In silico characterization of a nitrate reductase gene family and analysis of the predicted proteins from the moss *Physcomitrella patens*

Rigoberto Medina-Andrés and Verónica Lira-Ruan*

Laboratorio de Fisiología y Desarrollo Vegetal; Departamento de Bioquímica y Biología Molecular; Facultad de Ciencias; Universidad Autónoma del Estado de Morelos; Cuernavaca, México

**Keywords:** basal embryophytes, nitrate reductase, *Physcomitrella patens*, posttranslational regulation

**Abbreviations:** Cyt *b*, cytochrome *b*; MoCo, molybdenum cofactor; NO, nitric oxide; NR, nitrate reductase

Assimilatory nitrate reductase (NR; EC 1.7.1.1–3) catalyzes the reduction of nitrate to nitrite. This enzyme has a conserved structure common to fungi, algae and plants. However, some differences in the amino acid sequence between plant and algal NR suggest that the activity regulation mechanisms have changed during plant evolution. Since only NRs from angiosperms have been studied, the search and analysis of NRs from the moss *Physcomitrella patens*, a basal land plant, was performed to widen the knowledge of land plant NR structure. A family of three *nr* genes, named *ppnia1*, *ppnia1* and *ppnia2*, was localized in the *P. patens* genome. The predicted proteins are canonical NRs with the conserved domains Molybdenum-Cytochrome *b*-Cytochrome *b* reductase and possess 20 amino acid residues important for the enzymatic function conserved in plant and algal NRs. Interestingly, moss NRs lack a consensus sequence, common to angiosperm NRs, that is a target for posttranslational regulation. A phyllogenetic tree with embryophyte and green algae NR sequences was constructed and *P. patens* NRs localized at the base of embryophyte NR evolution. The data presented here suggest that bryophytes and vascular plants have different systems to regulate NR activity.

**Introduction**

Nitrogen is a fundamental macronutrient present in nucleic acids, proteins and cofactors in all organisms. An important way by which plants incorporate nitrogen to their metabolism is the assimilation of nitrate from the soil; nitrate is then converted to ammonia and further incorporated into amino acids. The first enzyme in this process is assimilatory nitrate reductase (NR; EC 1.7.1.1–3), a protein present in fungi, algae and plants.¹ NR is a dimeric soluble protein of around 100 kDa composed of three conserved domains forming an internal electron transport system. A Cyt *b* reductase domain, located at the C-terminus of the protein, takes electrons from NAD(P)H through its FAD cofactor; the electrons are channeled through a Cyt *b* domain to the MoCo domain, in the N-terminus of the protein, where they are used to reduce nitrate to nitrite.² ⁴ Beside its nitrate assimilatory function, NRs also use nitrite as substrate to produce nitric oxide (NO) in higher plants.⁵ ⁶ During the last decade, it has been shown that NO derived from NR is involved in root growth, stomatal closure, flower development and other responses in higher plants.⁷ ⁹ The green algae *Chlamydomonas reinhardtii* also produce NO in a nitrite-dependent manner; a mutant line lacking functional NR lost the capacity to produce NO, indicating that green algae NRs also reduce nitrate to NO.¹⁰ Thus, there is evidence that NR activity is important not only for nitrogen metabolism, but also for NO production in photosynthetic eukaryotes.

Expression of NR is tightly regulated by external and internal cues such as nitrate and sugar levels, photosynthesis rate and light.¹¹ ¹² In higher plants, NR activity is downregulated by low CO₂ levels and transition from light to dark.¹¹ ¹² The mechanism for NR inactivation in all known plant NRs is the reversible binding of a 14-3-3 protein to a phosphorylated Serine residue located in a conserved consensus sequence between the Cyt *b* and Cyt *b* reductase domains (see the review by MacKintosh and Meek).¹³ NR enzymes have a long evolutionary history; Stolz and Basu¹⁴ analyzed the phylogenetic relations of NRs from all organisms, demonstrating that eukaryotic NR have a monophyletic origin and that a three domain protein structure is conserved among fungi, algae and land plants, as well as the amino acid residues important for nitrate reduction. However, there is evidence that the NR gene structure and protein activity regulation has changed during plant evolution. The NR from the green alga *C. reinhardtii* is not regulated by a 14-3-3 protein.¹⁵ Moreover, NRs from *C. reinhardtii* and other green, brown and red algae lack the consensus sequence recognized by 14-3-3 proteins,¹⁶ ¹⁷ suggesting that this regulation mechanism is exclusive of plant NRs. However, the activity of red and green algae is
also regulated at the posttranscriptional level by unknown mechanisms.16,18-20 Furthermore, green algae nr genes are interrupted by a variable number of introns,17,20,21 while angiosperm nr genes have only three introns in conserved positions,22 indicating that during land plant evolution, the intron number in nr genes has been reduced.

