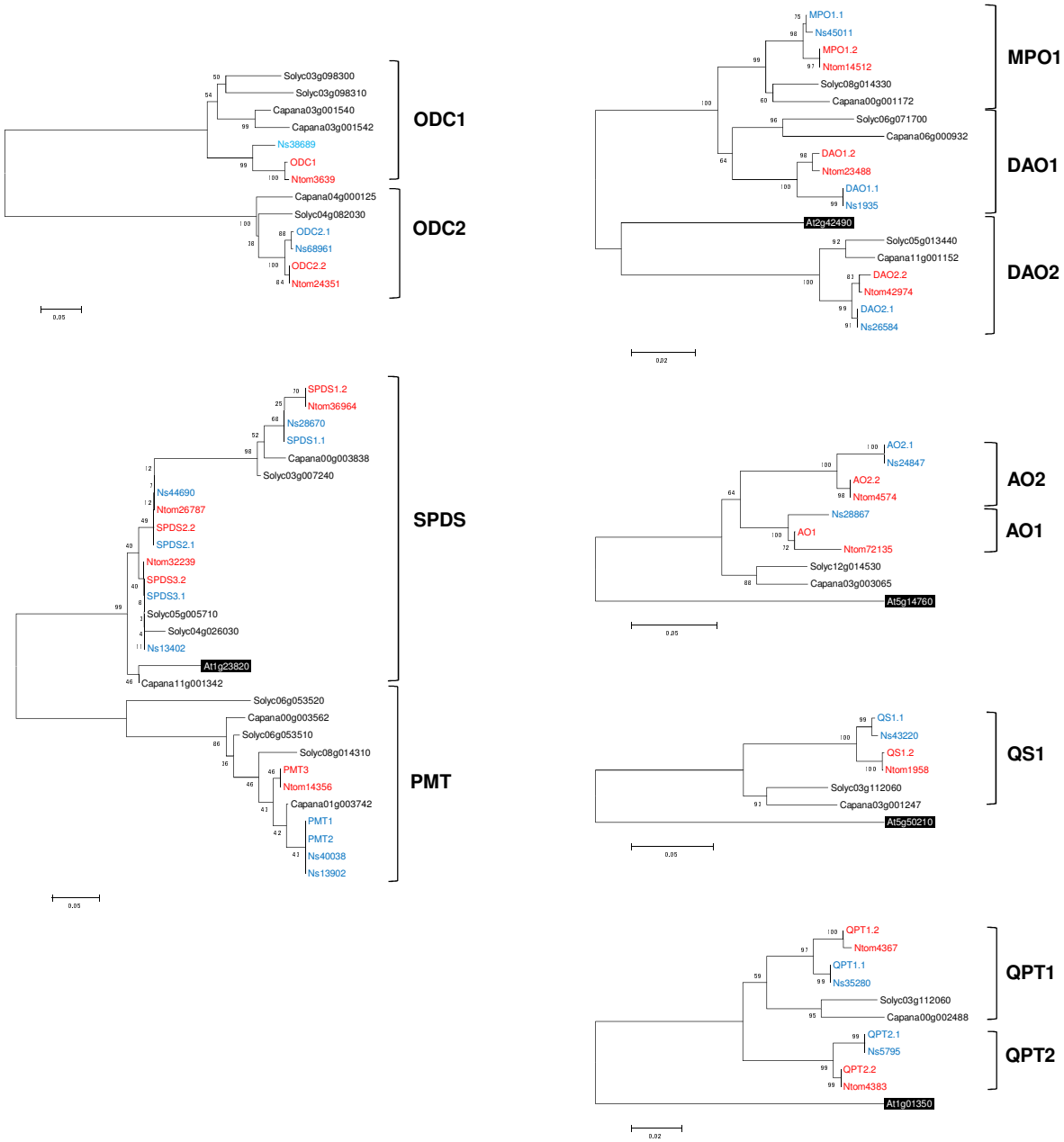
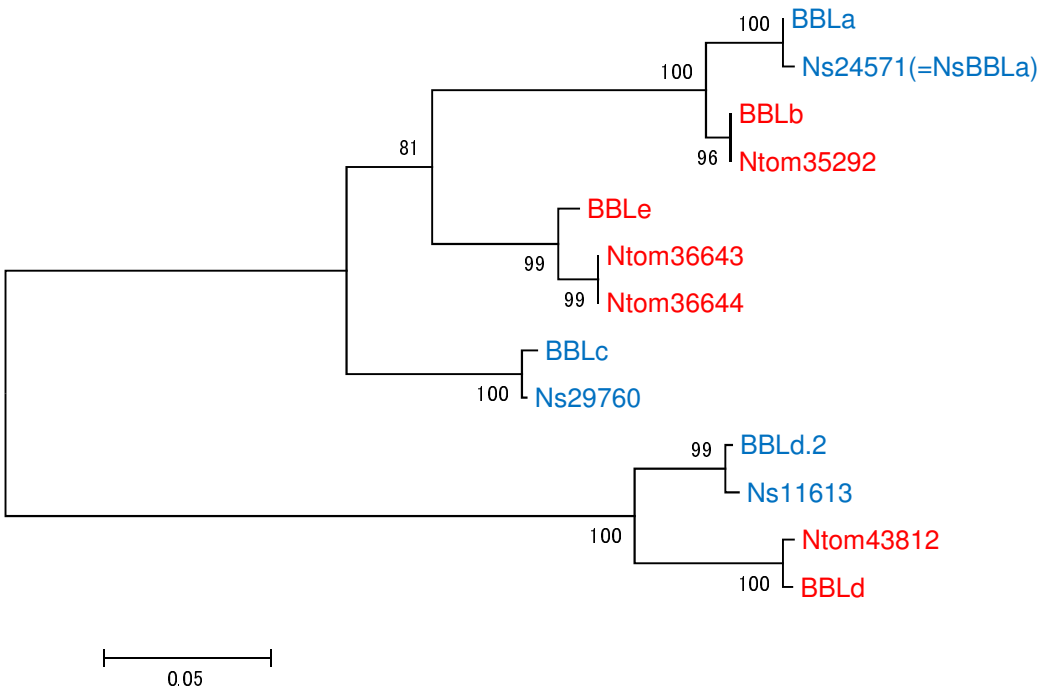
Figure S1



