

SCIENTIFIC REPORTS



OPEN

AOD: the antioxidant protein database

Pengmian Feng¹, Hui Ding², Hao Lin² & Wei Chen³

An antioxidant is a molecule that can prevent free radicals from causing damages in organisms. The increasing studies on antioxidants calls for a specialized database that is not readily available yet. To this end, in the present study, the Antioxidant Database (AOD) was developed to help researchers understand and reveal the biological functions of antioxidant proteins. AOD is freely available at <http://lin.uestc.edu.cn/AODdatabase/index.aspx>. The current release of AOD consists of 710 antioxidant proteins. Information including taxonomy, source organism, subcellular location, gene ontology, catalytic activity and function of antioxidant proteins are all extracted from UniProtKB/Swiss-Prot and captured in AOD. In addition, two web-based tools for performing sequence similarity search and computationally identification of antioxidants were also integrated in AOD. We believe that AOD will greatly facilitate the researches on antioxidants.

An antioxidant, also known as the free radical scavenger, is a molecule that can neutralize free radicals and thus preventing them from causing damages¹. Recent studies have demonstrated that antioxidants play important roles in the management or prevention of cancers², coronary heart disease³, macular degeneration⁴, Alzheimer's disease⁵, arthritis-related conditions⁶, and longevity⁷. Therefore, antioxidants have attracted considerable attentions of scientists who focus on cancer prophylaxis and therapy and human health.

Our human body naturally produces antioxidants to counteract the damaging effects of free radicals⁸. Endogenous antioxidants are enzymes, such as superoxide dismutase⁹, catalase¹⁰, glutathione peroxidase¹¹ or nonenzymatic compounds including uric acid¹², bilirubin¹³, metallothioneins¹⁴. However, the amount of free radicals is often greater than that of the naturally occurring antioxidants in organisms. In order to balance the disequilibrium, it's necessary to obtain antioxidants from external sources.

In the past several decades, multiple types of exogenous antioxidant proteins have been detected, such as vitamin A, vitamin C, and vitamin E^{15,16}, and a huge amount of studies about antioxidants have also been reported. However, there isn't a database systematically collecting and compiling the information of antioxidants at present. Therefore, it is urgent need to develop a database, where the researchers could obtain the precise information of known antioxidant proteins.

The enormous amount of data on antioxidants had motivated us to develop a general database. Hence, in the present study, we established the antioxidant database (AOD) with an objective to provide useful insights to the study of antioxidants.

Results and Discussions

Data statistics. Among the 710 antioxidant proteins of AOD, 458 are from eukaryota, 221 from bacteria, 28 from archaea and 3 from virus (Fig. 1a). In AOD, 94% of the antioxidant proteins contain more than 100 amino acids with an average length of 227. The longest antioxidant protein (A11GV8) has 1,463 amino acids and the shortest one (P83233) merely 8 amino acids. The length distribution of antioxidant proteins in AOD is shown in Fig. 1b.

The ratio of the frequency of amino acids in AOD to the frequency of amino acids in UniProtKB/Swiss-Prot protein knowledgebase is shown in Fig. 1c. We found that antioxidant proteins are enriched in residues Cys, Gly, His and Val.

¹Hebei Province Key Laboratory of Occupational Health and Safety for Coal Industry, School of Public Health, North China University of Science and Technology, Tangshan, 063000, China. ²Key Laboratory for Neuro-Information of Ministry of Education, School of Life Science and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, 610054, China. ³Department of Physics, School of Sciences, and Center for Genomics and Computational Biology, North China University of Science and Technology, Tangshan, Tangshan, 063000, China. Correspondence and requests for materials should be addressed to H.L. (email: hlin@uestc.edu.cn) or W.C. (email: greatchen@ncst.edu.cn)

