

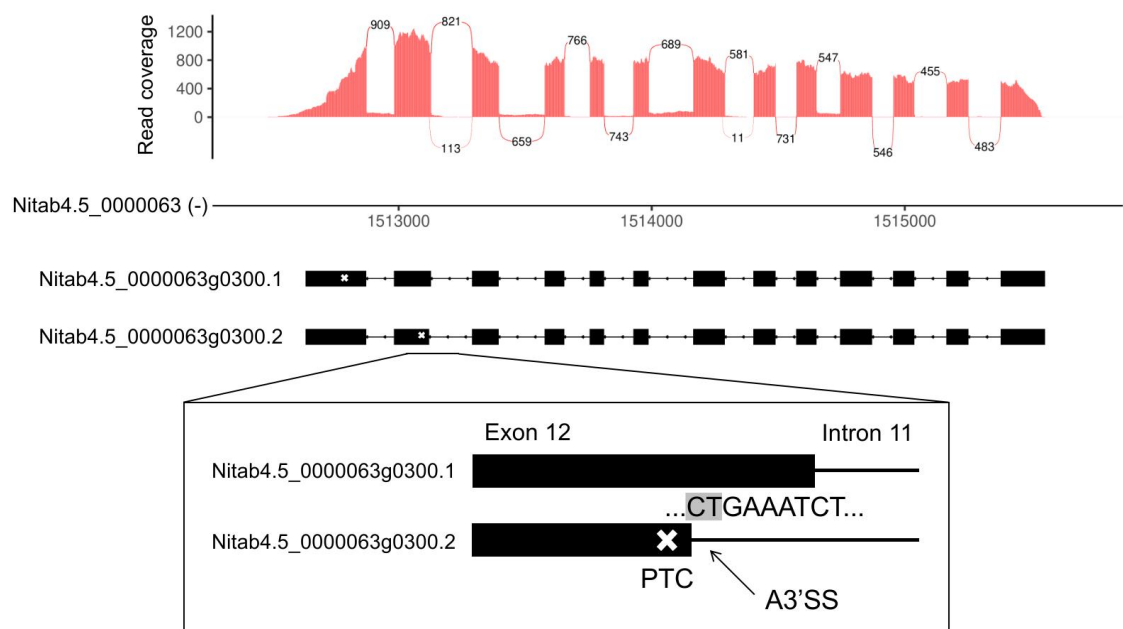
## Supplementary Figures

```

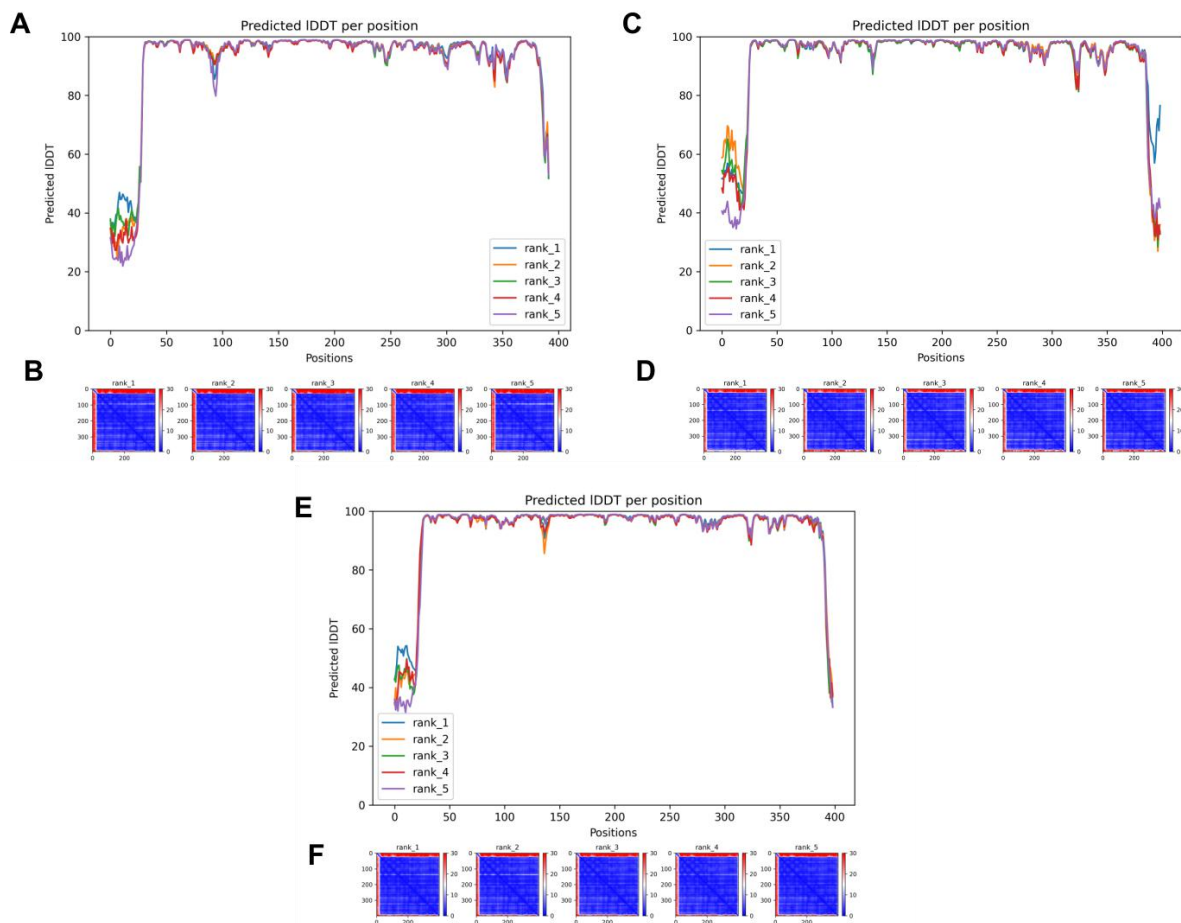
1   ctttacagcttttcaaactcttgttcaacttcttctgctatagaaagagagaaagacaaat   61
62   agcccttaacttattttattgcttctggctgttagaatggagaaagtgatggattgtaac 121
                                     M E K V M D C N
122  tcaatttctcatggctgcagtgcttttttctctactgtgtcctttggacaacaactgaagga 181
      S I L M A A V L F S L L C L W T T T E G
182  gacctcattgttaatatataacctacctcgagactgccatatctaagggtgcagtatgcttg 241
      D L I V N I T Y L E T A I S K G A V C L
242  gatggaagtgcctccagctttccactttgatcctggaactgggttcaggaaatttagctgg 301
      D G S A P A F H F D P G T G S G I Y S W
302  cttattcaccttcagggtggaggatgggtgtgaaaccatttctaactgccaatatcggtgca 361
      L I H L Q G G G W C E T I S N C Q Y R A
362  cttacggacttatcttcaaaaaatgcaaaatatatcttattttactggagttttaagt 421
      L T D L S S K N M I S Y F T G V L S
422  agcgatcctcaacttaatccagacttctacaactggaacagagtcagggttaagtattgt 481
      S D P Q L N P D F Y N W N R V K V K Y C
482  gatgggtcatcgatactggcgatattgaagaagttgatcccgatactaacctccacttt 541
      D G S S Y T G D I E E V D P D T N L H F
542  agaggagcaagaattttgaagctatcatggagtatttttaggtcaagggtgattttat 601
      R G A R I F E A I M E Y F L G Q G M I Y
602  gctcaaaatgccatcctcagtggaacttcagcaggagggttgggtgcaatcatacttgt 661
      A Q N A I L S G T S A G G L G A I I H C
662  gataagttcagactgtttctcccatctaactgctagagtgaatgcattctcggtgctggt 721