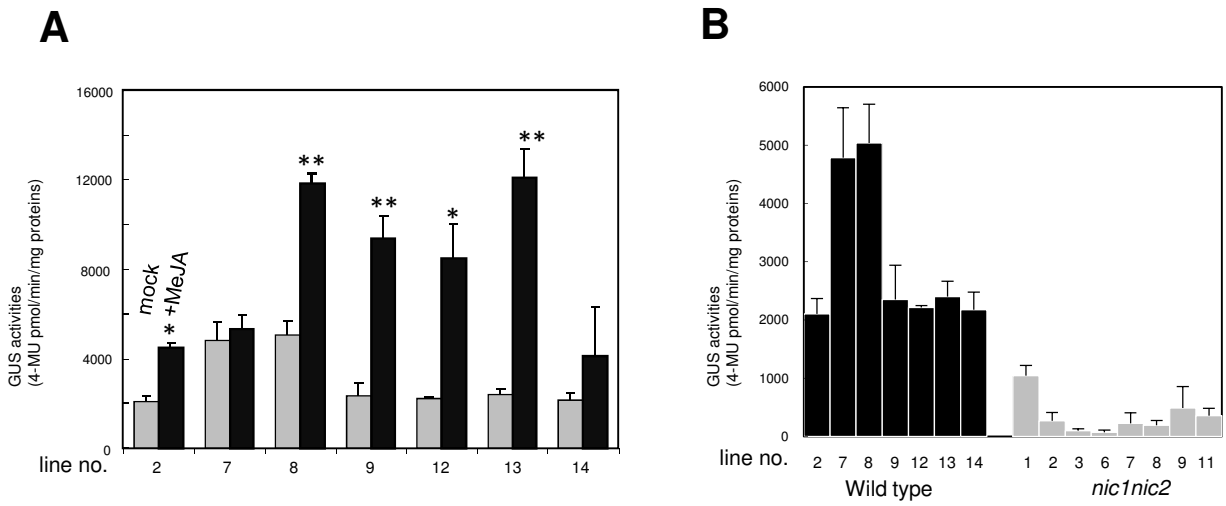
Supplemental Figure S1 Phylogenetic trees of enzyme proteins involved in nicotine and related primary metabolism. Bootstrap values are indicated at branch nodes and the scale bar indicates the number of amino acid substitutions per site. Proteins from *N. sylvestris* (blue) and *N. tomentosiformis* (red) are represented with gene no. denoted with prefixes, Ns and Ntom, respectively. Tobacco proteins (**Supplemental Table S1**) are colored same with their ancestral counterparts. Proteins from tomato (Solyc), pepper (Capana), and Arabidopsis (At, in black box), are included.