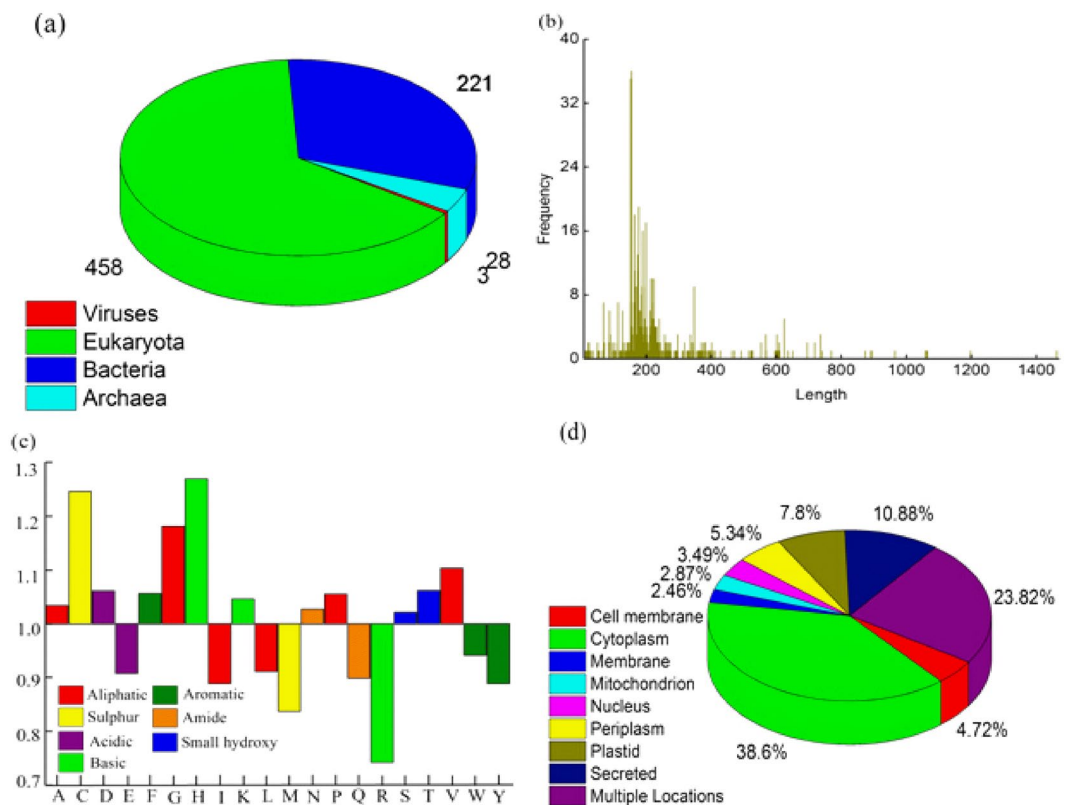


Figure 1. The statistical analysis of AOD. **(a)** Organism distribution of antioxidant proteins in AOD; **(b)** Length distribution of antioxidant proteins in AOD; **(c)** Relative amino acid composition of antioxidant proteins in AOD as compared with Swiss-Prot composition; **(d)** Subcellular location distribution of antioxidant proteins in AOD.

The subcellular location of antioxidant proteins is closely correlated with its biological activities¹⁷. Therefore, the information about subcellular location of antioxidant proteins is also provided in AOD. Four hundred and eighty seven antioxidant proteins in AOD have annotated subcellular location information, and 76% of them only reside in a single subcellular localization and the other 24% in multiple subcellular locations (Fig. 1d).

Quality assessment. In order to evaluate the quality of AOD, the 710 antioxidant proteins in the database were manually checked. It was found that there are 609 antioxidant proteins that have been reported in previous studies as indicated in the UniProtKB/Swiss-Prot. The Uniprot IDs for these 609 proteins were listed in Supplementary Table S1. As indicated in UniProtKB/Swiss-Prot, the remaining 101 antioxidant proteins were evaluated by experimental methods at transcript level, which supports the existence of the protein.

