Genome-wide identification and characterization of an amino acid permease gene family in Nicotiana tabacum†

Yingying Zhao,‡ab Yalong Xu,‡b Zhong Wang, b Jianfeng Zhang, b Xia Chen, b Zhengfeng Li, c Zefeng Li, b Lifeng Jin, b Pan Wei, b Lin Zhang, d Xiaoquan Zhang, d Ran Wang ‡ b and Fang Wei a

The amino acid permease (AAP) protein family is an important class of amino acid transporter. However, little information about this gene family is available in tobacco. Here, a total number of 15, 7, and 6 full-length putative AAP genes have been identified in the genomes of Nicotiana tabacum and progenitors of modern tobacco, Nicotiana sylvestris and Nicotiana tomentosiformis, respectively. We performed a multiple analyses of the AAP gene family in tobacco and reported data including phylogenetic relationships, gene structures, predicted protein structures and information about conserved motifs. These tobacco AAP family members shared high levels of similarity in their nucleotide and amino acid sequences. The gene structures of the NtAAP loci were all separated by introns. Phylogenetic analyses indicated that these AAP genes were clustered into three groups. We also performed expression profiling analyses of the NtAAP genes in various tobacco tissues/organs using quantitative PCR. The differential expression patterns of the NtAAP genes suggested that the family members might have different biological functions in tobacco plants. To study the function of NtAAP2-2, its coding sequence was cloned and overexpression/RNA interference vectors were constructed and these vectors were transformed into tobacco. Compared with wild-type tobacco, the contents of Glu, Gin, Asp and Asn were found to change in the transgenic tobacco leaves, indicating that NtAAP2-2 was involved in the transportation of Glu, Gin, Asp and Asn in tobacco. Our results can serve as fundamental research for future evolutionary and functional characterization studies of the AAP genes in tobacco.

Introduction

Nitrogen (N) is an important macronutrient and has been reported to be essential for the growth, development and reproduction of higher plants. N is used for the synthesis of essential compounds such as chlorophyll, nucleic acids (DNA and RNA), and amino acids (peptides and proteins). Limitations of the supply of N will lead to severe metabolic disturbance and growth inhibition in plants. In the majority of plants, the assimilation of nitrogen is not only in the form of inorganic-nitrogen such as nitrate and ammonium, which is converted into amino acids in root or shoot tissues, but also in the form of organic-nitrogen including urea, amino acids, etc.1 The assimilation form of nitrogen is affected by many factors, and it mainly depends on the plant species, amino acid concentrations and soil pH values.2,4 The inorganic nitrogen absorbed from the soil is used in the synthesis of amino acids, which are immediately used for metabolism, transiently stored (e.g. in the vacuole) or transported in the phloem to developing vegetative or reproductive sink tissues. In most plants, the absorption and redistribution of amino acids depend on the plasma membrane transporters. The amino acids synthesized in the leaves were transported to the sink tissues mainly through the phloem, but the amino acids synthesized in the roots were transported to the mature leaves with the transpiration stream through the xylem. The functions of amino acid transporters are mainly involved in three fields: notably in xylem and phloem loading, source to sink transportation of amino acid, and the uptake of amino acid from external environment.5,6 Therefore, amino acid transporters play an indispensable role in the regulation of plant growth and development.
It has been demonstrated that the amino acid transporters are divided into two major superfamilies: the amino acid, polyamine and choline transporters superfamilly (APC) and the amino acid transporter family (ATF). The APC superfamily consists of at least five families: amino acid permeases (AAPs), proline transporters (AtProTs), lysine and histidine transporters (AtLHTs), aromatic and neutral amino acid transporters (AtANTs) and auxin transporters (AUXs). Glutamine, glutamate and asparagine are the major transport forms of organic nitrogen, comparatively, aspartate is transported by most AAPs with an extremely low affinity and efficiency. In Arabidopsis, there are at least 53 genes annotated as amino acid transporters, which are present at the plasma membrane or tonoplast. Among all the transporters, 14 APC-type genes have been identified. 9 APC proteins (CAT-CAT9) have a 14 transmembrane (TM) topology, while the remaining 5 only have a 12 TM topology. Furthermore, all the characterized ATF transporters are proton coupled, and have a 9 to 11 transmembrane (TM) domain topology.