In view of the fundamental role of NR in nitrogen metabolism and differences in gene structure and posttranslational regulatory mechanisms between green algae and angiosperms, the interesting question of how these characteristics evolved during land plant evolution emerges. In virtue of this, we present the first analysis of NRs from a basal embryophyte, as a way to better understand the NRs evolution in land plants. Bryophytes (liverworts, hornworts and mosses) emerged ~450 million years ago and are considered the basal clade in the land plant tree of life.23,24 Among the bryophytes, the mosses are the most diverse group and have been studied since 1930s. During the past decade, the moss Physcomitrella patens has emerged as very useful model for the analysis of all aspects of plant biology and evolution, reviewed by Cove et al.25 and Cove,26 and the release of the P. patens complete genome sequence27 boosted its use as model plant. In this work we present the identification of a family of three nr genes from P. patens, named ppnia1;1, ppnia1;2 and ppnia2. The predicted NR sequences showed that P. patens NRs are canonical NR proteins composed by the three conserved domains, MoCo, Cyt b and Cyt b reductase and possess 20 conserved amino acid residues important for enzyme function. Interestingly, P. patens NRs lack the consensus sequence recognized by 14-3-3 proteins to all studied angiosperms, suggesting that its activity is not downregulated by this mechanism.

Results and Discussion

Identification of the nitrate reductase genes from Physcomitrella patens genome. The nitrate reductase 1 from Nicotiana benthamiana (accession number BAE46746.1) was used as probe to perform a search for nr genes in the P. patens genome database cosmoss (www.cosmoss.org). Four putative genes were found, two of them in scaffold 58 (named Pp1s58_249v6.1 and Pp1s58_252v6.1) and two in scaffold 79 (named Pp1s79_76v6.1 and Pp1s79_76v6.2). The last two sequences occupy the same position in the scaffold and the sequences are identical, demonstrating that only one gene exists in that position. Thus, Pp1s79_76v6.1 was used to continue this work. The protein sequence deduced from Pp1s58_252v6.1 was used to search for similar genes in the GenBank database. The sequences annotated as PpNia1;1 (accession number BAE19754.1) and PpNia1;2, (accession number BAE19755.1) corresponded with Pp1s58_252v6.1 and Pp1s79_76v6.1, respectively. A third sequence (accession number XP_001762979.1) corresponded to sequence Pp1s58_249v6.1 from the cosmoss database and we called it PpNia2. The size and position of the three nr genes from P. patens are shown in Figure 1. The genes ppnia1;1 and ppnia2 are localized in the scaffold 58, they are 50.89 kb apart and have opposite translation orientation; ppnia1;2 is located in a different scaffold. The three genes are predicted to have six introns located at the same positions. Interestingly, nr genes in angiosperms only have three introns of variable length in conserved positions;22 the only data available from a non-seed vascular plant, the Lycophyte Selaginella moellendorffii, indicate that its nr genes have five and four introns in nia1 and nia2, respectively (http://genome.jgi-psf.org/Selmo1/Selmo1.home.html); finally, the number of introns in nr genes from green algae varies from two in Dunaliella tertiolecta17 to 10 and 18 in Volvox carteri and Chlorella vulgaris, respectively.20,21 This suggests that during Viridiplantae evolution the nr genes have lost some introns. The study of introns loss/gain has become a useful tool to analyze gene family evolution.28 Thus, the apparent tendency to the reduction of intron number in nr genes during land plant evolution deserves further research as new nr gene sequences from non-seed vascular plants and bryophytes are available.

