.M. Leinonen, E. Kansanen, P. Pölonen, M. Heinäniemi, A.L. Levonen, Role of the Keap1-Nrf2 pathway in cancer, *Adv. Cancer Res.* 122 (2014) 281–320.
- [16] F. Ronquist, M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, et al., MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space, *Syst. Biol.* 61 (2012) 539–542.
- [17] S. Cosconati, S. Forli, A.L. Perryman, R. Harris, D.S. Goodsell, A.J. Olson, Virtual screening with AutoDock: theory and practice, *Expert Opin. Drug Discov.* 5 (2010) 597–607.
- [18] D.L. Wheeler, D.M. Church, S. Federhen, A.E. Lash, T.L. Madden, J.U. Pontius, et al., Database resources of the National Center for Biotechnology, *Nucleic Acids Res.* 31 (2003) 28–33.
- [19] S. Federhen, The NCBI Taxonomy database, *Nucleic Acids Res.* 40 (2012) D136–D143.
- [20] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res.* 28 (2000) 27–30.
- [21] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25 (1997) 3389–3402.
- [22] S.R. Eddy, Accelerated profile HMM searches, *PLoS Comput. Biol.* 7 (2011) e1002195.
- [23] T.F. Smith, M.S. Waterman, Identification of common molecular subsequences, *J. Mol. Biol.* 147 (1981) 195–197.
- [24] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, et al., Clustal W and Clustal X version 2.0, *Bioinformatics* 23 (2007) 2947–2948.
- [25] A. Gelman, D. Rubin, Inference from iterative simulation using multiple sequences, *Stat. Sci.* 7 (1992) 457–472.
- [26] Zmasek, C. Archaeopteryx [Internet]; 2014. <https://sites.google.com/site/cmzmasek/home/software/archaeopteryx>.
- [27] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, *Mol. Biol. Evol.* 30 (2013) 2725–2729.
- [28] S.C. Lo, X. Li, M.T. Henzl, L.J. Beamer, M. Hannink, Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling, *EMBO J.* 25 (2006) 3605–3617.
- [29] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, et al., The Protein Data Bank, *Nucleic Acids Res.* 28 (2000) 235–242.
- [30] D.T. Moustakas, P.T. Lang, S. Pegg, E. Pettersen, I.D. Kuntz, N. Brooijmans, et al., Development and validation of a modular, extensible docking program: DOCK 5, *J. Comput. Aided Mol. Des.* 20 (2006) 601–619.
- [31] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multi-threading, *J. Comput. Chem.* 31 (2009) 455–461.
- [32] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminform.* 4 (2012) 1–17.
- [33] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard, W.M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *J. Am. Chem. Soc.* 114 (1992) 10024–10035.
- [34] S.A. Adcock, J.A. McCammon, Molecular dynamics: survey of methods for simulating the activity of proteins, *Chem. Rev.* 106 (2006) 1589–1615.
- [35] ZINC Catalog UFEFS Natural Products [Internet] <http://zinc.docking.org/catalogs/ufefsp/>.
- [36] ZINC Catalog Nubbe Natural Products [Internet] <http://zinc.docking.org/pbcs/nubbenp/>.
- [37] J.J. Irwin, B.K. Shoichet, ZINC—a free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model.* 45 (2005) 177–182.
- [38] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E. C. Meng, et al., UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612.
- [39] J.M. Shick, W.C. Dunlap, Mycosporine-like amino acids and related gadusols: biosynthesis, accumulation, and UV-protective functions in aquatic organisms, *Annu. Rev. Physiol.* 64 (2002) 223–262.
- [40] R.P. Sinha, S.P. Singh, D. Ha, Database on mycosporines and mycosporine-like amino acids (MAAs) in fungi, cyanobacteria, macroalgae, phytoplankton and animals, *J. Photochem. Photobiol.* 89 (2007) 29–35.
- [41] AGN-PC-07CJ71 (CID 45378270)—Compound BioActivity Data [Internet] <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?cid=45378270>.
- [42] W. Xu, X. Li, Antifungal compounds from piper species, *Curr. Bioact. Compd.* 7 (2011) 262–267.
- [43] K. Gull, A.P. Trinci, Griseofulvin inhibits fungal mitosis, *Nature* 244 (1973) 292–294.
