

RESEARCH

Open Access



Evolutionary aspects of the Viridiplantae nitroreductases

Siarhei A. Dabravolski

Abstract

Background: Nitroreductases are a family of evolutionarily related proteins catalyzing the reduction of nitro-substituted compounds. Nitroreductases are widespread enzymes, but nearly all modern research and practical application have been concentrated on the bacterial proteins, mainly nitroreductases of *Escherichia coli*. The main aim of this study is to describe the phylogenetic distribution of the nitroreductases in the photosynthetic eukaryotes (Viridiplantae) to highlight their structural similarity and areas for future research and application.

Results: This study suggests that homologs of nitroreductase proteins are widely presented also in Viridiplantae. Maximum likelihood phylogenetic tree reconstruction method and comparison of the structural models suggest close evolutionary relation between cyanobacterial and Viridiplantae nitroreductases.

Conclusions: This study provides the first attempt to understand the evolution of nitroreductase protein family in Viridiplantae. Our phylogeny estimation and preservation of the chloroplasts/mitochondrial localization indicate the evolutionary origin of the plant nitroreductases from the cyanobacterial endosymbiont. A defined high level of the similarity on the structural level suggests conservancy also for the functions. Directions for the future research and industrial application of the Viridiplantae nitroreductases are discussed.

Keywords: Nitroreductase, Viridiplantae, Cyanobacteria, Phylogeny, Evolution

Background

Nitroreductases are a family of closely related proteins that catalyze the reduction of nitro-substituted compounds, using FMN (Flavin mononucleotide) or FAD (Flavin adenine dinucleotide) as a cofactor and NADH (Nicotinamide adenine dinucleotide) or NADPH (Nicotinamide adenine dinucleotide phosphate) as a reducing agent. Nitroreductases are ancient enzymes, with approximate evolutionary age ~ 2.5 billion years and represented by more than 26,000 known sequences [19]. In addition to the nitroreduction reaction, nitroreductases are known to catalyze a wide range of other reactions (dehalogenation, dehydrogenation, flavin fragmentation) and apply a wide range of substrates (metal ions, quinone, flavin, nitroaromatic, and enone compounds) (reviewed in [61]). While nitroreductase enzymes are

widespread and their reactions diversity is well-characterized, nearly all modern research for has been concentrated on the proteins of *Escherichia coli*. Thus, elucidating the evolutionary relations of the plant nitroreductases could facilitate their further research and industrial application.

Nitroreductases have high potential in their utility in activating prodrugs in directed anticancer therapies (reviewed in [61]). Bacterial nitroreductase NfsB (*Escherichia coli*) was applied with positive clinical outcomes for the treatment of prostate cancer and brain tumors [45]. Recently, *Mycobacterium smegmatis* nitroreductase NfnB was used as a pharmaceutical and chemicals synthesis agent to obtain new compound BTZ043 for the treatment of tuberculosis [37]. To bioremediate and degrade the world-wide use pollutant polychlorinated biphenyl was successfully created transgenic tobacco plants, expressing nitroreductase *bphC* gene from *Pandoraea pnomenusa* [44]. Also, nitroreductase NfsA from

Correspondence: sergedobrowski@gmail.com
Department of Clinical Diagnostics, Vitebsk State Academy of Veterinary
Medicine [UO VGAVM], 7/11 Dovatora St., 210026 Vitebsk, Belarus

Escherichia coli was successfully used in the biocatalysis of several nitroaromatic compounds and quinones [56].

The current classification includes two classes of nitroreductases: type I (oxygen-insensitive) catalyze the reduction of organic nitro compounds using a two-electron transfer mechanism to primary amines [31] and type II (oxygen-sensitive) catalyzes a one-electron reduction of the nitro group to produce nitro anion radicals that may react with oxygen, form superoxide and cause oxidative stress [48]. In the yeast, *Saccharomyces cerevisiae*, 2 genes, *frm2* (YCL026c-A) (fatty acid repression mutant) and *hbn1* (YCL026c-B) (homologous to bacterial nitroreductases), encoding putative nitroreductase-like proteins were identified by in silico analysis [15]. The biological functions of the yeast nitroreductase family of proteins are not well studied; however, their possible involvement in oxidative stress responses has been suggested [3]. Experimental data on Frm2 protein indicate that Frm2 may be involved in the lipid signalling pathway and cellular homeostasis [40]. Also, solved crystal structure [52] supports this finding and provides insights into the molecular mechanism of the yeast Frm2 activity.