Figure S2



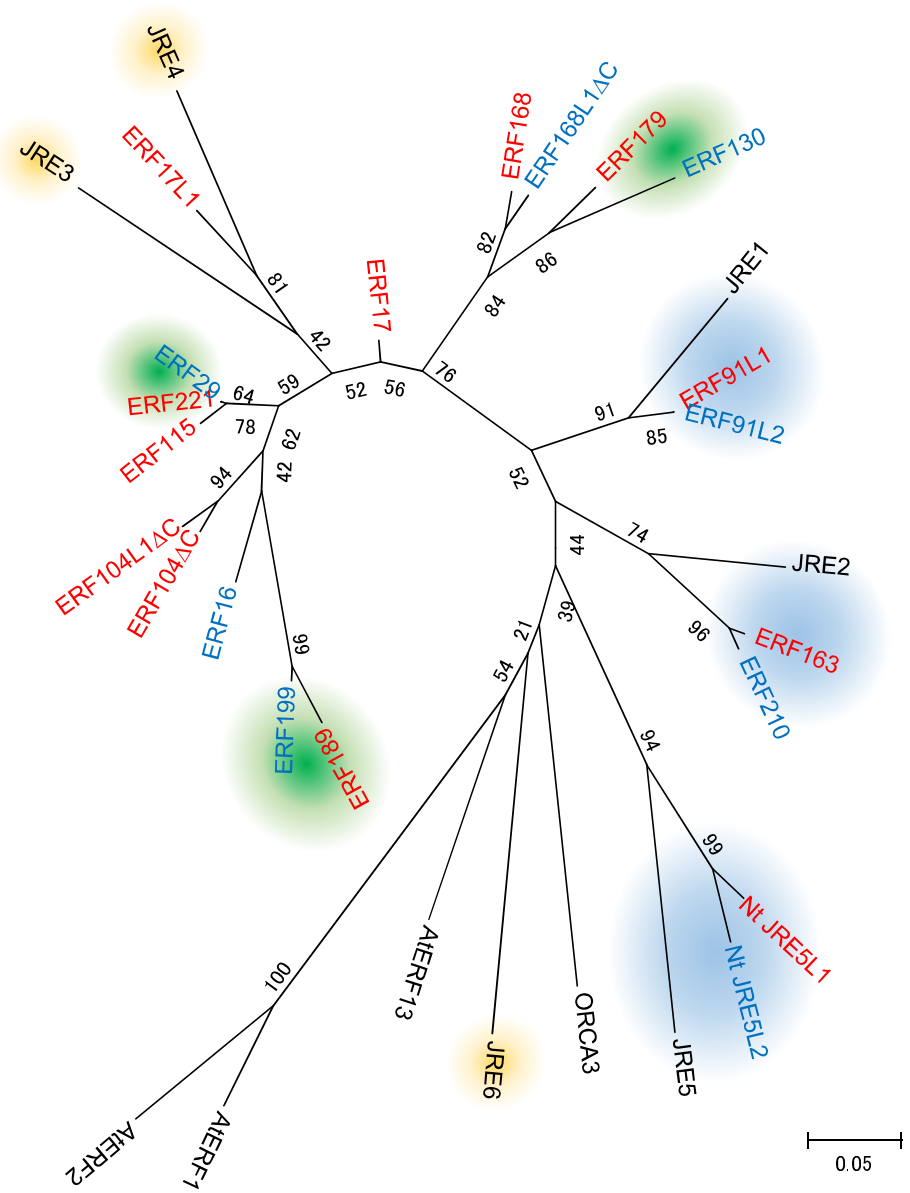
Supplemental Figure S2 A phylogenetic tree of BBL proteins from tobacco and its ancestral *Nicotiana* diploids. Bootstrap values are indicated at branch nodes and the scale bar indicates the number of amino acid substitutions per site. Names of BBLs from *N. sylvestris* (blue) and *N. tomentosiformis* (red) are represented with gene no. denoted with prefixes, Ns and Ntom, respectively. Tobacco BBLs, except BBLd (Kajikawa et al. 2009) that could not be retrieved from the genomic database, are listed in **Supplemental Table S1**, and are colored same with ancestral counterparts.