      D K F R L F L P L T A R V K C I S D A G
722  tctttgtcgacttagaaacaatatctggcgaaccacttattgaggaatgtataaaaga 781
      F F V D L E T I S G E P L I E E M Y K R
782  gttggtgctttgcacggatctgccaaacactaccactctcttgactgccttgagctca 841
      V V A L H G S A K N L P L S C T A L S S
842  gacccaaatttgtgcttcttccctcagtatgtagcaagacataatttgacaccacttttc 901
      D P N L C F F P Q Y V A R H I C T P L F
902  atcgtaactcggcctatgattcttggcagataaacaacagtttggttctctcgagttgct 961
      I V N S A Y D S W Q I N N S L V P R V A
962  gatcctgacgatgagtgacgtatttgaagaagatataaatgtatgctcaccaagtcag 1021
      D P D D E W T Y C K K D I N V C S P S Q
1022 atccaaactctccaagatttcaggttgacatttttgaggcattgaacgaactaggacca 1081
      I Q T L Q D F R L T F L E A L N E L G P
1082 agtacatcaaggatatttcatcttctgtcattcccaccacggcattgaaatacaa 1141
      S T S R G Y F I S S C H S H H G I E I Q
1142 aattattggtcctccaaaactctccaacttagataacaagacgatcgctgaagctgtt 1201
      N Y W S S K N S P T L D N K T I A E A V
1202 ggagattggttttcgaccgaagtgggtccaagaagtttattgcccttatcccttgggc 1261
      G D W F F D R S G F Q E V Y C P Y P C G
1262 aaattttgacagtagagttttctcattttgtactactatttaagattatTTTTTTTatta 1321
      K F C Q *
1322 ttatgtttttttacaagtggagtatctagcaatgcaagaagttagttagctctagtatct 1381
1382 tagatatttgttatccaactactcatgtaattgcttaacattgttcgtcaaacttgcatt 1441
1442 tcgtttttcacgtttttaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1485