Human DEHAL1 (iodotyrosine dehalogenase 1) is a well-characterized member of the nitroreductase family responsible for iodide recycle [26] and thyroid hormone synthesis [8]. It was shown that nitroreductase and dehalogenase activities are closely related to the sequence level [43].

The recent advantage in the sequencing technologies and genes annotation with an automatic pipeline allows identifying many genes as “nitroreductase family member” in Viridiplantae. Despite the undoubted importance of the nitroreductases, their characterization in Viridiplantae is missing. Thus, although a recent study has provided a deep insight into the understanding of the evolution of nitroreductases [1], the evolutionary history of the green lineage nitroreductases has not been addressed as broadly as in other kingdoms.

In the present work, all currently available genomic resources were used to explore the diversity and the phylogenetic distribution of the nitroreductase domain-containing proteins in Viridiplantae. This study represents the first step toward understanding the evolution of the nitroreductase proteins in the green lineage. Altogether, results of this study could facilitate further research and industrial application of the Viridiplantae nitroreductases.

Methods

Identification of the nitroreductases in the Viridiplantae clade

Nitroreductases were identified with keyword search and following BLAST (Basic local alignment search tool) [51] searches in NCBI (National Center for Biotechnology

Information), InterPro 77 [42], Pfam 32.0 [19], and Phytosome 12.1 [21] databases. The consensus sequence of the nitroreductase domain (PF00881/IPR029479) was used in on-line BLASTP (Basic Local Alignment Search Tool Protein) searches. All partial and fragmented sequences were eliminated. Presence of the nitroreductase domain was checked with CD-search (NCBI) [38] and MOTIF search (KEGG 93) [30] tools with E-value (≤ 0.001). Domains, fused to the nitroreductase domain, were verified with the same tools and threshold.

Multiple sequence alignments and phylogenetic analysis

Nitroreductase domain sequences were extracted from Pfam database [19] (for proteins with 2 nitroreductase domains, the sequence of N-terminal one was used) and multiple sequence alignments were performed with MUSCLE [18]. The test of substitution models and phylogenetic analysis were carried out using the MEGA X software [32]. For maximum likelihood tree [60], the LG substitution model [34] was selected assuming an estimated proportion of invariant sites and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data. Reliability for the internal branch was assessed using the bootstrapping method (1000 bootstrap replicates). The same settings with the JTT substitution model [28] were used for reconstruction with the Neighbor-Joining [50] method.

Localization prediction and structure modeling

Subcellular localization was predicted with TargetP online tool [20], “Plant Organism group” settings were used. Model of *Arabidopsis* nitroreductase (At1G02020) was built with SWISS-MODEL [59] and matched with *Anabaena variabilis* nitroreductase (PDB 3EO7 (<https://doi.org/10.2210/pdb3EO7/pdb>)) in Chimera software [49]. Quality of the created models was verified with QMEAN [5]. VAST+ was used to search for the structure similarity [36], iPBA webserver was used for the pdb structures alignment (https://www.dsimb.inserm.fr/dsimb_tools/ipba/index.php). The quality of the structures alignments was evaluated with RMSD and normalized score [55, 64].

Results

Exploring the distribution of nitroreductases in Viridiplantae

All currently available Viridiplantae species genomes were checked, and 97 proteins containing nitroreductase domain (Supplementary Table 1) were identified. Most nitroreductases are single-domain proteins, containing only nitroreductase domain, but there are some examples of nitroreductases with doubled nitroreductase domain. In 13 proteins, this additional C-terminal domain

Table 1 Structural alignment of the Arabidopsis O23673 protein (Locus At1g02020) with nitroreductases with defined structures from different species

| Organism | PDB ID | Normalized score | RMSD | GDT TS |
|---|--------|------------------|------|--------|
| <i>Geobacter Metallireducens Gs-15</i> | 4DN2 | − 204.97 | 1.67 | 16.72 |
| <i>Clostridium difficile R20291</i> | 5J62 | − 277.48 | 2.75 | 2.14 |
| <i>Chlorobium tepidum</i> | 2R01 | − 231.45 | 1.71 | 15.53 |
| <i>Ralstonia eutropha jmp134</i> | 3HJ9 | − 177.6 | 1.82 | 20.26 |
| <i>Bacteroides fragilis NCTC 9343</i> | 3EK3 | − 313.01 | 2.67 | 2.73 |
| <i>Vibrio parahaemolyticus RIMD 2210633</i> | 5UU6 | − 246.08 | 1.77 | 15.48 |
| <i>Saccharomyces Cerevisiae</i> | 4URP | − 230.67 | 2.4 | 13.71 |
| <i>Anabaena variabilis atcc 29413</i> | 3EO7 | 130.13 | 0.73 | 73.69 |

was significantly above threshold (designated as $\times 2$ in Supplementary Table 1), in all other cases, the domain was partial and below a defined threshold. Two proteins (D7KP50 and A0A2P5WXT2) have N-terminal C2H2-type zinc finger (in D7KP50) and Myb/SANT-like DNA-binding (in A0A2P5WXT2) domains (from *Arabidopsis lyrata* L. and *Gossypium barbadense* L., respectively) that may suggest their additional function as transcription factors.