The AAPs have been suggested to be involved in a number of physiological processes in plants. Up to now, 8, 11, 19 AAP genes have been identified in Arabidopsis thaliana, Medicago sativa, Oryza sativa, respectively. AtAAP1 belongs to a family of eight members (AtAAP1-8) in Arabidopsis, and yeast complementation experiment showed that AAP1 mediated transport of neutral and acidic amino acids. The expression of PsAAP1 in Pisum sativum L. was found in the seed coat, cotyledon surface and the phloem throughout the plant. It was suggested the PsAAP1 might be the major amino acid transporter in the phloem, and might also function on seed development and the accumulation of storage protein in seeds, since transgenic PsAAP1 in pea limited seed protein levels. The importance of StaAAP1 in long-distance transport of amino acids from source leaves to sink organs has also been pointed out in Solanum tuberosum. The Arabidopsis amino acid transporters display tissue- and organ-specific expression patterns during the development. AtAAP1 is localized to the root tip and epidermis cells including root hairs, and it is important in the uptake of neutral and acidic amino acids when soil solution concentration is higher than 50 mM. Zhang et al. demonstrated that AtAAP2 located to the phloem throughout the plant, and was important for amino acid transportation from xylem to phloem and the sink N/C supply. AtAAP3 is presented in the root phloem, and transports neutral and basic amino acids, but atAAP3 mutants showed no obvious phenotypic defects. AtAAP4 is expressed in leaves, stems, and flowers, whereas AtAAP5 transcripts is found in all tissues, but with different levels. Furthermore, AtAAP6 and AtAAP8 were characterized as high affinity amino acid transporters. The expression of AtAAP8 was mainly in the veins of young flowers, peduncle, siliques and in very young seeds. The atAAP8 mutants showed a reduction in seed numbers per silique, suggesting that AtAAP8 was responsible for the import of amino acid into the seed. Tobacco is one of the most important economic crops in the world. The content of amino acids will affect the quality of tobacco. Since the functions of tobacco AAP genes are unknown, it is necessary to identify and characterize the AAP gene family.

In our study, we identified fifteen AAP genes in the allotetraploid Nicotiana tabacum (N.tabacum), and 7, 6 AAP genes in the diploid tobacco progenitor Nicotiana sylvestris (N.sylvestris) and Nicotiana tomentosiformis (N.tomentosiformis), respectively. A systematic analysis of the NtAAP genes in N. tabacum was presented, including the location of each member of NtAAP gene family in the chromosome, the organization of their intron/exon, their evolutionary relationships and the expression profiles of the NtAAP genes in tobacco tissues. The findings of this study will serve as a foundation for further studies of NtAAP genes at physiological and molecular levels.

Experimental section

Materials

N. tabacum (Honghua Dajinyuan) was used in the analysis of the expression profiles of the NtAAP genes. Tissue materials including roots, stems, leaves at tobacco resetting stage, and vigorous stage, flowering bud stage, flowering stage, topping stage, and middle leaf maturing stage, were collected from the experimental field in the Yunnan province of China. In addition, sepalas, corollas, anthers, capillaments, stigmas, styles and ovaries at flowering period were also harvested. All the plant materials were immediately frozen in the liquid N2 after collection, and then stored at −80 °C prior to RNA extraction.

To clarify the function of the NtAAP2 gene under N starvation, tobacco seedlings were grown under N deficient treatment for two weeks. The N stress nutrient solution contained the following compounds: KH2PO4 0.136 g L⁻¹, K2SO4 0.261 g L⁻¹, CaCl2 0.222 g L⁻¹, MgSO4 0.493 g L⁻¹, H2BO3 6.2 mg L⁻¹, KI 0.83 mg L⁻¹, MnSO4 22.3 mg L⁻¹, ZnSO4 8.6 mg L⁻¹, Na2MO4 0.25 mg L⁻¹, CuSO4 0.025 mg L⁻¹, CoCl2 0.025 mg L⁻¹, Fe-EDTA 0.042 g L⁻¹, 0.101 g L⁻¹ KNO3, 0.264 g L⁻¹ (NH4)2SO4. Tobacco seedlings, were treated for 2 weeks, the fifth leaf was used for total RNA isolation and the analysis of amino acid content.

Tobacco AAP gene identification

We downloaded 8 AtAAPs and 19 OsAAPs protein sequences from Arabidopsis thaliana genome database website TAIR (http://arabidopsis.org) and Oryza sativa genome database website TIGR (http://rice.plantbiology.msu.edu) as the query sequences, respectively. We then performed BLASTX search against N. tabacum, N. sylvestris and N. tomentosiformis protein sequences in China Tobacco Genome Database (V3.0) respectively, with the parameter E-value 1.0 × 10⁻⁵. Non-redundant candidate AAP were obtained from the blastp result. The Pfam database (http://pfam.sanger.ac.uk/search) and SMART (http://smart.embl-heidelberg.de/) were used to further verify the reliability of these candidate sequences and confirm each candidate NtAAP protein as a member of AAP family. The AAP protein sequences of other corresponding plant species were collected by searching the NCBI GeneBank database using “amino acid permease” as a query keyword.
Gene structure analysis

The gene structures were drawn by the Gene Structure Display Server (GSDS, http://gsds.cbi.pku.edu.cn/) using mRNA and genomic sequence obtained from the genome sequences.