Amino acid sequences analysis of NRs from P. patens. The predicted amino acid sequence from P. patens NRs were obtained and compared. PpNia1;1 and PpNia2 have 892 amino acids, their sequences are almost identical (99.8%), they differ only in two amino acids at positions 39 and 61 where Asn and Thr are found in PpNia1;1 while Lys and Ala are found in PpNia2, respectively. PpNia1;2 has 893 amino acids and has 89.5% similarity and 80.1% identity with PpNia1;1. The sequence of PpNia1;1 and PpNia1;2 were compared with Selaginella moellendorffii NR 1, (Protein ID XP_002972481.1) and with selected angiosperm species (Fig. 2). The alignment shows that P. patens NRs share the general structure of canonical eukaryotic NRs, they contain the MoCo-Cyt b-Cyt b reductase domains. P. patens NRs possess 20 conserved amino acids involved in NR function (Table 1 and Fig. 2). Among them the most important are a Cys residue (C200) that binds the MoCo; two His residues (H592 and H615) in the Cyt b domain that coordinate the heme iron, and three amino acids in the Cyt b domain that...
reductase domain that are involved in the NAD(P)H union, two of them, K757 and G820, are conserved in the three P. patens enzymes, but C919 is absent in PpNia1;2. This residue is considered to be the active site for electron transfer from NAD(P)H to FAD. Studies with Cyt b reductase fragment mutants from corn NR and Neurospora crassa NADPH:NR showed that any substitution of the invariant Cys severely reduce the NAD(P)H oxidase activity (Reviewed by Campbell). These data suggest that PpNia1;2 may have different catalytic rates than the other two enzymes.

An interesting difference between P. patens and vascular plant NRs is that moss enzymes lack the consensus sequence (K/R) (S/T)XS*(T/S)XP recognized by a 14-3-3 protein during NR activity downregulation (amino acids 544 to 549 in Fig. 2). P. patens NRs have Glu in position 547 instead of the Ser which is phosphorylated before union with a 14-3-3 protein in NR from vascular plants. When this Ser residue was changed for an Asp in NR from Nicotiana plumbaginifolia the mutated plant had constitutive NR activity even in the dark and produced more NR from Nicotiana plumbaginifolia than those plants that lack the 14-3-3 recognition sequence. Moreover, it was demonstrated that C. reinhardtii synthesizes a 14-3-3 protein that does not regulate its own NR but can bind and inactivate spinach NR. However, red and green algae downregulate NRs activity during light/dark transitions, high nitrate and ammonium levels. Upregulation of NRs activity in Rhodophyta, green algae and higher plants is mainly directed by expression of the nr genes and de novo synthesis of the protein. The downregulation of NR in red and green algae that lacks the 14-3-3 mechanism to downregulate NR activity in mosses to three in angiosperms. Moreover, our analysis shows that the known mechanism to downregulate NRs activity in angiosperms could be common to vascular plants, but not with the most basal land plants, the mosses. To confirm this it will be necessary to analyze the NR sequences from other bryophytes that unfortunately, are not available at the moment. Furthermore, this work opens the door for future detailed research on the mechanisms that regulate P. patens NRs activity that may be common to green algae.

The results presented here are startling because they suggest that the known mechanism that regulates NRs activity in vascular plants did not appear when the embryophytes emerged. Due to the novelty of this finding it is necessary to start experimental work to exhaustively analyze the P. patens NRs regulation at the transcriptional and posttranslational level to unravel the mechanisms that downregulate NR activity in mosses and green algae.

Evolutionary relationships of NRs from land plants. A phylogenetic analysis of NR from Viridiplantae, including green algae, P. patens, S. moellendorffii and angiosperms was performed (Fig. 3). The sequences used to construct the phylogenetic tree are listed on Table 2. Tree topology is consistent with the known algae a plant evolutionary history; embryophytes share a common ancestor with green algae. Then, at the base of land plants, NRs from P. patens form a first branch that separated before the division of non-seed vascular plants and angiosperms. The position of ppnia1;1 and ppnia2 genes in the P. patens genome (Fig. 1), the high identity (99.8%) of their amino acid sequences and their position in the tree suggest a recent duplication of this gene, as they form a group separated from ppnia1;2. Vascular plants are separated in two branches, one containing non-seed vascular plants, represented by S. moellendorffii, and another that includes all angiosperms. The topology of angiosperm and algae branches in Figure 3 is similar to those reported before. The result presented here, that summarizes the evolutionary history of Embryophytes NRs will be better defined in the basal branches as more sequences of Bryophytes and Pteridophytes become available.

