

Table S1 Metabolic enzyme and transporter genes involved in nicotine and related metabolism in tobacco. Metabolic enzymes are abbreviated as in **Figure 1**. Same SSs appearing multiple times are indicated with numbered asterisks in brackets. A phylogenetic tree for BBLs supporting the gene naming is shown in **Supplemental Figure S3**. NA; not available, chr.; chromosome, SS; super scaffold. T; *N. tomentosiformis*, S; *N. sylvestris*, MATE; multidrug and toxic compound extrusion

protein	gene	gene no.	origin	chr. no.	no.	Super Scaffold (SS)		orientation
						start position on chr.	end position on chr.	
ODC	<i>ODC1</i>	53859	T	4	3473	NA	NA	NA
	<i>ODC2.1</i>	59117	S	NA	4493	NA	NA	NA
	<i>ODC2.2</i>	36820	T	12	17585	3053449	3401286	+
PMT	<i>PMT1</i>	51171	S	5	293	NA	NA	NA
	<i>PMT2</i>	11501	S	5	12228	NA	NA	NA
	<i>PMT3</i>	64369	T	15	5496	24558026	25158092	+
SPDS	<i>SPDS1.1</i>	85439	S	13	9517	65137943	66066006	-
	<i>SPDS1.2</i>	66703	T	4	5927	8070558	8503316	-
	<i>SPDS2.1</i>	31582	S	20	16761	11140291	11479445	+
	<i>SPDS2.2</i>	38827	T	24	18021	73714954	74588983	-
	<i>SPDS3.1</i>	55613	S	16	3829	43241401	44313357	-
	<i>SPDS3.2</i>	7022	T	12	11406	NA	NA	NA
MPO	<i>MPO1.1</i>	67187	S	NA	6051	NA	NA	NA
	<i>MPO1.2</i>	3417	T	24	10683	NA	NA	NA
DAO	<i>DAO1.1</i>	62765	S	8	517	73159306	74639532	+
	<i>DAO1.2</i>	7666	T	22	11487	41494549	42231281	-
	<i>DAO2.1</i>	75226	S	20	7543	16088260	17549898	-
	<i>DAO2.2</i>	24485	T	24	1540	69072361	69815524	-
AO	<i>AO1</i>	57190	T	21	4133	51719172	52188197	+
	<i>AO2.1</i>	19078	S	7	147079	NA	NA	NA
	<i>AO2.2</i>	71591	T	21	6835	NA	NA	NA
QS	<i>QS1.1</i>	54977	S	6	371	79139200	79805586	+
	<i>QS1.2</i>	27785	T	4	1625	72847380	73170704	-
QPT	<i>QPT1.1</i>	26947	S	NA	160281	NA	NA	NA
	<i>QPT1.2</i>	18160-2	T	NA	1382(*1)	NA	NA	NA
	<i>QPT2.1</i>	74444	S	NA	74055	NA	NA	NA
	<i>QPT2.2</i>	18160	T	NA	1382(*1)	NA	NA	NA
A622	<i>A622</i>	75882	S	NA	7679	NA	NA	NA
	<i>A622L</i>	75125	T	12	753(*2)	57347634	58008935	-
BBL	<i>BBLa</i>	36833	S	22	17588	73071000	73701890	+
	<i>BBLb</i>	44985	T	19	2106	NA	NA	NA
	<i>BBLc</i>	72778	S	NA	7078	NA	NA	NA
	<i>BBLd.2</i>	75618	T	NA	7622	NA	NA	NA
	<i>BBLe</i>	8212	T	19	11571	NA	NA	NA
MATE	<i>MATE1</i>	13706	S	16	12706	45652827	45980361	+
	<i>MATE2</i>	75126	T	12	753(*2)	57347634	58008935	-
NUP	<i>NUP 1</i>	690	T	24	10130	NA	NA	NA
	<i>NUP2</i>	81472	S	16	8708	NA	NA	NA

Table S2 Primer sequences for qRT-PCR analysis. These primers were used for the experiments shown in **Figure 2** and **Supplemental Figure S6**.