Protein sequences of NtAAP genes characteristic analysis

The conserved motifs in the protein sequences were identified with the Multiple Em (Expectation Maximization) for Motif Elicitation (MEME) program Version 4.11.2 (http://meme-suite.org/tools/meme). The MEME program was employed using the following parameters: number of repetitions, zero or one occurrence per sequence; maximum number of motifs, 10; optimum motif width set to >40 and <120. The protein structure of each NtAAP protein was confirmed by SMART (http://smart.embl-heidelberg.de/). In addition, protein sequences of NtAAP genes were also analyzed by using the online ExPASy programs (http://www.expasy.org/tools/protparam.html). The number of amino acids, amino acid composition, molecular weight, theoretical isoelectric point (pI), and instability index were obtained by this tool. Furthermore, the topology of NtAAP in the N. tabacum was predicted by combining the results from multiple sequence alignment, secondary structure prediction, and transmembrane helix prediction. Multiple sequence alignment was performed using DNAMAN with default parameters. Accordingly, hydrophobicity analysis was performed using HMMTOP (http://www.enzim.hu/hmmtop/) and TMRPres2D.

Phylogenetic analysis

To examine the evolutionary relationships among the AAP genes of tobacco and other plant species, multiple sequences alignment was performed using Clustal W with the AAP protein sequences of N. tabacum, Arabidopsis thaliana, Zea mays, Brassica rapa and Oryza sativa. Gaps and ambiguously aligned regions were removed. Next, an unroot phylogenetic tree was constructed based on the alignment using ML method implemented in MEGA5.1. Bootstrap testing was performed with 1000 re-samplings, and then colored with the iTOL tool (http://itol.embl.de/).

Gene expression analysis

Total RNA was extracted using a Superpure Plantpoy RNA kit (Gene Answer), and then treated by RNA-free DNase I. The concentration and quality of RNA samples were evaluated with Nanodrop 2000 (Thermo), and the integrity of RNA was checked by 1% agarose gel electrophoresis. The first-stand of cDNA was synthesized from 1 μg total RNA using Reverse Transcript M-MLV (Gene Answer) with Random Primer. The concentration of cDNA was measured by Nanodrop 2000, and then the volume was diluted to the concentration of 50 ng μL⁻¹.

For quantitative PCR analysis, the gene specific primers were designed based on the tobacco coding sequence, and 26 s RNA was used as a reference gene. The primer sequences were listed in Table S1.† qPCR was performed using a 2× SYBR Green qPCR Mix (Gene Answer) on a Bio-Rad CFX96 TM instrument (Bio-Rad). The qPCR program was set as follows: 95 °C for 3 min, 40 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 40 s. Each reaction was carried out in 25 μL reaction volume containing 12.5 μL SYBR Green Mix, 150 ng cDNA sample, 1 μL upstream/downstream primer. Three replicates were set for each sample. The average value of the three Cps was obtained and then calculated with the method of $2^{-ΔΔC_t}$.

Cloning of NtAAP2-2, vector construction, plant transformation and confirmation

Primers for NtAAP2-2 cloning were designed by DNAMAN (V6.0) software, and the coding sequence of NtAAP2-2 were cloned by PCR amplification with high-fidelity DNA polymerase (TransGen Biotech). The 50 μL PCR amplification reaction mixture contained 10× PCR buffer II, 10 mM dNTP mix, 1 μM of each primer, 1 μL HiFaiase, and 150 ng of template cDNA. PCR amplification was performed on a Bio-Rad DNA thermal cycler.

To clarify the physiological role of NtAAP2-2 in N. tabacum, we designed gene-specific primers [NtAAP2-2-CD5 list in Table S1†] to amplify the full length CDS of NtAAP2-2 gene. The sequence was then cloned into the overexpression (OE) vector (Super pCAMBIA1300) with SpeI and KpnI restriction sites. Besides, we constructed the NtAAP2-2-RNAi (RNA interference) vector. The primers with attB adapters were designed from the partial CDS of NtAAP2-2, listed in the ESI Table S1.† After amplification, the fragment was then integrated into the expression vector pHellsgate2 by BP Clonase enzyme mix (invitrogen).