- [44] S. BouzBouz, J. Cossy, Total synthesis of (+)-strictifolione, *Org. Lett.* 5 (2003) 1995–1997.
- [45] AGN-PC-077JEH | C13H14O4 - PubChem [Internet] http://pubchem.ncbi.nlm.nih.gov/compound/1_-Acetoxychavicol_acetate.
- [46] W.M. Abdel-Mageed, E.Y. Backheet, A.A. Khalifa, Z.Z. Ibraheem, S.A. Ross, Antiparasitic antioxidant phenylpropanoids and iridoid glycosides from *Tecoma mollis*, *Fitoroterapia.* 83 (2012) 500–507.
- [47] S. De Marino, C. Festa, F. Zollo, A. Nini, L. Antenucci, G. Raimo, et al., Antioxidant activity and chemical components as potential anticancer agents in the olive leaf (*Olea europaea* L. cv Leccino.) decoction, *Anticancer Agents Med. Chem.* 14 (2014) 1376–1385.
- [48] L.Y. Wang, N. Unehara, S. Kitanaka, Lignans from the roots of *Wikstroemia indica* and their DPPH radical scavenging and nitric oxide inhibitory activities, *Chem. Pharm. Bull. (Tokyo)* 53 (2005) 1348–1351.
- [49] J.K. DaSilva, E.H.A. Andrade, M.J. Kato, L.M. Carreira, E.F. Guimaraes, J.G.S. Maia, Antioxidant capacity and larvicidal and antifungal activities of essential oils and extracts from *Piper krukoffii*, *Nat. Prod. Comm.* 6 (2011) 1361–1366.
- [50] J.S. Choi, H.J. Lee, S.S. Kang, Alaternin, cassiaside and rubrofusarin gentiobioside, radical scavenging principles from the seeds of *Cassia tora* on 1,1-diphenyl-2-picrylhydrazyl (DPPH) radical, *Arch. Pharm. Res.* 17 (1994) 462–466.
- [51] X.L. Piao, I.H. Park, S.H. Baek, H.Y. Kim, M.K. Park, J.H. Park, Antioxidative activity of furanocoumarins isolated from *Angelica dahurica*, *J. Ethnopharmacol.* 93 (2004) 243–246.
- [52] W. Jin, P.T. Thuong, N.D. Su, B.S. Min, K.H. Son, H.W. Chang, et al., Antioxidant activity of cleomiscosins A and C isolated from *Acer okamotoanum*, *Arch. Pharm. Res.* 30 (2007) 275–281.
- [53] X. Li, Y. Gao, F. Li, A. Liang, Z. Xu, Y. Bai, et al., Maclurin protects against hydroxyl radical-induced damages to mesenchymal stem cells: antioxidant evaluation and mechanistic insight, *Chem. Biol. Interact.* 219 (2014) 221–228.
- [54] CHEMBL239211 | C14H10O5—PubChem [Internet] http://pubchem.ncbi.nlm.nih.gov/compound/4_4_-Oxybis_benzoic_acid.
- [55] F. Epifano, S. Genovese, E. James Squires, M.A. Gray, Nelumal A, The active principle from *Ligularia nelumbifolia*, is a novel farnesoid X receptor agonist, *Bioorg. Med. Chem. Lett.* 22 (2012) 3130–3135.
- [56] I. Rodríguez-Ramiro, S. Ramos, L. Bravo, L. Goya, M.Á. Martín, Procyandin B2 induces Nrf2 translocation and glutathione S-transferase P1 expression via ERKs and p38-MAPK pathways and protect human colonic cells against oxidative stress, *Eur. J. Nutr.* 51 (2012) 881–892.
- [57] T. Esatbeyoglu, A.E. Wagner, V.B. Schini-Kerth, G. Rimbach, Betanin—a food colorant with biological activity, *Mol. Nutr. Food Res.* (2014) 36–47.
- [58] L. Li, X. Zhang, L. Cui, L. Wang, H. Liu, H. Ji, et al., Ursolic acid promotes the neuroprotection by activating Nrf2 pathway after cerebral ischemia in mice, *Brain Res* 1497 (2013) 32–39.
- [59] U. Boettler, K. Sommerfeld, N. Volz, G. Pahlke, N. Teller, V. Somoza, et al., Coffee constituents as modulators of Nrf2 nuclear translocation and ARE (EpRE)-dependent gene expression, *J. Nutr. Biochem.* 22 (2011) 426–440.