```

**Figure S1. cDNA sequence of *NtPAE1* and corresponding predicted amino acid sequence.** The numbers represent nucleotide positions in the sequence (1 through 1485). The CDS spans nucleotides 98-1276, producing a predicted protein of 392 amino acids. The start codon ATG and the stop codon TAG are highlighted with underlined nucleotides. The region spanning nucleotides 98-897 (bold and italic highlight) corresponds to the region cloned to produce RNAi transgenic plants. In the amino acid sequence, underlined residues 1-28 correspond to the predicted signal peptide of a secretory pathway substrate cleaved by SPase I (Sec/SPI) (SignalP6 prediction with probability of 0.9998). The red arrow indicates the predicted cleavage site of SPaseI between positions G28 and D29 (SignalP6 prediction with probability of 0.9821). In grey highlight (residues 29-376) is the pectin acetyltransferase (PAE) domain (PFAM accession pfam03283), CD-search result with e-value of 0.0. A predicted phosphorylation of Serine 356 is highlighted in dark grey (MusiteDeep score 0.634). The predicted active site triad is highlighted in black (amino acids S178, D275 and H342).



**Figure S2. Transcripts of NtPAE1 in the *N. tabacum* genome.** RNA-seq coverage of region Nitab4.5\_0000063:1512267-1515863 (antisense strand), showing read coverage and exon/intron structure of transcripts from gene Nitab4.5\_0000063g0300 (NtPAE1). Transcript Nitab4.5\_0000063g0300.1 is the original annotated transcript from Sol Genomics Network database, while Nitab4.5\_0000063g0300.2 was manually annotated by predictions made from publicly available RNA-seq data from flowers of *N. tabacum*. Transcript Nitab4.5\_0000063g0300.2 is a result of an alternative 3' splice site (A3'SS, black arrow and grey highlight in the zoomed box) in exon 12, and this event is supported by 113 read junctions. Additionally, the same event is present in *NtPAE1* homologue in *N. tomentosiformis* (gene LOC104098660), which supports our finding. This alternative splicing event creates a premature stop codon (PTC - white "X" in the zoomed box) in exon 12, which could either create a truncated PAE domain in the protein product, or direct the transcribed RNA to the Nonsense Mediated Decay pathway. The graph shows the number of reads mapped to each region of the transcript, while connecting curves represent exon-exon junctions and the number of reads that support them. In the scheme of transcripts, boxes represent exons and lines represent introns.



**Figure S3. Structure prediction metrics for NtPAE1 (A,B), PvPAE8 (C,D), and VrPAE1 (E,F).** A, C, and E: Predicted Local Distance Difference Test (pLDDT) values per amino acid position for five predicted models (rank\_1-5). B, D, and F: Predicted Aligned Error (PAE) for the five predicted models. y axis = aligned residue, x axis = scored residue. Gradient scale = PAE in Ångströms.