gene	direction	sequence
<i>ODC1</i>	F	AATCTTGTCAAATATGTGAGTC
	R	ATGTTATAAAACACTATCAG
<i>ODC2</i>	F	ATCGTAGTAACGTTACCTGCG
	R	AACCTGCAGCTCCGGTAACTG
<i>SPDS2</i>	F	AGGAGGMGATGGTGGTGTCTTGCG
	R	AGMAACMTCAAYTACCATCTTG
<i>PMT</i>	F	GAGGAATTGGTTTTACATTATTCG
	R	CTRGATACATCAACTACCACGTC
<i>DAO1</i>	F	ATCAGCAGAAACAAACGCCG
	R	TGGCATGTATGAGCTCTTGG
<i>MPO1</i>	F	CYATCCTCCGTCGTGAGGCACAG
	R	ACGCTAGAGCTGATGAGGCCGTC
<i>AO1</i>	F	ATTCTTGCCAGAGAGATGGC
	R	TTCCATGTTTGGCAACTTCAAGAG
<i>AO2</i>	F	GAGATGGATCAGTTACTAGGTAC
	R	CCACAGTTCCATGCTTGGCAACC
<i>QS1.1</i>	F	GATTCAGTTTGCAAACCCCAAT
	R	AAGCGGAGCTTGATCGTTG
<i>QS1.2</i>	F	GATTCAGTTTGCGAACCCCAAG
	R	GAGCTTGATCGTCCATTATCACG
<i>QPT1</i>	F	CAATAGCCACCAAAAATGCAG
	R	AGAGGGCAAGTTGAATAACA
<i>QPT2</i>	F	CCACAAAATCATTTCACARCC
	R	TTTTCACCACCAACCTTGGA
<i>ERF189</i>	F	GCAGCTTCGACTGCAGCTTCCTC
	R	CTCCTCGGACTCGGAGCACTTC
<i>ERF199</i>	F	TTAGCAGCTTCGACTTCGAC
	R	CTCGGAGTACTTTTCATGGG
<i>ERF16</i>	F	GGCGGCCATGGGGGACGTTTTTCAG
	R	GATTTTGAAAGCGGCTTGATCGTAC
<i>ERF29</i>	F	ATTCGATGTTTAGTATTAGAATCCC
	R	ATTATCATGATCGAACAGTACTAC
<i>ERF115</i>	F	CAAGTTTTGGCAGCTTCCCTTCG
	R	GGACTCGGACTTTTCTTCAAGTTG
<i>ERF221</i>	F	GAACCTACAAATGTTGCTATAACTG
	R	GACAGAGAGATGTTTACAAGTTTCG
<i>ERF163</i>	F	CTTGTAATGAGAAATACATAAAGTAG
	R	TATTAAATTGGATAAAACAGTAGAATAG
<i>ERF91</i>	F	TACAAGGAGGAAACAAAGAAGAACC
	R	TATCTCCGCCGCAAATTTACCCACG
<i>COP9 sununit7</i>	F	AAGAAATTGTTATCCATCTACAGC
	R	TTTGCTTTTGGAGTTGGTAAGATCTTC
<i>JRE5L1 ΔC</i>	F	CACCATATTTGCACAAAACATGG
	R	TGGCTCCTTATCATTTTCTGACGAAGTC
<i>ERF104L1 ΔC</i>	F	GGCGGCCATGGGGGACGTTTTTCAC
	R	CGGAGCCAATGCTGCATCCTCAGGAG
<i>ERF168L1 ΔC</i>	F	ACCCCGACAGGAGAGGTGCCAGAC
	R	GCGGCTTGATCGTAAGCGTTAC
<i>EF1a</i>	F	AAGCCCATGGTTGTTGAGAC
	R	GTCAACGTTCTTGATAACAC

Table S3 *NIC2* -locus *ERF* genes and their homologs in tobacco. Clustering of the genes were schematically shown in **Figure 5**. Unplaced SSs are marked with asterisks. The origin of *ERF221* marked with double asterisk was presumed based

<u>gene name</u>	<u>gene no.</u>	<u>mRNA no</u>	<u>SS no.</u>	<u>origin</u>
<u>genes on Chr. 19</u>				
<i>ERF163</i>	NA	NA	10805	T
<i>ERF91L1</i>	4016	8541	10805	T
<i>ERF104 ΔC</i>	NA	NA	10805	T
<i>ERF221</i>	NA	NA	10805	T**
<i>ERF115</i>	NA	NA	10805	T
<i>ERF168</i>	NA	NA	10805	T
<i>ERF17</i>	NA	NA	10805	T
<i>ERF179</i>	NA	NA	10805	T
<i>ERF17L1</i>	4018	8543	10805	T
<i>ERF189</i>	4020	8545	10805	T
<i>ERF17L2 ΔC</i>	NA	NA	10805	T
<i>JRE5L1</i>	NA	NA	15353	T
<i>ERF104L1 ΔC</i>	NA	NA	3881*	T
<i>ERF168L1 ΔC</i>	55871	119118	3881*	T
<u>genes on Chr. 7</u>				
<i>ERF199</i>	52368	111850	3181	S
<i>ERF17L3 ΔN</i>	52369	111851	3181	S
<i>ERF210</i>	ND	NA	3181	S
<i>ERF130</i>	ND	NA	3181	S
<i>ERF16</i>	52370	111852	3181	S
<i>ERF29</i>	ND	NA	3181	S
<i>ERF91L2</i>	ND	NA	6062*	S
<i>JRE5L2</i>	16863	36133	134732	S

Table S4 Primer sequences for genomic PCR analysis to detect the chromosomal delation around *NIC2* locus in the mutants. These primers were used for the experments shown in **Supplemental Figure S5**.