Several recent studies have provided important advances in our understanding of the structure and distribution of nitroreductases in many species: Bacteria, Fungi (Oliveira 2007); yeast (*Saccharomyces cerevisiae*) (Song 2015); *Clostridium difficile* (Wang [57]); nematode *Caenorhabditis elegans* [14]; mouse (*Mus musculus*) [54]; human [25]. However, phylogenetic analysis of nitroreductases in the whole green lineage (Viridiplantae) is missing.

Genome searches revealed that nitroreductases are present in the Chlorophyta (green algae) (11 proteins), mosses (1 protein in *Physcomitrella patens* Hedw [33]., liverwort (1 protein in *Marchantia polymorpha* L. [10],

clubmosses (2 proteins in *Selaginella moellendorffii* [4]. No nitroreductases were found in Red Algae. Also, nitroreductase proteins were found in Acrogymnospermae: *Wollemia nobilis* [47, 62] and *Araucaria cunninghamii* Mudie. (one protein in each). Ancient flowering plant *Amborella trichopoda* Baill [2]., *Cinnamomum micranthum* (Hayata) [13], and *Macleaya cordata* Willd [35]. were containing one nitroreductase protein each, but three proteins were found in *Nelumbo nucifera Gaertn* [41]. Monocotyledons and eudicotyledons were represented with a higher number of proteins, 10 and 64 respectively (Supplementary Table 1).

To gain insight into the phylogenetic relationships between nitroreductases from green algae (Chlorophyta) and land plants (Streptophyta), the robust phylogenetic tree after multiple alignments of 100 predicted sequences (Supplementary Figure 1 and 2) was generated. Nitroreductases from Cyanobacteria (*Anabaena cylindrical* and *Nostoc punctiforme*) and Chloroflexi (*Chloroflexus islandicus*) were added as green photosynthetic bacteria outgroup. As expected, nitroreductases from Streptophyta and Green algae were clustered in

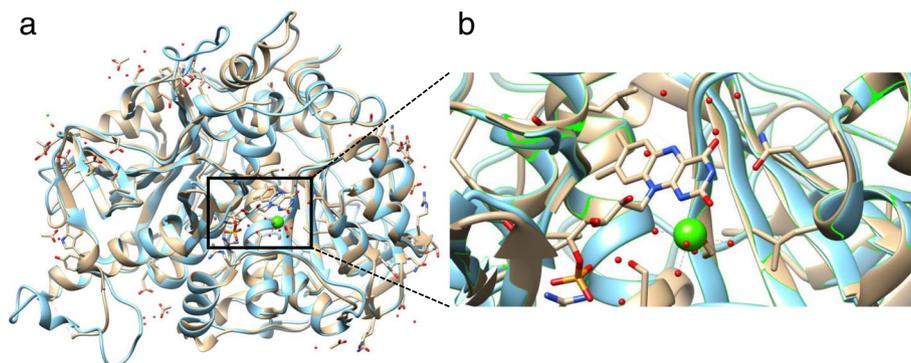


Fig. 1 **a** Matched crystal structures of putative nitroreductase from *Arabidopsis thaliana* (blue) and *Anabaena variabilis* (PDB 3EO7) (pale pink). **b** Closer look at the FMN (in the centre) binding site and Ca^{2+} ion (green ball)

separated branches, with closer relation of Bacterial nitroreductases to Green algae. Also, high homology on the level of the family that decreasing dramatically in higher taxa could be noticed. Thus, this represents the first description of nitroreductase homologs in Green Algae and Streptophyta.

Structural models comparison

Structural alignment of the *Arabidopsis thaliana* chloroplasts/mitochondria-localized nitroreductase (O23673) with different known nitroreductases (Table 1) confirms results obtained with the phylogenetic tree. In comparison to different bacteria and yeast, cyanobacteria *Anabaena variabilis* shown the highest structural similarity to the nitroreductase from *Arabidopsis thaliana* (Fig. 1a). It is important to notice, that amino acids, required for the co-factor (FMN) binding are conserved in cyanobacteria and *Arabidopsis* (Fig. 1b): 98R, 101P, 102S, 188D, 191H, and 334R. These features suggest that plant nitroreductases are most likely to perform functions, similar to bacterial. The main difference between plant and bacterial nitroreductases is the presence of the N-terminal peptide (1–40 in O23673), required for the protein translocation to the chloroplasts and/or mitochondria.