The data presented here provides compelling evidence that during land plant evolution the nr genes have lost introns, from six in mosses to three in angiosperms. Moreover, our analysis shows that the known mechanism to downregulate NRs activity in angiosperms could be common to vascular plants, but not with the most basal land plants, the mosses. To confirm this it will be necessary to analyze the NR sequences from other bryophytes that unfortunately, are not available at the moment. Furthermore, this work opens the door for future detailed research on the mechanisms that regulate P. patens NRs activity that may be common to green algae. Finally, we addressed interesting questions on Viridiplantae NR evolution that deserve to be investigated.

Methods

Search for NR genes. The protein sequence of Nicotiana benthamiana NR1 (GenBank accession number BAE46746.1) was used as probe to run a BLASTP search to find putative nr genes in the P. patens genome database (www.cosmoss.org). The accession numbers of the retrieved sequences were used to get the complete gene sequences. The position in the genome and the structure of the genes was analyzed.
Figure 2. Amino acid sequence comparison between *P. patens* NR and selected vascular plant NR. Underlined sections indicate the protein domains: MoCo, solid line; Cyt b, dotted line and Cyt b reductase, dashed line. Conserved amino acid residues important for NR function are underlined and in bold type. Gray shaded sequence indicates the binding site for 14-3-3 proteins. Sequences were obtained from the GenBank database using the following accession numbers: *Nicotiana benthamiana* nia1 (BAE46746.1); *Cucumis sativus* nia2 (ADN96689.1); *Arabidopsis thaliana* nial (NP_177899.1); *Zea mays* nia1 (AAD38068.1); *Selaginella moellendorffii* nia1 (XP_002972481.1); *Physcomitrella patens* nia1 (BAE19754.1) and nia12 (BAE19755.1). Amino acid numbering corresponds to the *N. benthamiana* sequence.

Analysis of NR protein sequences. Deduced amino acid sequences from the three *P. patens* NRs were aligned using the ClustalW2 software and the percentage of similarity and identity between the sequences were determined using the MatGAT2.01 software. To compare the protein sequence of *P. patens* NRs with other plant NRs, a BLASTP search was done using the PpNia1;1 protein sequence (Accession number BAE19754.1) in the GenBank database (www.ncbi.nlm.nih.gov). Some of the retrieved sequences, four from angiosperms and one from a Lycophyte were aligned with PpNia1;1 and PpNia1;2 using the ClustalW software. An evolutionary tree of algae and land plants NRs amino acid sequences was constructed using the MEGA5 software. A bootstrap analysis with 1,000 replicates was carried on the trees inferred from the neighbor-joining method.

Disclosure of Potential Conflicts of Interest

There are no conflicts of interest to disclose.

Acknowledgments

Authors are grateful to an anonymous reviewer and editor in chief of CIB for they enlighten comments that help us to improve this work. Also we want to thank to Dr. Jorge Luis Folch Mallol and Dr. Ernesto Ortiz Suri for they helpful comments to this manuscript.
Figure 3. For figure legend, see page 5.
was designed to organisms that have only one
Sequences were obtained from GenBank and the accession number is provided.

