

SUPPLEMENTARY INFORMATION

This file contains the Supplementary Figures and Supplementary Tables corresponding to the manuscript: **“Synthesis of crocin I and crocin II by multigene stacking in *Nicotiana benthamiana*”**

Lei Xie¹, Zuliang Luo¹, Xunli Jia¹, Changming Mo², Xiyang Huang³, Yaran Suo⁴, Shengrong Cui¹, Yimei Zang¹, Jingjing Liao^{2*} & Xiaojun Ma^{1*}

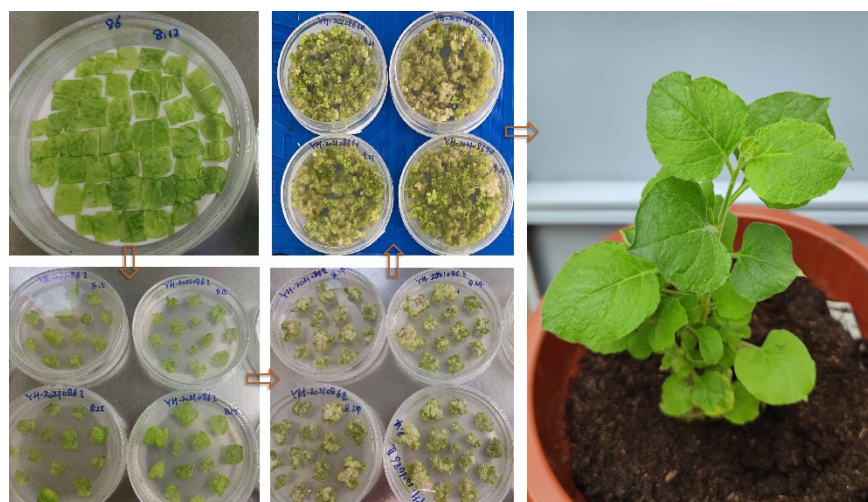


Figure S1. Process of agrobacterium tumefaciens-mediated transformation of multi-gene vector AU-CU into *Nicotiana benthamiana*.

TGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTC
 ACTTTATTGTGAAGATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGG
 AAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAG
 GAGCATCGTGGAAGAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATC
 TCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCCTCCTCTATATAAG
 GAAGTTTCATTTCAATTTGGAGAGAACAATGGCCGTCCAAAGCAACGGCAATGGAAGCTCGG
 ATTCTCATGTTAAGGTTCCCTCAGATAAAGTTACCAAGCTCTTCATCAATGGAGAGTT
 CGTTGATTCTGTTTCAGGAAAGACGTTTGAGACAATAGATCCAAGAAATGGGGAGGT
 GATTGCAGAAATTGCCGAGGGGGGCAAGGAAGATGTTGATTGGCTGTCAAAGCTGC
 CCGAGAAGCTTTTGACCACGGCCCATGGCCTCGTTTACCTGGCTCTGAGAGGCGAAA
 GATAATGATGAAGTTTGCAGACTTAATAGATGAAAATACCCAAGAATTGGCCACCTT
 GGATGCAATTGATGCTGGAAAATTGTTTTACCTTTGCAAGATTATGGATATTCCAGGA
 GCAGCAGAAACAATTGCTATTATGCGGGTGCGGCAGATAAGATTTCATGGGGAGACT
 CTAATAATGTCAACCGCGCTACAAGGATACACATTGCATGAGCCCATTTGGAGTTGTT
 GGACACATCATTCCCTGGAACCTCCCCAGTCAAATGTTTGCTATGAAGGTTGGCCCTG
 CATTAGCAGCTGGCTGCACCATGGTAGTCAAGCCTGCTGAACAGACGCCTCTGTCAG
 CCCTCTATTACGCTCATTTAGCAAAGCTGGCTGGAATCCCCGATGGAGTGCTAAACGT