Based on the obtained phylogenetic results and predominant localization of the plant nitroreductases in chloroplasts, we could suggest the general line of evolution from cyanobacteria via endosymbiotic event to the modern chloroplasts. Overall, our assumption is well-supported by the modern theory of the chloroplasts and mitochondria origin from bacterial ancestors [39].

Discussion

In this study, for the first time, predicted nitroreductases in unexplored eukaryotic Viridiplantae supergroup were described. As it was shown in previous studies [15, 22] nitroreductases have very low sequence identity/similarity. This fact may explain why their presence in Viridiplantae was overlooked. Based on the ancient nature of the nitroreductases (approximate evolutionary age ~ 2.5 billion years [58]), wide representation among different taxa, we assume that nitroreductases are omnipresent enzymes and also presented in Viridiplantae. Application of different BLAST search strategies allows to identify nitroreductases literally in all Viridiplantae species but mostly with partial nitroreductase domain or below a threshold value.

Up to date, not much is known about localization of the nitroreductase activity in mammalian cells. The recently developed fluorescent sensor allows to image sub-mitochondrial localization of nitroreductase activity in live HEK 293 cells (human embryonic kidney cells) [53]. Similarly, bacterial nitroreductases, transformed into

plants, have higher activity in case of chloroplast and mitochondrial localization [65]. Most probably such localization is connected to the role in oxidative stress response and regulation of antioxidant enzymes as it was shown in yeast [3, 16] and detoxification of the photosynthesis by-products [7]. Predicted localization of the defined plant nitroreductases corresponds with this assumption, and majority of proteins were predicted to have chloroplast and/or mitochondrial localization (Supplementary Table 1). The only exception with no organelle localization is the secreted nitroreductase (A0A0D2VDD8) from cotton *Gossypium raimondii* Ulbr.

Most probably, plant nitroreductases are participating in the oxidative stress, pollutant, and herbicide responses [7, 65]. Also, some connections to circadian rhythms or the efficiency of the photosynthetic machine are possible. In particular, by-products of many pollutant and herbicide are known to degrade in the mitochondria and cause the production of the highly toxic reactive superoxide. The reduction of the superoxide is catalyzed by several families of closely related reductases localized in the mitochondria and chloroplasts, like, for example, monodehydroascorbate reductases [27] and quinone oxidoreductase [6]. Most probably, plant nitroreductases are also participating in the superoxide reduction. Also, it is known that some pollutants could damage photosynthetic apparatus and decrease the content of chlorophyll, but these negative effects are neutralized by the overexpression of transgenic bacterial nitroreductase [23], thus, suggesting photoprotective role.

The main focus of modern nitroreductases research is oriented on substrate recognition specificity, kinetic parameters related to prodrug activation or antibiotic resistance, but missing detailed characterization of regulatory mechanisms (reviewed by [61]). Based on available data, bacterial nitroreductases are induced by oxidative stress or decreases intracellular NAD(P)H to NAD(P)⁺ ratio [46]. Expression of the yeast nitroreductases is constitutive and does not depend on the cell physiological status [16]. Without any experimental data available, it is hard to predict regulatory mechanisms for the plant nitroreductases expression and functioning.

Plant nitroreductases have high potential in industrial application and biotechnology. It is known that substrate specificity of the nitroreductase could be changed just by the replacement of a single amino acid [43]. Thus, it is possible to adjust co-factor binding site, size, and shape of the substrate-binding pocket and create an enzyme suitable for degradation of the nearly any compound [17]. Bacterial enzymes are not much suitable for such purposes, because they lack membrane anchor and their localization is not

specific. Partially, this problem could be solved by transplastomic transformation, providing production of the high amount of functional enzyme [65]. From the other side, this method has several limitations, mainly: (1) many pollutants are absorbed by roots, where chloroplastic enzymes are not presented, or their activity is very low [11]; (2) plastid genes are greatly downregulated in fruits, where pollutants are often concentrated [29]; in general, plastid transformation is well-established only in the limited number of species, but agricultural and industrial plants species are rather recalcitrant [9]. Further research in this area would allow engineering plant species resistant to herbicides and with target phytoremediation properties.