- [60] C.H. Park, J.S. Noh, J.H. Kim, T. Tanaka, Q. Zhao, K. Matsumoto, et al., Evaluation of morroniside, iridoid glycoside from corni fructus, on diabetes-induced alterations such as oxidative stress, inflammation, and apoptosis in the liver of type 2 diabetic db/db mice, *Biol. Pharm. Bull.* 34 (2011) 1559–1565.
- [61] S. Kim, K. Kim, J. Park, J. Kwak, Y. Shik, S. Lee, Geniposidic acid protects against D-galactosamine and lipopolysaccharide-induced hepatic failure in mice, *J. Ethnopharmacol.* 146 (2013) 271–277.
- [62] B.M. Choi, S.M. Kim, T.K. Park, G. Li, S.J. Hong, R. Park, et al., Piperine protects cisplatin-induced apoptosis via heme oxygenase-1 induction in auditory cells, *J. Nutr. Biochem.* 18 (2007) 615–622.
- [63] A.N. Kong, E. Owuor, R. Yu, V. Hebbar, C. Chen, R. Hu, et al., Induction of xenobiotic enzymes by the MAP kinase pathway and the antioxidant or electrophile response element (ARE/EpRE), *Drug Metab. Rev.* 33 (2001) 255–271.
- [64] X. Jin, Q. Liu, L. Jia, M. Li, X. Wang, Pinocembrin attenuates 6-OHDA-induced neuronal cell death through Nrf2/ARE pathway in SH-SY5Y cells, *Cell. Mol. Neurobiol.* 35 (2014).

- [65] D.S. Lee, B. Li, K.S. Kim, G.S. Jeong, E.C. Kim, Y.C. Kim, Butein protects human dental pulp cells from hydrogen peroxide-induced oxidative toxicity via Nrf2 pathway-dependent heme oxygenase-1 expressions, *Toxicol. Vitro* 27 (2013) 874–881.
- [66] C.S. Huang, C.K. Lii, A.H. Lin, Y.W. Yeh, H.T. Yao, C.C. Li, et al., Protection by chrysin, apigenin, and luteolin against oxidative stress is mediated by the Nrf2-dependent up-regulation of heme oxygenase 1 and glutamate cysteine ligase in rat primary hepatocytes, *Arch. Toxicol.* 87 (2013) 167–178.
- [67] S. Tanigawa, M. Fujii, D.X. Hou, Action of Nrf2 and Keap1 in ARE-mediated NQO1 expression by quercetin, *Free Radic. Biol. Med.* 42 (2007) 1690–1703.
- [68] B. Podder, H. Song, Y. Kim, Naringenin exerts cytoprotective effect against paraquat-induced toxicity in human bronchial epithelial BEAS-2B cells through Nrf2 activation, *J. Microbiol. Biotechnol.* 24 (2014) 605–613.
- [69] P.T. Thuong, Y.R. Pokharel, M.Y. Lee, S.K. Kim, K. Bae, N.D. Su, et al., Dual antioxidant effects of fraxetin isolated from *Fraxinus rhynchophylla*, *Biol. Pharm. Bull.* 32 (2009) 1527–1532.
- [70] W. Liao, Z. Ning, L. Chen, Q. Wei, E. Yuan, J. Yang, et al., Intracellular antioxidant detoxifying effects of diosmetin on 2,2-azobis(2-amidinopropane) dihydrochloride (AAPH)-induced oxidative stress through inhibition of reactive oxygen species generation, *J. Agric. Food Chem.* 62 (2014) 8648–8654.
- [71] K.S. Bhullar, H.P.V. Rupasinghe, Antioxidant and cytoprotective properties of partridgeberry polyphenols, *Food Chem.* 168 (2015) 595–605.
- [72] J. Barillari, D. Canistro, M. Paolini, F. Ferroni, G.F. Pedulli, R. Iori, et al., Direct antioxidant activity of purified glucoerucin, the dietary secondary metabolite contained in rocket (*Eruca sativa* Mill.) seeds and sprouts, *J. Agric. Food Chem.* 53 (2005) 2475–2482.