Figure S3



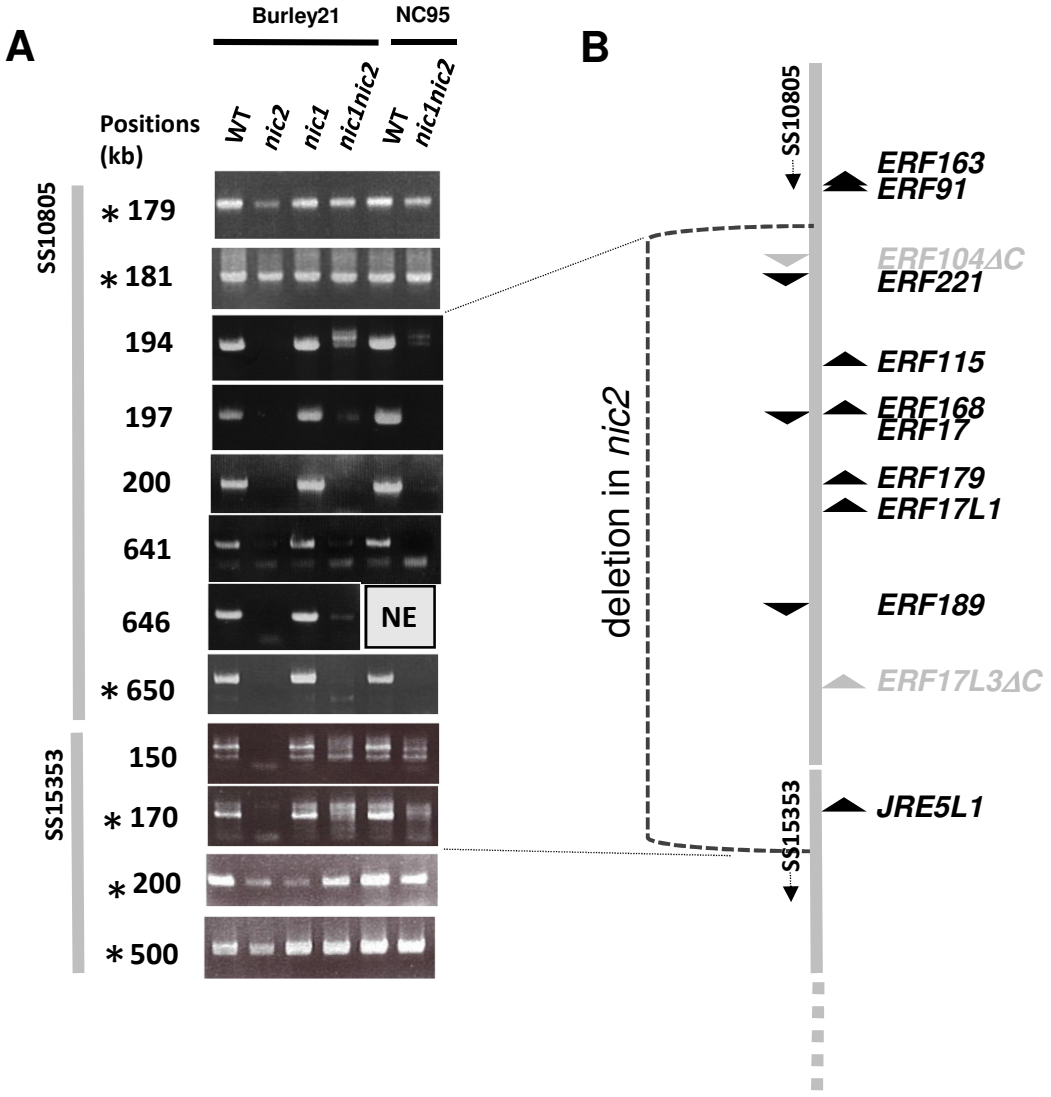
Supplemental Figure S3 GUS activities in tobacco hairy roots transformed with the reporter gene driven by *NsBBLa* promoter.
The *GUS* reporter gene was driven by a promoter region from -1,162 to -1 (numbered from the first ATG) of *NsBBLa* gene from *N. sylvestris*. The error bars indicate standard deviations for more than three biological replications. A, Activities in transgenic tobacco hairy roots treated with 100 μM MeJA for 24 h. Significant differences were determined by Student's t-test. $^{**}P<0.01$, $^{*}P<0.05$. B, Activities in transgenic tobacco hairy roots of wild-type and *nic1nic2* mutant.