Interestingly, that comparison of homologs on the level of protein structures provided a possible evolutionary relation of plant and bacterial nitroreductases (Table 1). Not surprisingly, the closest bacterial homolog of nitroreductase for *Arabidopsis thaliana* (L.) Heynh. (O23673) was nitroreductases from *Anabaena variabilis* (3eo7) (Fig. 1). Most probably that Viridiplantae has obtained nitroreductase via cyanobacterial endosymbiont [39].

Finally, this study reports the first description of the nitroreductases in the Viridiplantae supergroup. The low level of similarity between identified nitroreductases in Viridiplantae species complicated the phylogenomic analysis, and it was possible to make only a general overview of the evolutionary relationships of nitroreductases in this supergroup. Nitroreductase proteins have been thought to be absent from photosynthetic eukaryotes although this conclusion was made in rather old studies [12, 24] without application of modern similarity search algorithms and when a small number of Viridiplantae genomes was available.

Based on obtained results, some research directions for future investigation could be suggested: (1) proper classification of the Viridiplantae nitroreductase proteins based on their biochemical features (type I or II); (2) functional characterization of the newly defined proteins; (3) how localization of the nitroreductases (chloroplast, mitochondrial, or other) is related to the recognized substrates. Current attempts to create transgenic plants suitable for phytoremediation are based on the application of bacterial nitroreductases [7, 63, 65], although plant nitroreductases could provide better results. In addition to the application in the phytoremediation, transgenic plants, overexpressing nitroreductases could provide significant improvement during stress adaptation and disease resistance.

Conclusions

The present study reports, for the first time, evolutionary relation between previously overlooked nitroreductases from Viridiplantae including Chlorophyta, Bryophyta,

Marchantiophyta, Lycopodiopsida, and Spermatophyta. Results of the phylogenetic tree reconstruction and structural models' comparison suggest that green algae and cyanobacteria are the closest relatives for the modern plant nitroreductases. Conserved active sites, required for the co-factor binding, and chloroplastic/mitochondrial localizations imply primary physiological function in the oxidative stress response. In total, results of this study provide the first theoretical background for the future research of the Viridiplantae-delivered nitroreductases and discuss prospective areas for their practical application.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s43141-020-00073-3>.

Additional file 1: Supplementary Table 1. Accessions (UniProt) list of bacterial, algal, and plant proteins, identified in this study. Sequences were used for phylogenetic tree reconstruction (Supplementary Figure 1 and 2). Taxonomy was simplified. Domains are named according to the Pfam database [21]. Presence of both N- and C-terminal Nitroreductase domains are depicted as "x2 domain". Predicted subcellular localization (TargetP): C – chloroplast, M – mitochondria, S – secreted, "-" – not defined.

Additional file 2: Supplementary Figure 1. Phylogeny estimation of identified nitroreductase domain-containing proteins. The Neighbor-Joining method and JTT model were used; 1000 bootstrap replicates. Only branches with bootstrap value >50 are shown.

Additional file 3: Supplementary Figure 2. Phylogeny estimation of identified nitroreductase domain-containing proteins. The Maximum Likelihood method and LG model were used; 1000 bootstrap replicates. Only branches with bootstrap value >50 are shown.

Abbreviations

FMN: Flavin mononucleotide; FAD: Flavin adenine dinucleotide; NADH: Nicotinamide adenine dinucleotide; NADPH: Nicotinamide adenine dinucleotide phosphate; BLAST: Basic local alignment search tool; NCBI: National Center for Biotechnology Information; JTT: Jones-Taylor-Thornton

Acknowledgements

Not applicable

Author's contributions

SAD has designed and performed analyses, drafted, and wrote the manuscript. The author read and approved the manuscript.

Funding

This work was not supported by any funding organization.