NtAAP2-2-RNAi and NtAAP2-2-OE plasmids were all transformed into Agrobacterium strain GV3101, which were used to generate N. tabacum transgenic lines. The transgenic method was described in the previous reports.6,23 Wild-type (K326) plants were used as controls in the below experiments. T₁ transgenic plants were grown in the greenhouse until seed harvesting. The NtAAP2-2-OE T₁ transgenic plants were obtained through the selection on MS medium containing 250 mg L⁻¹ cephalosporin and 5 mg L⁻¹ hygromycin. The NtAAP2-2-RNAi T₁ transgenic plants were selected by 150 mg L⁻¹ kanamycin and 250 mg L⁻¹ cephalosporin. To confirm the integration of transgenic plants, PCR was used to screen the genomic DNA. The NtAAP2-2-RNAi transgenic plants were selected by using primers of marker gene Kanamycin (npII). The NtAAP2-2-OE transgenic plants were selected by hygromycin primers, NtAAP2-2 flanking primers and flag-R, listed in Table S1.† The amplification program as list: 94 °C, 3 min; 94 °C, 30 s, 55 °C, 30 s, 72 °C, 2 min, 28 cycles; 72 °C, 10 min; 4 °C.

Free amino acid analysis

The amino acid contents in tobacco leaves from transgenic individuals and wild type plants were evaluated by amino acid analyzer method. Free amino acid was extracted from 0.1 g freeze-dried leaves with 2 mL lithium citrate. The mixture was extracted with ultrasonic for 30 min followed by centrifugation (10 000 rpm, 10 min), the supernatant was collected and filtered through a 0.22 μm filter, and then transferred to amino acid analyzer. The standard amino acid mixture was supplemented
with 2.5 μmol mL⁻¹ asparagine, asparagine, alanine, arginine, glycine, glutamic, glutamine, histidine, isoleucine, leucine, lysine, benzene alanine, proline, serine, threonine, tyrosine and tryptophan.

Results and discussion

Tobacco putative AAP gene identification

A total of 18 AAP potential genes were identified from the N. tabacum from the tobacco genome databases, and then were used for multiple sequence alignment by DNAMAN software. Three NtAAP potential genes (the information of these genes were showed in Table S2†) were removed for their obvious lower similarity (less than 40%) compared to other gene sequences, and finally 15 putative NtAAP genes were retained (Table 1). Among the 15 putative NtAAP sequences, NtAAP3, NtAAP6 and NtAAP2 have 7, 6 and 2 allelic haplotypes, respectively. The allelic haplotypes of each AAPs were indicated additional “−1” to “−7” behind the gene name. Sequence identity analysis showed that the amino acid sequences of NtAAP proteins shared 69.96% identity, and the nucleotide identity was up to 71.04% (Table S3†). NtAAP proteins were analyzed by ExPaSy programs tool. As shown in Table 1, almost all the putative NtAAP CDS were about 1500 bp in length, which gave rise to about 500 amino acid sequences of putative NtAAP proteins. But the gene and CDS length of NtAAP6-3 was 12 586 bp and 1332 bp, respectively. Instability index of all the predicted proteins were inferior to 40. Isoelectric point was estimated among 8.36 to 9.52, and the tobacco putative NtAAP proteins were all alkaline. Furthermore, we also identified 7 and 6 putative AAP genes in diploid ancestor N. sylvestris and N. tomentosiformis, respectively.