- [73] I.C. Lee, S.H. Kim, H.S. Baek, C. Moon, S.H. Kim, Y.B. Kim, et al., Protective effects of diallyl disulfide on carbon tetrachloride-induced hepatotoxicity through activation of Nrf2, *Environ. Toxicol.* (2013).
- [74] J.O.A. Narciso, R.O.D. de Araújo Soares, J.R. dos Santos Mallet, A.É. Guimarães, M.C. de Oliveira Chaves, J.M. Barbosa-Filho, Burchellin: study of bioactivity against *Aedes aegypti*, *Parasit. Vectors* 7 (2014) 172.
- [75] S.P. Singh, S. Kumari, R.P. Rastogi, K.L. Singh, R.P. Sinha, Mycosporine-like amino acids (MAAs): chemical structure, biosynthesis and significance as UV-absorbing/screening compounds, *Indian J. Exp. Biol.* 46 (2008) 7–17.
- [76] J.M. Batista, A.N.L. Batista, D. Rinaldo, W. Vilegas, D.L. Ambrósio, R.M. B. Cicarelli, et al., Absolute configuration and selective trypanocidal activity of gaudichaudianic acid enantiomers, *J. Nat. Prod.* 74 (2011) 1154–1160.
- [77] D.Y. Wang, S. Kumar, S.B. Hedges, Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi, *Proc. Biol. Sci.* 266 (1999) 163–171.
- [78] G. Giribet, C. Dunn, G. Edgecombe, G. Rouse, A modern look at the animal tree of life, *Zootaxa* 79 (2007) 61–79 <http://dx.doi.org/10.11211.9797>.
- [79] R.P. Oliveira, J.P. Abate, K. Dilks, J. Landis, J. Ashraf, C.T. Murphy, et al., Condition-adapted stress and longevity gene regulation by *Caenorhabditis elegans* SKN-1/Nrf, *Aging Cell* 8 (2009) 524–541.
- [80] K.P. Choe, A.J. Przybysz, K. Strange, The WD40 repeat protein WDR-23 functions with the CUL4/DDB1 ubiquitin ligase to regulate nuclear abundance and activity of SKN-1 in *Caenorhabditis elegans*, *Mol. Cell. Biol.* 29 (2009) 2704–2715.
- [81] S. Chowdhry, Y. Zhang, M. McMahon, C. Sutherland, Cuadrado, a and Hayes, J. D. Nrf2 is controlled by two distinct β -TrCP recognition motifs in its Neh6 domain, one of which can be modulated by GSK-3 activity, *Oncogene* 32 (2013) 3765–3781.
- [82] P. Rada, A.I. Rojo, N. Evrard-Todeschi, N.G. Innamorato, A. Cotte, T. Jaworski, et al., Structural and functional characterization of Nrf2 degradation by the glycogen synthase kinase 3 β -TrCP Axis, *Mol. Cell. Biol.* 32 (2012) 3486–3499.
- [83] G. Wu, G. Xu, B.A. Schulman, P.D. Jeffrey, J.W. Harper, N.P. Pavletich, Structure of a β -TrCP1-Skp1- β -catenin complex: destruction motif binding and lysine specificity of the SCF(β -TrCP1) ubiquitin ligase, *Mol. Cell* 11 (2003) 1445–1456.
- [84] J.A. Powell-Coffman, Hypoxia signaling and resistance in *C. elegans*, *Trends Endocrinol. Metab.* 21 (2010) 435–440.
- [85] G. Rimbach, Betanin—a food colorant with biological activity, *Mol. Nutr. Food Res.* 59 (2015) 36–47.
- [86] Y. Sato, S. Itagaki, T. Kurokawa, J. Ogura, M. Kobayashi, T. Hirano, et al., In vitro and in vivo antioxidant properties of chlorogenic acid and caffeic acid, *Int. J. Pharm.* 403 (2011) 136–138.
- [87] M. Yoshiki, K. Tsuge, Y. Tsuruta, T. Yoshimura, K. Koganemaru, T. Sumi, et al., Production of new antioxidant compound from mycosporine-like amino acid, porphyra-334 by heat treatment, *Food Chem.* 113 (2009) 1127–1132.
- [88] W.C. Dunlap, Y. Yamamoto, Small-molecule antioxidants in marine organisms: antioxidant activity of mycosporine-glycine, *Comp. Biochem. Physiol.* 112B (1995) 105–114.