Availability of data and materials

Available upon request

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

None

Received: 10 June 2020 Accepted: 14 September 2020

Published online: 06 October 2020

References

- Akiva E, Copp JN, Tokuriki N, Babbitt PC (2017) Evolutionary and molecular foundations of multiple contemporary functions of the nitroreductase superfamily. *Proc Natl Acad Sci USA* 114:E9549–E9558. <https://doi.org/10.1073/pnas.1706849114>
- Amborella Genome Project, Albert VA, Barbazuk WB et al (2013) The Amborella genome and the evolution of flowering plants. *Science* 342:1241089–1241089. <https://doi.org/10.1126/science.1241089>
- Bang SY, Kim JH, Lee PY et al (2012) Confirmation of Frm2 as a novel nitroreductase in *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* 423:638–641. <https://doi.org/10.1016/j.bbrc.2012.05.156>
- Banks JA, Nishiyama T, Hasebe M et al (2011) The selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332:960–963. <https://doi.org/10.1126/science.1203810>
- Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27:343–350. <https://doi.org/10.1093/bioinformatics/btq662>
- Biniek C, Heyno E, Kruk J et al (2017) Role of the NAD(P)H quinone oxidoreductase NQR and the cytochrome b AIR12 in controlling superoxide generation at the plasma membrane. *Planta* 245:807–817. <https://doi.org/10.1007/s00425-016-2643-y>
- Bo Z, Hongjuan H, Xiaoyan F et al (2018) Degradation of trinitrotoluene by transgenic nitroreductase in *Arabidopsis* plants. *Plant Soil Environ* 64:379–385. <https://doi.org/10.17221/655/2017-PSE>
- Bobyk KD, Ballou DP, Rokita SE (2015) Rapid kinetics of dehalogenation promoted by iodotyrosine deiodinase from human thyroid. *Biochemistry* 54:4487–4494. <https://doi.org/10.1021/acs.biochem.5b00410>
- Bock R (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annu Rev Plant Biol* 66:211–241. <https://doi.org/10.1146/annurev-arplant-050213-040212>
- Bowman JL, Kohchi T, Yamato KT et al (2017) Insights into land plant evolution garnered from the marchantia polymorpha genome. *Cell* 171:287–304.e15. <https://doi.org/10.1016/j.cell.2017.09.030>
- Brentner LB, Mukherji ST, Walsh SA, Schnoor JL (2010) Localization of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) and 2,4,6-trinitrotoluene (TNT) in poplar and switchgrass plants using phosphor imager autoradiography. *Environ Pollut* 158:470–475. <https://doi.org/10.1016/j.envpoll.2009.08.022>
- Bryant C, DeLuca M (1991) Purification and characterization of an oxygen-insensitive NAD(P)H nitroreductase from *Enterobacter cloacae*. *J Biol Chem* 266:4119–4125
- Chaw S-M, Liu Y-C, Wu Y-W et al (2019) Stout camphor tree genome fills gaps in understanding of flowering plant genome evolution. *Nat Plants* 5:63–73. <https://doi.org/10.1038/s41477-018-0337-0>
- de la Cruz IP, Ma L, Horvitz HR (2014) The *Caenorhabditis elegans* iodotyrosine deiodinase ortholog SUP-18 functions through a conserved channel SC-Box to regulate the muscle two-pore domain potassium channel SUP-9. *PLoS Genet* 10:e1004175. <https://doi.org/10.1371/journal.pgen.1004175>
- De Oliveira IM, Henriques JAP, Bonatto D (2007) In silico identification of a new group of specific bacterial and fungal nitroreductases-like proteins. *Biochemical and Biophysical Research Communications* 355(4):919–925. <https://doi.org/10.1016/j.bbrc.2007.02.049>
- de Oliveira IM, Zanotto-Filho A, Moreira JCF et al (2009) The role of two putative nitroreductases, Frm2p and Hbn1p, in the oxidative stress response in *Saccharomyces cerevisiae*. *Yeast*. <https://doi.org/10.1002/yea.1734>
- Durchschein K, Hall M, Faber K (2013) Unusual reactions mediated by FMN-dependent ene- and nitro-reductases. *Green Chem* 15:1764. <https://doi.org/10.1039/c3gc40588e>
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. <https://doi.org/10.1186/1471-2105-5-113>
- El-Gebali S, Mistry J, Bateman A et al (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* 47:D427–D432. <https://doi.org/10.1093/nar/gky995>
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300:1005–1016. <https://doi.org/10.1006/jmbi.2000.3903>