Browse. By clicking the ‘Browse’ button, the antioxidant proteins in AOD database will be shown on the computer screen and can be displayed page by page. The antioxidant proteins could be selected by clicking the buttons on their left. The information of protein name, taxonomy, organism, subcellular location, gene ontology, catalytic activity, function, amino acid sequence and links to external database such as Swiss-Model Repository, Uniprot, Gene Ontology and NCBI for each entry can be viewed by clicking on the sequence ID or by clicking the ‘Show’ button on the top right corner. The selected antioxidant proteins could also be downloaded to local machine by clicking the ‘Download’ button and saved in fasta format.

Search. The AOD can be searched in multiple ways. Users can basically search antioxidant proteins by limiting the searching filed to UniProt ID, Taxonomy, Organism, Subcellular locations or Protein names and entering the related query keyword. To perform conditional search, user can add (or remove) the searching filed using the Add (or Remove) button and then join the multiple query keywords by logical operators like AND/OR.

Tools. AOD integrates two web-based tools for performing further analyses, i.e. sequence similarity search and identification of antioxidants. We have integrated BLAST¹⁸ in AOD that allows users to perform the BLAST search against sequences deposited in AOD, which will facilitate finding sequences in the database that have high sequence similarity with the query sequence.

To the best of our knowledge, **AodPred**¹⁹ is the smartest computational tool for identifying antioxidants at present. For the convenience of experimental scientists, the **AodPred** predictor was also integrated in AOD. Users can employ it to predict whether a query protein sequence is antioxidant or not. After inputting the query

sequences in fasta format and clicking the ‘Submit’ button, the probability of the predictions being antioxidant or non-antioxidant will be shown in a new page.

Conclusions

In the present study, a comprehensive database called AOD is built with the aim to provide useful insights to the study of antioxidant proteins and to help researchers understand the role of different properties of antioxidant in their antioxidative activities. AOD is the first database providing information on antioxidants from multiple perspectives. We hope the AOD will better serve the research on antioxidant proteins. In order to make an encyclopedia-like database for antioxidants, we will continue to accumulate the information of new antioxidant proteins and add them into AOD.

Methods and Materials

Data collection. Sequences of antioxidant proteins together with their information including taxonomy, source organism, subcellular location, gene ontology, catalytic activity and function were obtained from the UniProtKB/Swiss-Prot database (release 2016_11)^{20,21} by searching the keyword “antioxidant”. In order to obtain the quality data, the following steps were performed: (i) only proteins with the experimentally confirmed antioxidative activities were included; (ii) proteins including illegal letters, i.e., “B”, “X” or “Z”, were excluded. This finally yields 710 proteins that have experimentally proven or confirmed antioxidative activity to be included into AOD.

Database structure and interface. As a user-friendly database, AOD is freely available at <http://lin.uestc.edu.cn/AODdatabase/index.aspx>. The database main page contains the following interfaces: Home, Browse, Search, Tools, Statistics, Links and Contact. The information related to protein name, taxonomy, origin of organism, subcellular location, sequence length, gene ontology, catalytic activity, protein sequence, function and links to external database such as Swiss-Model Repository, Gene Ontology and NCBI are all provided in AOD.