Table 2. NR sequences used to construct the phylogenetic tree

| Abbreviation | Scientific name                  | Accession number |
|--------------|----------------------------------|------------------|
| Rcnia2       | Ricinus communis                 | XP_002513830.1   |
| Rcnia1       | Ricinus communis                 | AAG30576.1       |
| Vitis vinifera|                                 | XP_002285831.1   |
| Ptnia2       | Populus trichocarpa              | XP_002307415.1   |
| Smnia2       | Selaginella moellendorffii       | XP_002984312.1   |
| Smnia1       | Selaginella moellendorffii       | XP_002974281.1   |
| Cnria2       | Cucumis sativus                  | ADN96689.1       |
| Cmnia1       | Cucurbita maxima                 | P17569.1         |
| Cnria1       | Cucumis sativus                  | ADK77877.1       |
| Bnnia1       | Brassica napus                   | P39867.1         |
| Ptnia1       | Populus trichocarpa              | XP_002301015.1   |
| StNr6        | Solanum tuberosum               | BABA93534.1      |
| Ntntia1      | Nicotiana tabacum                | 227925           |
| Slia1        | Solanum lycopersicum             | P17570.1         |
| NbNR1        | Nicotiana benthamiana            | BAE46746.1       |
| MtNR2        | Medicago truncatula              | ADV03139.1       |
| NbNR2        | Nicotiana benthamiana            | BAE96752.1       |
| Brassica      | Brassica rapa subsp. chinensis  | AF93242.1        |
| Beta         | vulgaris                         | ABW05098.1       |
| Nicotiana     | sylvestris                       | 227926           |
| Ntntia2      | Nicotiana tabacum                | P08509.2         |
| Gmnia2       | Glycine max                      | P39870.1         |
| StNr2        | Solanum tuberosum               | AAB18985.1       |
| StNr3        | Solanum tuberosum               | AAB52786.1       |
| Bnnya2       | Brassica napus                   | P39868.1         |
| Alnia1       | Arabidopsis lyrata subsp. lyrata| XP_002889158.1   |
| Atntia1      | Arabidopsis thaliana             | NP_177899.1      |
| Pvinia2      | Phaseolus vulgaris               | P39866.1         |
| Mtnia1       | Medicago truncatula              | ADV03138.1       |
| Pxhnia2      | Petunia x hybrida                | P36859.1         |
| Solnia1      | Spinacia oleracea                | P23312.1         |
| Lijnia1      | Lotus japonicus                  | P39869.1         |
| Cichorium     | Intybus                          | P43101.1         |
| Pxhnia1      | Petunia x hybrida                | AAA33712.1       |
| Solnia2      | Spinacia oleracea                | BAA13047.1       |
| Thellungiella| halophila                       | BAJ33682.1       |
| Os(j)nia1     | Oryza sativa Japonica Group      | NP_001062006.1   |
| Zmnia1       | Zea mays                         | AAD38068.1       |

Sequences were obtained from GenBank and the accession number is provided. No abbreviation was designed to organisms that have only one NR sequence.
Table 2. NR sequences used to construct the phylogenetic tree

| Genus            | Species Name | Accession Number |
|------------------|--------------|------------------|
| Attnia           | Arabidopsis thaliana | NP_174901.1 |
| Alnia            | Arabidopsis lyrata subsp. lyrata | XP_002891229.1 |
| Pnav1            | Phaseolus vulgaris | P39865.1 |
| Sbnia1           | Sorghum bicolor | XP_002444490.1 |
| Oryza sativa Indica Group | EEC74079.1 |
| Osjinya3         | Oryza sativa Japonica Group | NP_001048235.1 |
| Hvnia1           | Hordeum vulgare subsp vulgare | P27967.1 |
| Sbnia3           | Sorghum bicolor | XP_002454625.1 |
| Osjinya2         | Oryza sativa Japonica Group | NP_001062009.1 |
| Gmnia1           | Glycine max | P54233.1 |
| Hvnia7           | Hordeum vulgare subsp. vulgare | P27968.1 |
| Sbnia2           | Sorghum bicolor | XP_002454083.1 |
| Zmnia3           | Zea mays | P49102.1 |
| Chlorella vulgaris | UTEX259 | ACP44801.1 |
| Chlorella vulgaris NJ-7 | ABJ912084.1 |
| Chlorella variabilis | EFNS2691.1 |
| Ectocarpus siliculosus | CBNZ7846.1 |
| Volvox carteri f. nagariensis | XP_002955156.1 |
| Gracilaria tenuistipitata | ACX31652.1 |
| Dunaliella tertiocota | AAL79356.1 |
| Dunaliella viridis | AATT72293.1 |
| Thalassiosira pseudonana | CCMP1335 | XP_0022994410.1 |
| Dunaliella salina | AAPP7505.1 |
| Phaeocystis tricornutum | AAV69996.1 |
| Ostreococcus lucimarinus | CCE9901 | XP_001420098.1 |

Sequences were obtained from GenBank and the accession number is provided. No abbreviation was designed to organisms that have only one NR sequence.