- Goodstein DM, Shu S, Howson R et al (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40:D1178–D1186. <https://doi.org/10.1093/nar/gkr944>
- Guillén H, Curiel JA, Landete JM et al (2009) Characterization of a nitroreductase with selective nitroreduction properties in the food and intestinal lactic acid bacterium *Lactobacillus plantarum* WCFS1. *J Agric Food Chem* 57:10457–10465. <https://doi.org/10.1021/jf9024135>
- Hannink N, Rosser SJ, French CE et al (2001) Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. *Nat Biotechnol* 19:1168–1172. <https://doi.org/10.1038/nbt1201-1168>
- Hecht HJ, Erdmann H, Park HJ et al (1995) Crystal structure of NADH oxidase from *Thermus thermophilus*. *Nat Struct Mol Biol* 2:1109–1114. <https://doi.org/10.1038/nsb1295-1109>
- Hu J, Chuenchor W, Rokita SE (2015) A switch between one- and two-electron chemistry of the human flavoprotein iodotyrosine deiodinase is controlled by substrate. *J Biol Chem* 290:590–600. <https://doi.org/10.1074/jbc.M114.605964>
- Hu J, Su Q, Schlessman JL, Rokita SE (2019) Redox control of iodotyrosine deiodinase: redox control of iodotyrosine deiodinase. *Protein Sci* 28:68–78. <https://doi.org/10.1002/pro.3479>
- Johnston EJ, Rylott EL, Beynon E et al (2015) Monodehydroascorbate reductase mediates TNT toxicity in plants. *Science* 349:1072–1075. <https://doi.org/10.1126/science.aab3472>
- Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275–282
- Kahlau S, Bock R (2008) Plastid transcriptomics and translationalomics of tomato fruit development and chloroplast-to-chromoplast differentiation: chromoplast gene expression largely serves the production of a single protein. *Plant Cell* 20:856–874. <https://doi.org/10.1105/tpc.107.055202>
- Kanehisa M, Sato Y, Kawashima M et al (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44:D457–D462. <https://doi.org/10.1093/nar/gkv1070>
- Kobori T, Sasaki H, Lee WC et al (2001) Structure and site-directed mutagenesis of a flavoprotein from *Escherichia coli* that reduces nitrocompounds: alteration of pyridine nucleotide binding by a single amino acid substitution. *J Biol Chem* 276:2816–2823. <https://doi.org/10.1074/jbc.M002617200>
- Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lang D, Ullrich KK, Murat F et al (2018) The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J* 93:515–533. <https://doi.org/10.1111/tpj.13801>
- Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* 25:1307–1320. <https://doi.org/10.1093/molbev/msn067>
- Liu X, Liu Y, Huang P et al (2017) The genome of medicinal plant *Macleaya cordata* provides new insights into benzylisoquinoline alkaloids metabolism. *Mol Plant* 10:975–989. <https://doi.org/10.1016/j.molp.2017.05.007>
- Madej T, Lanczycki CJ, Zhang D et al (2014) MMDB and VAST+: tracking structural similarities between macromolecular complexes. *Nucleic Acids Res* 42:D297–D303. <https://doi.org/10.1093/nar/gkt1208>
- Manina G, Bellinzoni M, Pasca MR et al (2010) Biological and structural characterization of the *Mycobacterium smegmatis* nitroreductase NfnB, and its role in benzothiazinone resistance: NfnB and BTZ resistance. *Mol Microbiol* 77:1172–1185. <https://doi.org/10.1111/j.1365-2958.2010.07277.x>
- Marchler-Bauer A, Bo Y, Han L et al (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203. <https://doi.org/10.1093/nar/gkw1129>
- Martin W, Rujan T, Richly E et al (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci USA* 99:12246–12251. <https://doi.org/10.1073/pnas.182432999>
- McHale MW, Kroening KD, Bernlohr DA (1996) Identification of a class of *Saccharomyces cerevisiae* mutants defective in fatty acid repression of gene transcription and analysis of the frm2 gene. *Yeast* 12:319–331. [https://doi.org/10.1002/\(SICI\)1097-0061\(19960330\)12:4%3C319:AIID-YEA914%3E3.0.CO;2-#](https://doi.org/10.1002/(SICI)1097-0061(19960330)12:4%3C319:AIID-YEA914%3E3.0.CO;2-#)
- Ming R, VanBuren R, Liu Y et al (2013) Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). *Genome Biol* 14:R41. <https://doi.org/10.1186/gb-2013-14-5-r41>