Figure S4



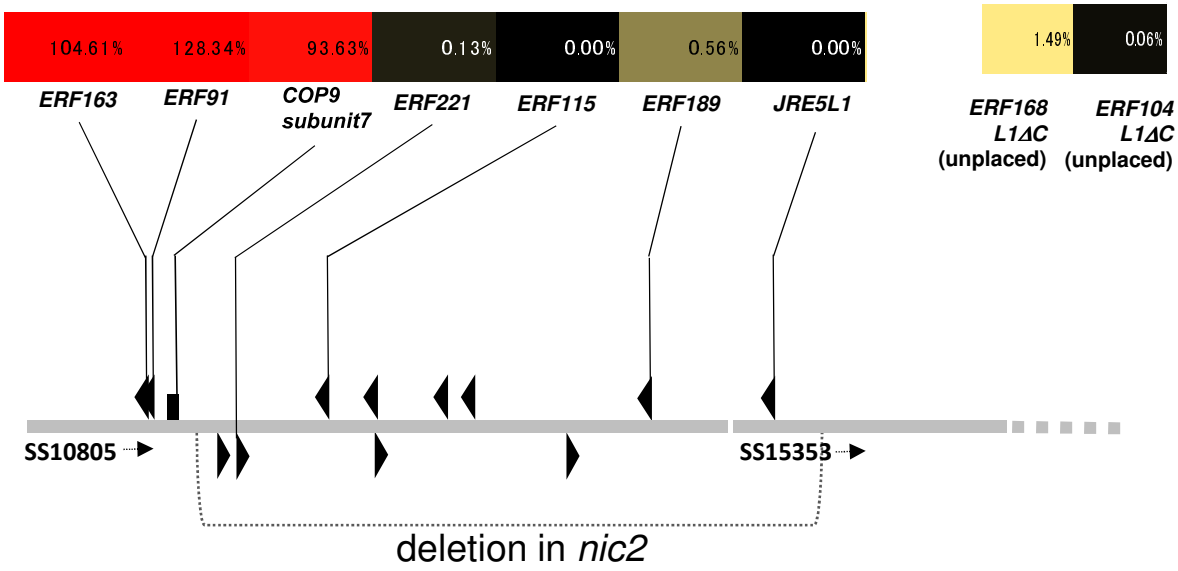
Supplemental Figure S4 A phylogenetic tree of *NIC2*-locus ERFs and related proteins. Bootstrap values are indicated at branch nodes and the scale bar indicates the number of amino acid substitutions per site. All ERFs from tobacco listed in **Supplemental Table S2** except ERF17L2 ΔC and ERF17L3ΔN, which are shorter than 120 amino acid residues, JREs from tomato (Thagun et al. 2016), ORCA3 from *Catharanthus roseus* (AJ251249), and AtERF1 (At4g17500), AtERF2 (At5g47220), and AtERF13 (At2g44840) from *Arabidopsis* are included. Tobacco ERFs presumably originated from *N. tomentosiformis* and *N. sylvestris* are shown in red and blue, respectively. The groups consisted with orthologous members are shown with colored backgrounds. The groups common in tomato, tobacco, and two ancestral diploids of tobacco are shown in blue, whereas those specific to the *Nicotiana* species and tomato are in green and yellow, respectively.