Table 1  Putative amino acid permease genes identified in tobacco

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene name</th>
<th>MW (kDa)</th>
<th>PI</th>
<th>Amino acid size (bp)</th>
<th>CDS size (bp)</th>
<th>gDNA size (bp)</th>
<th>Instability index</th>
<th>Conserved domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. tabacum</td>
<td>NtTab0475230 NtAAP3-1</td>
<td>52.73</td>
<td>8.69</td>
<td>480</td>
<td>1443</td>
<td>4376</td>
<td>31.47</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0186340 NtAAP3-2</td>
<td>52.73</td>
<td>8.69</td>
<td>480</td>
<td>1443</td>
<td>4440</td>
<td>31.47</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0484720 NtAAP3-3</td>
<td>52.49</td>
<td>8.69</td>
<td>478</td>
<td>1437</td>
<td>4583</td>
<td>32.12</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0049050 NtAAP3-4</td>
<td>52.48</td>
<td>8.69</td>
<td>478</td>
<td>1437</td>
<td>2303</td>
<td>32.12</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0126970 NtAAP3-5</td>
<td>52.75</td>
<td>8.52</td>
<td>480</td>
<td>1443</td>
<td>2384</td>
<td>30.68</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0889300 NtAAP3-6</td>
<td>52.76</td>
<td>8.36</td>
<td>480</td>
<td>1443</td>
<td>2384</td>
<td>30.68</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0329990 NtAAP3-7</td>
<td>52.55</td>
<td>8.36</td>
<td>479</td>
<td>1440</td>
<td>2383</td>
<td>31.81</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0126360 NtAAP3-8</td>
<td>55.04</td>
<td>9.09</td>
<td>501</td>
<td>1506</td>
<td>4922</td>
<td>38.55</td>
<td>2 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0759460 NtAAP3-9</td>
<td>51.97</td>
<td>8.73</td>
<td>475</td>
<td>1428</td>
<td>5056</td>
<td>36.3</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0648520 NtAAP3-10</td>
<td>48.62</td>
<td>9.07</td>
<td>443</td>
<td>1332</td>
<td>12 586</td>
<td>37.65</td>
<td>2 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0439440 NtAAP3-11</td>
<td>53.87</td>
<td>8.89</td>
<td>487</td>
<td>1464</td>
<td>4407</td>
<td>37.08</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0331750 NtAAP3-12</td>
<td>51.23</td>
<td>8.58</td>
<td>464</td>
<td>1395</td>
<td>3864</td>
<td>32.61</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0475230 NtAAP3-1</td>
<td>52.73</td>
<td>8.69</td>
<td>480</td>
<td>1443</td>
<td>4376</td>
<td>31.47</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0186340 NtAAP3-2</td>
<td>52.73</td>
<td>8.69</td>
<td>480</td>
<td>1443</td>
<td>4440</td>
<td>31.47</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0484720 NtAAP3-3</td>
<td>52.49</td>
<td>8.69</td>
<td>478</td>
<td>1437</td>
<td>4583</td>
<td>32.12</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0049050 NtAAP3-4</td>
<td>52.48</td>
<td>8.69</td>
<td>478</td>
<td>1437</td>
<td>2303</td>
<td>32.12</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0126970 NtAAP3-5</td>
<td>52.75</td>
<td>8.52</td>
<td>480</td>
<td>1443</td>
<td>2384</td>
<td>30.68</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0889300 NtAAP3-6</td>
<td>52.76</td>
<td>8.36</td>
<td>480</td>
<td>1443</td>
<td>2384</td>
<td>30.68</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0329990 NtAAP3-7</td>
<td>52.55</td>
<td>8.36</td>
<td>479</td>
<td>1440</td>
<td>2383</td>
<td>31.81</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0126360 NtAAP3-8</td>
<td>55.04</td>
<td>9.09</td>
<td>501</td>
<td>1506</td>
<td>4922</td>
<td>38.55</td>
<td>2 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0759460 NtAAP3-9</td>
<td>51.97</td>
<td>8.73</td>
<td>475</td>
<td>1428</td>
<td>5056</td>
<td>36.3</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0648520 NtAAP3-10</td>
<td>48.62</td>
<td>9.07</td>
<td>443</td>
<td>1332</td>
<td>12 586</td>
<td>37.65</td>
<td>2 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0439440 NtAAP3-11</td>
<td>53.87</td>
<td>8.89</td>
<td>487</td>
<td>1464</td>
<td>4407</td>
<td>37.08</td>
<td>1 Aa_trans</td>
</tr>
<tr>
<td></td>
<td>NtTab0331750 NtAAP3-12</td>
<td>51.23</td>
<td>8.58</td>
<td>464</td>
<td>1395</td>
<td>3864</td>
<td>32.61</td>
<td>1 Aa_trans</td>
</tr>
</tbody>
</table>

Exon–intron structure and chromosome location of NtAAP genes

Based on phylogenetic analysis, the AAP genes in the tobacco could be divided into two major clades (Fig. 1a). The exon–intron structures of NtAAP genes were analyzed by Gene Structure Display Server (GSDS), which were showed in the Fig. 1c. All of the NtAAP genes were disrupted by introns. In the clade A, the CDS of NtAAP3-1 to NtAAP3-4 were comprised by 5 exons, and the first exon was larger than other NtAAP3 exons, which had 6 exons. NtAAP6 had 7 exons that belonged to the clade B, in addition, NtAAP2 had 8 exons. Comparative analysis among the NtAAP family suggested that the first and the last exon of NtAAP3 were more likely to be splitted by introns. In addition,
we ranked the intron positions relative to the conserved exons, and finally identified 7 putative positions for introns, among which 5 introns (except introns 1 and 7) were found in most of the NtAAP sequences (Fig. S1†). The lack of one or more introns, which mainly occurred in the 1st, 2nd, 7th intron positions, lead to the formation of larger exons, including the 333 bp and/or 653 bp exons in some of the NtAAP genes (Fig. S1†).