42. Mitchell AL, Attwood TK, Babbitt PC et al (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* 47:D351–D360. <https://doi.org/10.1093/nar/gky1100>
43. Mukherjee A, Rokita SE (2015) Single amino acid switch between a flavin-dependent dehalogenase and nitroreductase. *J Am Chem Soc* 137:15342–15345. <https://doi.org/10.1021/jacs.5b07540>
44. Novakova M, Mackova M, Antosova Z et al (2010) Cloning the bacterial *bphC* gene into *Nicotiana tabacum* to improve the efficiency of phytoremediation of polychlorinated biphenyls. *Bioengineered Bugs* 1:419–423. <https://doi.org/10.4161/bbug.1.6.12723>
45. Patel P, Young JG, Mautner V et al (2009) A phase I/II clinical trial in localized prostate cancer of an adenovirus expressing nitroreductase with CB1984. *Mol Ther* 17:1292–1299. <https://doi.org/10.1038/mt.2009.80>
46. Paterson ES, Boucher SE, Lambert IB (2002) Regulation of the *nfsA* Gene in *Escherichia coli* by SoxS. *J Bacteriol* 184:51–58. <https://doi.org/10.1128/JB.184.1.51-58.2002>
47. Peakall R, Ebert D, Scott LJ et al (2003) Comparative genetic study confirms exceptionally low genetic variation in the ancient and endangered relictual conifer, *Wollemia nobilis* (Araucariaceae): exceptionally low diversity genetic in Wollemi pine. *Mol Ecol* 12:2331–2343. <https://doi.org/10.1046/j.1365-294X.2003.01926.x>
48. Peterson FJ, Mason RP, Hovsepian J, Holtzman JL (1979) Oxygen-sensitive and -insensitive nitroreduction by *Escherichia coli* and rat hepatic microsomes. *J Biol Chem* 254:4009–4014
49. Pettersen EF, Goddard TD, Huang CC et al (2004) UCSF Chimera? A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>
50. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
51. Shiryev SA, Papadopoulos JS, Schaffer AA, Agarwala R (2007) Improved BLAST searches using longer words for protein seeding. *Bioinformatics* 23:2949–2951. <https://doi.org/10.1093/bioinformatics/btm479>
52. Song HN, Jeong DG, Bang SY, Paek SH, Park BC, Park SG, Woo EJ. (2015) Crystal structure of the fungal nitroreductase *Frm2* from *Saccharomyces cerevisiae*. *Protein Science* 24(7):1158–1163. <https://doi.org/10.1002/pro.2686>
53. Thiel Z, Rivera-Fuentes P (2019) Single-molecule imaging of active mitochondrial nitroreductases using a photo-crosslinking fluorescent sensor. *Angew Chem Int Ed*. <https://doi.org/10.1002/anie.201904700>
54. Thomas SR, McTamney PM, Adler JM et al (2009) Crystal structure of iodotyrosine deiodinase, a novel flavoprotein responsible for iodide salvage in thyroid glands. *J Biol Chem* 284:19659–19667. <https://doi.org/10.1074/jbc.M109.013458>
55. Tyagi M, de Brevern AG, Srinivasan N, Offmann B (2008) Protein structure mining using a structural alphabet. *Proteins* 71:920–937. <https://doi.org/10.1002/prot.21776>
56. Valiauga B, Williams EM, Ackerley DF, Čenas N (2017) Reduction of quinones and nitroaromatic compounds by *Escherichia coli* nitroreductase A (*NfsA*): Characterization of kinetics and substrate specificity. *Arch Biochem Biophys* 614:14–22. <https://doi.org/10.1016/j.abb.2016.12.005>
57. Wang B, Powell SM, Hessami N et al (2016) Crystal structures of two nitroreductases from hypervirulent *Clostridium difficile* and functionally related interactions with the antibiotic metronidazole. *Nitric Oxide* 60:32–39. <https://doi.org/10.1016/j.niox.2016.09.003>
58. Wang M, Jiang Y-Y, Kim KM et al (2011) A universal molecular clock of protein folds and its power in tracing the early history of aerobic metabolism and planet oxygenation. *Mol Biol Evol* 28:567–582. <https://doi.org/10.1093/molbev/msq232>
59. Waterhouse A, Bertoni M, Bienert S et al (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46:W296–W303. <https://doi.org/10.1093/nar/gky427>
60. Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 18:691–699. <https://doi.org/10.1093/oxfordjournals.molbev.a003851>
61. Williams EM, Little RF, Mowday AM et al (2015) Nitroreductase gene-directed enzyme prodrug therapy: insights and advances toward clinical utility. *Biochem J* 471:131–153. <https://doi.org/10.1042/BJ20150650>
62. Yap J-Y, Rohner T, Greenfield A et al (2015) Complete chloroplast genome of the wollemi pine (*Wollemia nobilis*): structure and evolution. *PLoS ONE* 10:e0128126. <https://doi.org/10.1371/journal.pone.0128126>
63. You S-H, Zhu B, Han H-J et al (2015) Phytoremediation of 2,4,6-trinitrotoluene by *Arabidopsis* plants expressing a NAD(P)H-flavin nitroreductase from *Enterobacter cloacae*. *Plant Biotechnol Rep* 9:417–430. <https://doi.org/10.1007/s11816-015-0379-y>
64. Zemla A (2003) LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res* 31:3370–3374. <https://doi.org/10.1093/nar/gkg571>
65. Zhang L, Rylott EL, Bruce NC, Strand SE (2017) Phytodetoxification of TNT by transplastomic tobacco (*Nicotiana tabacum*) expressing a bacterial nitroreductase. *Plant Mol Biol* 95:99–109. <https://doi.org/10.1007/s11103-017-0639-z>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)