References

- Lobo, V., Patil, A., Phatak, A. & Chandra, N. Free radicals, antioxidants and functional foods: Impact on human health. *Pharmacognosy reviews* **4**, 118–126, doi:10.4103/0973-7847.70902 (2010).
- Dreher, D. & Junod, A. F. Role of oxygen free radicals in cancer development. *European journal of cancer* **32A**, 30–38 (1996).
- Maxwell, S. R. Coronary artery disease-free radical damage, antioxidant protection and the role of homocysteine. *Basic research in cardiology* **95**(Suppl 1), 165–71 (2000).
- Rojas-Fernandez, C. H. & Tyber, K. Benefits, Potential Harms, and Optimal Use of Nutritional Supplementation for Preventing Progression of Age-Related Macular Degeneration. *The Annals of pharmacotherapy* **51**, 264–270, doi:10.1177/1060028016680643 (2017).
- Galasko, D. R. *et al.* Antioxidants for Alzheimer disease: a randomized clinical trial with cerebrospinal fluid biomarker measures. *Archives of neurology* **69**, 836–841, doi:10.1001/archneurol.2012.85 (2012).
- van Vugt, R. M., Rijken, P. J., Rietveld, A. G., van Vugt, A. C. & Dijkmans, B. A. Antioxidant intervention in rheumatoid arthritis: results of an open pilot study. *Clinical rheumatology* **27**, 771–775, doi:10.1007/s10067-008-0848-6 (2008).
- Mecocci, P. *et al.* Plasma antioxidants and longevity: a study on healthy centenarians. *Free radical biology & medicine* **28**, 1243–1248 (2000).
- Pham-Huy, L. A., He, H. & Pham-Huy, C. Free radicals, antioxidants in disease and health. *International journal of biomedical science: IJBS* **4**, 89–96 (2008).
- Mruk, D. D., Silvestrini, B., Mo, M. Y. & Cheng, C. Y. Antioxidant superoxide dismutase - a review: its function, regulation in the testis, and role in male fertility. *Contraception* **65**, 305–311 (2002).
- Kang, M. Y. *et al.* The critical role of catalase in prooxidant and antioxidant function of p53. *Cell death and differentiation* **20**, 117–129, doi:10.1038/cdd.2012.102 (2013).
- Espinoza, S. E. *et al.* Glutathione peroxidase enzyme activity in aging. *The journals of gerontology. Series A, Biological sciences and medical sciences* **63**, 505–509 (2008).
- Ames, B. N., Cathcart, R., Schwiers, E. & Hochstein, P. Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 6858–6862 (1981).
- Sedlak, T. W. *et al.* Bilirubin and glutathione have complementary antioxidant and cytoprotective roles. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5171–5176, doi:10.1073/pnas.0813132106 (2009).
- Viarengo, A., Burlando, B., Ceratto, N. & Panfoli, I. Antioxidant role of metallothioneins: a comparative overview. *Cellular and molecular biology* **46**, 407–417 (2000).
- Pisoschi, A. M., Negulescu, G. P. & Methods for Total Antioxidant Activity Determination: A Review. *Biochemistry and Analytical Biochemistry* **1** (2011).
- Stratil, P., Klejdus, B. & Kuban, V. Determination of total content of phenolic compounds and their antioxidant activity in vegetables—evaluation of spectrophotometric methods. *Journal of agricultural and food chemistry* **54**, 607–616, doi:10.1021/jf052334j (2006).
- Green, T. R., Fellman, J. H., Eicher, A. L. & Pratt, K. L. Antioxidant role and subcellular location of hypotaurine and taurine in human neutrophils. *Biochimica et biophysica acta* **1073**, 91–97 (1991).
- Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389–3402 (1997).
- Feng, P., Chen, W. & Lin, H. Identifying Antioxidant Proteins by Using Optimal Dipeptide Compositions. *Interdisciplinary sciences, computational life sciences* **8**, 186–191, doi:10.1007/s12539-015-0124-9 (2016).
- Pundir, S., Martin, M. J. & O’Donovan, C. UniProt Protein Knowledgebase. *Methods in molecular biology* **1558**, 41–55, doi:10.1007/978-1-4939-6783-4_2 (2017).
- The UniProt, C. UniProt: the universal protein knowledgebase. *Nucleic acids research* **45**, D158–D169, doi:10.1093/nar/gkw1099 (2017).

Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive comments. This work was supported by the Program for the Top Young Innovative Talents of Higher Learning Institutions of Hebei Province (No. BJ2014028), the Outstanding Youth Foundation of North China University of Science and Technology (No. JP201502) and the Natural Science Foundation for Distinguished Young Scholar of Hebei Province (No. C2017209244).

Author Contributions

W.C. and H.L. conceived and designed the experiments; P.M. and H.D. analyzed the data and built the web server; W.C. and H.L. wrote the paper.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-08115-6](https://doi.org/10.1038/s41598-017-08115-6)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017