Structure and motif analysis of the NtAAP proteins
We also analyzed the amino acid sequences of the NtAAPs using the MEME motif search tool, and identified a total of seven distinct motifs. The length and sequence information of these 7 motifs were listed in Table S4.† The position of each motif in the NtAAP protein was listed in Fig. 1b. Motifs 1–4 were transmembrane domain with transmembrane transporter activity, and were present in all NtAAPs. Motif 5 was existed in all the NtAAP proteins except for NtAAP6-6. Motif 6 was only presented in the NtAAP2, and was located in its N-terminus. Moreover, the motif 5 and 6 were functional unknown domains, since there was no hit when searching the sequence in the InterProScan.

In addition, all the NtAAP protein sequences were searched against the SMART database to predict the putative protein structures, and the results revealed that each of the proteins contained at least one Aa_trans domains (Table 1).

Transmembrane structure of the NtAAPs
The hydrophobic transmembrane domain of the NtAAPs was predicted by HMMTOP. It was shown that the family members contained 9–12 transmembrane helices (Fig. S2†). According to the position of the N/C-terminal and the number of predicted transmembrane helices, the transmembrane structure could be divided into 3 types: N/C terminals were all predicted in cytoplasmic contained 10 TMs, including NtAAP3-1 to NtAAP3-4, NtAAP6-1, NtAAP6-5 and NtAAP6-6; N/C terminals were all predicted in extracellular contained 12 TMs, including NtAAP2, and NtAAP3-5 to NtAAP3-7; N/C-terminals of the other NtAAP proteins were predicted in the extracellular and cytoplasmic, respectively, containing 9 TMs (Fig. S2†). Multiple sequences alignment displayed that the limits of the TM 1 segment (H, I, I, T, A, V, I, G, S, G, V, L, S, L, A, W, A) were presented in all the NtAAPs (Fig. S3†), therefore, this region was highly conserved.

Phylogenetic analysis of the tobacco AAP genes
Classification of genes and phylogenetic relationship are important for the functional analysis of a gene family. To examine the evolutionary relationship among the AAP genes of tobacco and other plant species, multiple sequences alignment was performed using Clustal W, and an unroot tree was constructed based on the alignment using ML method implemented in MEGA 5.1. Based on the phylogenetic analysis, the AAP proteins from N. tabacum, Arabidopsis thaliana, Zea mays, Brassica rapa and Oryza sativa were classified into 4 main clusters (1, 2, 3, 4), which was consistent with the previous report.14 The cluster 1 was sub-divided into three subclusters (A, B, C), and the cluster 4 was sub-divided into two subclusters (A and B) (Fig. 2). The NtAAPs were divided into clusters 1 and clusters 4, which contained 9 and 6 NtAAP proteins, respectively. All of the NtAAP3 and NtAAP2 proteins belonged to the cluster 1A and cluster 1C, respectively. It is well known that gene structural diversity drives the evolution of multigene families. In Cluster 1A, NtAAP3-1 to NtAAP3-4 in the same branch contained different exons numbers with other NtAAP3 genes (Fig. 1c), indicating that these NtAAP3 genes might have different evolutionary processes. In addition, cluster 1A also included Arabidopsis AtAAP3 and AtAAP5. Cluster 1C only contained one rice AAP protein (and multiple eudicot sequences), which was distributed to the same branch with NtAAP2 proteins in tobacco. Notes that NtAAP2 gene has been appeared before the evolution of dicotyledonous and monocotyledonous. All of these genes in cluster 1C appeared to be involved in the loading of amino acid into phloem.13 NtAAP6 proteins in tobacco were divided into cluster 4B, which also contained the Arabidopsis
AtAAP1, AtAAP6, AtAAP8, and OsAAP6. NtAAP6-1/NtAAP6-2 and OsAAP6, other NtAAP6 proteins and AtAAP1, AtAAP6, AtAAP8 were branched into a group, respectively. These genes might function in the transportation of amino acid from xylem to phloem.

Interestingly, cluster 1B and cluster 2 only contained monocots protein sequences, it is possible that this large group of AAPs might be evolved after monocot–dicot divergence, and might have specialized roles in monocots or dicots.