primer target				
SS name	position on SS (kb)	primer name	sequence	
10805	179	179F	GATATAAATCTCAGCCAACCTCTCCG	
		180R	TCCTTGGTTTGGCAGGGTATTACCG	
	181	181F	CCAACTCCCTGGCCATACTGGTCCG	
		182R	ATGGGAAGACCGATGATCAGGTACC	
	193	193F	TTGTTGTTGCAGCAGTGGAAAATGG	
		193R2	TCATACTTAAACATTTGAATGAGGC	
	196	196F	GAAACCCTGTTCCCGGAACTGTTGC	
		197R2	GTTTCAGTTAAGGTGAAACTCAGTGC	
	200	200F	GGCATCTGGTAAGAAAATCTAGACC	
		200R	GTCCAAGTGAAAGGGGCCATCGC	
	641	641F	TTGCTAATTATGTAGACATTTGC	
		641R	TTCATGCACATGTTTCTATCAGCTC	
	646	646F	CTAAAAATATAACACGAACTCACTCG	
		646R	TGTGGTCACAAGGACATATATGACG	
	650	650F	CGGAAAAGATTGTACATATTTTAGAG	
		650R	TGCGATCACACCAGCAACCCAGCTGC	
	15353	150	150F	GTGAAGGTCGTTTACGCGAGGTGACG
			150R	CTCGCTTTTGTGAGCCAAAATGCG
170		170F	AGCTCCTCTTCGCGAACGTGATGG	
		170R	ACGTCTCACCTGTCGTGATGGTGC	
200		200F	GTGTGGTTCTAGGCTTATCCGC	
		201R	TCAGATTTGATTTCTAAAGTTGC	
500		500F	TCTTCCCACAAGAAGGTCACGG	
		501R	CCTCTCAACCCGCTTCTAATTGC	

Table S5 Genes found to be deleted in LA Burley21. Supporting data are in **Supplemental Fig. S7**. Deletions of genes of which no. are marked with asterisks were assumed based on the results for the other genes on the same SS. The descriptions in the right three columns are based on previous microarray data (Shoji et al., 2010). Relative positions on chromosomes are estimated on nucleotide counts and presented in percentages (%). ND; not defined.

chr. no.	SS no.	chromosomal position (%)	gene no.	description	EST accession	ranking	signal ratio (<i>nic1nic2</i> to WT)
10	438	60.9	58364	GroES-like family protein	EH613852	19	0.084
17	15390	0.72	24440	alpha/beta-hydrolases superfamily protein	DW004070	4	0.015
		0.81	24438*	C2H2 and C2HC zinc fingers superfamily protein	DV999683	25	0.1
	4711	1.91	60293	unknown	EB443175	1	0.013
		1.94	602868	adenine nucleotide alpha hydrolases-like superfamily prote	EH618232	30	0.111
	17049	2.05	33909	alpha/beta-Hydrolases superfamily protein	EB445921	5	0.021
		2.11	33912	Unknown	TA19154_4097	2	0.015
2.25		33908*	zinc finger (CCCH type) helicase family protein	DW004917	38	0.133	
2.29		33905*	unknown	TA12325_4097	16	0.062	
19	3704	92.4	54926	basic region/leucine zipper motif 53	FG635137	7	0.021
		92.4	54927*	secretion-associated RAS super family 2	EH618182	6	0.021
	1234	77.4	12119	light harvesting complex photosystem II subunit 6	TA12004_4097	14	0.062
ND	ND	ND	ND	ND	FG644812	10	0.028

Table S6 Primer sequences for genomic PCR analysis to detect genes deleted in LA Burley21. These primers were used for the experiments shown in **Supplemental Figure S7**.

gene ID (name/no./accession)	forward or reverse (F or R)	sequence
<i>ERF189</i>	F	GGGCAATGGAAATGAATCTAGC
	R	CTTCCTTCCTTTCACATAG
58364	F	CTCTGACATCAACGCTTATGAG
	R	CTCAACACTTCATAACATCCAC
24440	F	CCTTGCCTAATGGAGGAGATAACG
	R	CCTCTCAAAGGTAGATAAACTCCC
60293	F	TCTACCACCAAATGCCAAG
	R	GCTTTGGGGATAGAGCGGAGC
33909	F	GGATTCCAAAAAATGGTGCC
	R	CATCCCATTCAGATTTCCC
33912	F	GGATTCTTCCTAAAGGGGTGACAG
	R	AACTTGGGCATTCCACAAAGC
54926	F	GCGATTGGGAGAATTGATGAGC
	R	GAAACCAAAGCTACATTGTCC
12119	F	GCTTGAGAGACTGAAGGTTG
	R	CTACCGTATCACTGCCTTTG
FG644812	F	CTACTGGAGGAGGAAGAGGAAAGG
	R	TGACCCTGTGCCATAACCATAGAC