Figure S5



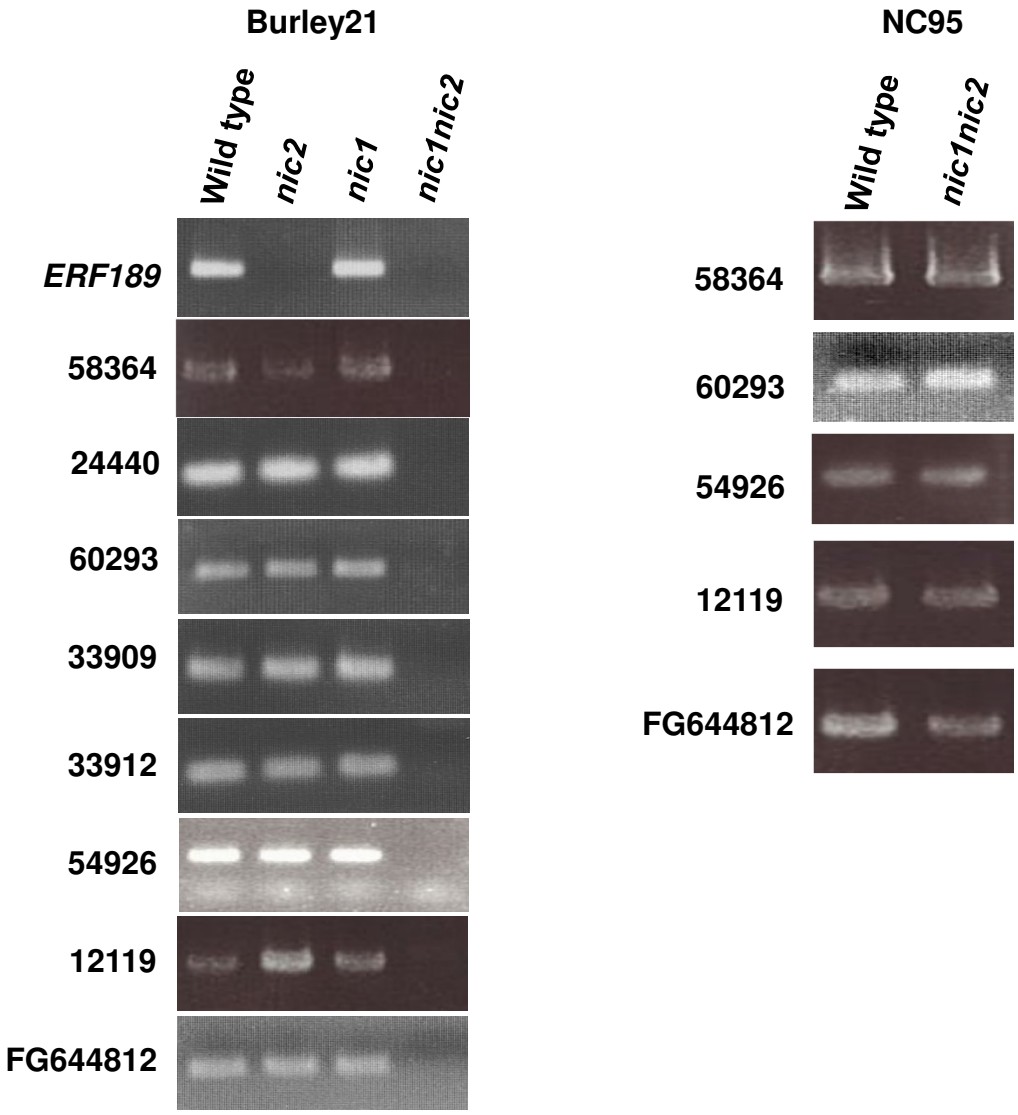
Supplemental Figure S5 A chromosomal region deleted in *nic2* mutant was delimitated with genomic PCR analysis. A, Genomic PCR analysis with primers targeted to the indicated genomic regions (counted from a one end of SS in kb in the directions shown with arrows in B) on SS10805 (664 kb) and SS15353 (515 kb). Genomic DNAs from different genotypes in Burley 21 and NC95 cultivars were used as templates. Specific amplifications were confirmed by sequencing the amplified fragments for those denoted with asterisks. B, A large part of the gene cluster indicated are deleted in *nic2* mutant. Arrowheads indicate positions and orientations of predicted open reading frames of *ERF*s. *ERF* genes, denoted with Δ and shown in gray, encode possibly non-functional proteins that don't include full-length DNA-binding domains. NE; not examined.

Figure S6



Supplemental Figure S6 Transcript levels of genes around *NIC2* locus in the roots of wild-type and *nic2* mutant tobacco were analyzed by qRT-PCR. The levels in *nic2* relative to wild type are shown in percentages (%). Average values for three biological replicates were adapted. Even for genes within the deleted region, trace amounts (less than 2 %) of transcripts were detected, but such detections is considered false and possibly caused by background amplification of non-specific targets. The levels of two *ERF* genes, *ERF168L1ΔC* and *ERF104L1ΔC*, on unplaced SS3881 (**Supplemental Table S2**) were analyzed as well.

Figure S7



Supplemental Figure S7 Genes deleted in LA Burley 21 of *nic1nic2* genotype were found by genomic PCR analysis. The genes are listed in **Table 1**. *NIC2*-locus *ERF189* deleted in *nic2* mutant was also analyzed. Primers used are listed in **Supplemental Table S5**.