Expression pattern analysis of the NtAAP genes in tobacco tissues

To detect the relative expression levels of each NtAAP gene in the tobacco leaves, stems, roots and flowers from different developmental stages, a comprehensive quantitative PCR analysis was further performed. As shown in Fig. 3, NtAAP6-1, NtAAP6-3 and NtAAP6-6 were highly expressed in stems and root, whereas NtAAP6-5 was mainly expressed in leaves. NtAAP2 was highly expressed in stems at resetting stage and vigorous stage. In addition, NtAAP2 gene was also ubiquitously expressed in the tissues of other growth stages. NtAAP3-1, NtAAP3-6 and NtAAP3-7 were highly expressed in roots and stems at resetting, vigorous and flowering stage (Fig. 3). However, their expression levels at flowering bud and topping stage were higher in leaves than those in other tissues. Furthermore, NtAAP3 gene was also expressed in the floral organ.

Phylogenetic analysis and expression patterns can be used to predict the molecular functions of genes involved in different physiological processes. AtAAP3, AtAAP5 and NtAAP3 are assigned to the cluster 1A, and these two AAP genes in Arabidopsis appear to be associated with loading of amino acid into phloem. The function of AtAAP3 in loading amino acid into phloem seems to be restricted to the roots. AtAAP5 is associated with root acquisition of amino acid. In our analysis, the transcripts of NtAAP3-6 and NtAAP3-7 genes are not only detected in the stems and roots at resetting and vigorous stage, but also in flowers at bud and flower stage. Therefore, NtAAP3 might play a role not only in phloem loading, but also in importation of amino acids into sink tissues. In addition, AtAAP2, AtAAP4 and NtAAP2 are classified into cluster 1C. AtAAP2 functions in phloem loading of amino acids along the transport pathway, and is essential for N supply to seed sinks. NtAAP2 is highly expressed in the stems at resetting stage and vigorous stage, and was also detected in the floral organ.
Therefore, NtAAP2 may be involved in the delivery of amino acids to seed sinks. NtAAP6 genes, AtAAP1, 6 and 8 are assigned into cluster 4B. AtAAP6 is mainly expressed in sink tissues, such as roots and sink leaves. GUS expression studies revealed that AtAAP6 is also expressed in the xylem parenchyma cell. It’s predicted to be important for xylem to phloem transfer of amino acids in Arabidopsis. Although AAP1, AAP5, LHT1 have been demonstrated to affect root uptake of amino acids, only AAP1 has been shown to be localized to the root epidermis and root hairs, and function in the import of glutamate and neutral amino acids into root hairs and epidermis at relatively high amino acid concentrations. In addition, AtAAP1-mediated...
uptake of amino acids by the embryo is important for the storage protein synthesis and seed yield.\textsuperscript{27} \textit{AtAAP8} plays a crucial role for the uptake of amino acids into the endosperm, and supplies the developing embryo with amino acids during early embryogenesis.\textsuperscript{28} The transcripts of \textit{NtAAP6-1}, \textit{NtAAP6-3}, and \textit{NtAAP6-6} are detected in the root at relatively high level, therefore, these \textit{NtAAP6} genes might have similar functions with the \textit{AAP} genes in \textit{Arabidopsis}.

Functional analysis of \textit{NtAAP2} in amino acid transportation

We successfully cloned the coding sequence and the gene sequence of \textit{NtAAP2-2} from tobacco (Honghua Dajinyuan) (Fig. S4A\textsuperscript{†}). Bioinformatics analysis indicated that the CDS of \textit{NtAAP2-2} was 1539 bp, encoding 512 amino acids with molecular weight 56.79 kDa and the pI 9.45 (Table 1). Based on the CDS of \textit{NtAAP2-2}, we successfully constructed \textit{NtAAP2-2} overexpression and RNA interference vectors (Fig. S4A\textsuperscript{†}). To better understand the biological functions of \textit{NtAAP2}, we constructed the overexpression (OE) and RNAi mutant lines for the \textit{NtAAP2}. When we obtained the T1 generation transgenic lines and controls, DNA was extracted and used as templates for PCR identification. \textit{NtAAP2-2-Flag} and \textit{nptII} were used as primers for RT-PCR. The \textit{NtAAP2-2-RNAi} lines had a specific 659 bp band, and the \textit{NtAAP2-2-OE} lines had a specific 320 bp band, while these two bands were not present in the controls (Fig. S4B and D\textsuperscript{†}). Six confirmed \textit{NtAAP2-2-OE} and RNA\textit{i} transgenic T1 lines were chosen for further analysis, respectively. Moreover, the total RNA of transgenic T1 lines and controls were isolated. The first strand cDNA were synthesized and then used to analyze the expression of \textit{NtAAP2}. The expression level of \textit{NtAAP2} was dramatically increased in the \textit{NtAAP2-2-OE} transgenic plants, but significantly reduced in the \textit{NtAAP2-2-RNAi} lines (Fig. S4C and E\textsuperscript{†}).

Tobacco is an important economic crop in China. It is also widely used as a model plant in gene transformation researches owing to its mature genetic transformation system. Amino acids are essential organic nitrogen compounds in plants. The content of amino acids in tobacco leaves not only affects the growth of tobacco, but also affects the quality and flavor of tobacco. Transgenic \textit{NtAAP2-2-RNAi} and OE lines exhibited obvious phenotypic modulations compared to those of WT plants. The leaf area increased in the \textit{NtAAP2-2-OE} lines, but decreased in the \textit{NtAAP2-2-RNAi} lines Fig. 4A. To further explore the function of \textit{NtAAP2} under N stress, \textit{NtAAP2-2-RNAi}, OE lines and controls were treated with low N nutrient solution. After one week of N-deficiency treatment, \textit{NtAAP2-2-RNAi} transgenic plants showed slightly yellow leaves. While the OE lines and controls grown well. After 3 weeks of treatment, as shown in the Fig. 4B, \textit{NtAAP2-2-OE} transgenic lines still had no phenotypic changes, while the leaves of WT began to turn yellow, the plant growth and leaf growth of WT lines was

![Fig. 4](image)

**Fig. 4** The phenotype of \textit{NtAAP2-2} transgenic lines. (A) The phenotype of \textit{NtAAP2-2} transgenic lines under normal condition. (B) The phenotype of \textit{NtAAP2-2} transgenic lines under N stress.

![Fig. 5](image)

**Fig. 5** The content of amino acid in transgenic and wild type lines. (A) Amino acid content in \textit{NtAAP2-2-OE} transgenic lines and wild types under normal and N stress. (B) Amino acid content in \textit{NtAAP2-2-RNAi} transgenic lines and wild types under normal and N stress. Error bars depict standard deviation (SD). Asterisks indicate significant differences from the wild-type detected by Student’s t-test (\(P < 0.01\)).
seriously retarded, but the NtAAP2-2-OE transgenic lines still showed healthy growth (Fig. 4B), while NtAAP2-2-RNAi transgenic plants showed significant dwarfism.

As amino acid permeases are principal mediators of N uptake and distribution in plants, we evaluated amino acid content from individual transgenic lines and wild type plants under normal growth conditions and N stressed treatment, respectively. Under normal growth condition, the content of Gln increased by 20% in NtAAP2-2-OE transgenic lines than that of control plants (Fig. 5A), the levels of other amino acid were slightly changed. However, under N stress growth condition, the content of Asp in NtAAP2-2-OE transgenic lines increased by 53.3%, Asn increased by 60%, Glu increased by 26%, Gln content increased 53%. In contrast, the Gln content in NtAAP2-2-RNAi transgenic lines decreased by 64.8% under normal growth conditions, while the Glu content decreased by 27.7%. The content of Asp decreased by 33.3%, and Gln content decreased by 26.3% in N stressed NtAAP2-2-RNAi transgenic lines (Fig. 5B). The results indicated that NtAAP2 might participate in the transport of Gln and Glu. In plants, glutamate and glutamine are not only raw materials for the synthesis of other amino acids, but also involved in the synthesis of other N-containing compounds, such as chlorophyll, pyridine nucleic acid. The decrease of Gln and Glu in NtAAP2-2-RNAi transgenic lines may lead to the decrease of chlorophyll content, which further leads to plant leaves loss of green. The contents of Gln in NtAAP2-2-OE transgenic lines were increased, which could be further used to synthesize other amino acids and N-containing compounds. Under N stress, overexpression of NtAAP2-2 could provide more N sources, which enhanced plant resistance to N stress.

Conclusions

In the present study, 15 NtAAP proteins containing conserved domain Aa_trans were identified and characterized. Phylogenetic analyses revealed that the NtAAP gene family members were divided into three clusters. The exon/intron structure, transmembrane sequence and motif compositions of the NtAAP genes and proteins are highly conserved in each class, indicating that the function was conserved. An additional comprehensive analysis of the expression profiles revealed that most NtAAP genes were expressed in different tissues under normal growth conditions. We successfully cloned the NtAAP2-2 gene from tobacco, and the overexpression vectors and the RNAi vector of NtAAP2-2 were constructed. In transgenic tobacco, the contents of Asp, Asn, Glu and Gln were changed, indicating that NtAAP2-2 was involved in the transport of Asp, Asn, Glu and Gln in tobacco plants. Thus, results of the present study provide the fundamental information for future functional investigations of this gene family.

Acknowledgements

This work was supported by the Science Project of the Zhengzhou Tobacco Research Institute (902013CZ0620). We also thank Dr John Hugh Snyder and Fangfei Liu for their kind help in discussion, writing and revising the manuscript.

References