TGTAACAGGATTCGGACAAACAGCTGGTGCGGCAATTAGCTCTCATATGGACATTGA
TATGGTAAGTTTCACGGGTTCTACAGAAGTAGGGCGTCTCGTCATGCAAGCTGCAGC
AAACAGCAACTTAAAGCCTGTGTGTCTAGAACTGGGAGGCAAGTCCCCATTTTAAT
CTTCGATGATGCAGATGTTGACAAAGTTACAGAACTCGCACTGCAGGGAAGTCTATTT
AAACAGGGTGAAATATGTGTTGCTGGCTCTCGTATTTTTGTTCAAGAGGGGATTTATG
ACAAGTTTTTAATCAAGTTGAAGCAGAAGGTGAAGAATTGGGTGGTCGGTGACCCCTT
TTGATCCAACATCTCATCAAGGACCCCAAGTTGACAAGAAGCAGTACGATAGAATAC
TTTCATACATCGAGCATGGGAAGAAGGAAGGTGCCACCTTATTCCACGGTGGCAAGC
CTTGTGACGGGAAGGGCTACTACATTGAGCCAACAATATTCACTGATGTCACGGATG
AGATGACTATTGCGAAGGAAGAAATATTTGGGCCTGTGATGTCAGTATTCAAATTCA
AGACTGTTGAGGAGGCAATCAAGAGAGCAAATGCGACCAAATATGGACTGGCAGCG
GGCGTCATGACTAACAACATTAACATCGCCAACACCGTCGCAAGATCAATCCGTGCG
GGTGCTATCTGGATAAATTGTTACTTTGCCTTCGACAGAGATAGCCCTTACGGAGGCT
ACAAAATGAGTGGTTTTCGGGAGAGACATGGGAATGGATGGCCTCAAGAAGTATCTTG
CAGTTAAAGCTGTTGCTACCCCCATTTATAATTCACCTTGGCTG**GAGCAGAACTCAT**
CTCTGAAGAGGATCTGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGG
AGACGTGGAGGAGAACCCTGGACCTATGAAGGTACTAATGTTGCCATGGTTAGCTCA
TGGTCACATCTCTCCTTTCTTAGAGCTAGCCAAGAGGCTGGCAAAGAAAACTTCCAC
ATCTACCTGTGTTCCACCTCTGTCAATCTCAGCTCCATCAAGAACAAGATTACCGGAG
AATATTCCGACTCAATAGAACCAGTTGAGCTACAGCTTCCATGCTTGCCTGATCTTCC
TCCTCATTACCACACCACCAATGGCCTCCCACCCCATCTCATGACCACCTCAAAACA
GCCTATGAACTGTCAGCTCCCGACTTTTCTAATATCCTCACCGCTCTGCATCCTGACTT
GGTTGTGTACGACTTTAACCAGCCGTGGGCAGCAGAAATTGCTTCGTCTAAAAACATC
CCAGCAGTTCAATTTCTCCCGTTGGGGCAACAATGATGGCTTTCAGTCTGCACATGC
TCAAGTATTCTGGAAAAGAATCCCCCACCCAGAAATTTATATCCGGGACTACGAGA
TGCTCAAGTTTCAGTCAAGGAACGATCAAGTTAATGATGTTAGTGATAGAGAAAGAA
TCCTGCAAGCTCTGGACCTGTCTGTGAAGATTTTGTGGTCAAGTCGTTCAAAGAAAT
CGAAGAAAAATTCATGAATACTCTCTCTGTTGCATCTGGTAAAAAGGTGGTCCCCGTT
GGCCCTCTTGTCCAAGACGTCAATATCGATGACATTCAAGATGAGGAAATGGAATC
ATCCATTGGCTTGACCAAAAGGAAAATGCTTCAGTGGTGTGTTGTCTCTTTTGGTAGTG
AGTATTTCTTGACTAAGGAGGAAAGAAACGAGATTGCTCGGGGTCTTGAGCTCAGCA
ATGTTAACTTTATATGGGTAATTAGGTTCCCTCTGGGAGGAAAAATAACTATGGAAGA
AGCATTGCCCCAAGGTTTTCTTGAGAGAGTGGGAGATAGAGGAAAAATGTGGATGG
ATGGGCTCCTCAGGCAAGAATTTTGAAGCATGCGAATACTGGTGCTTTTTTGAGCCAT
TGTGGATGGAGTTCAATGATGGAGAGTATGAAATTTGGTGTTCCAATAATAGCTATGC
CTATGAGTGTTGATCAACCTGTGAATGCTAGATTGATTGAAGCAGTTGGAGTGGGATT
GGAGCCCTTGAGAGACGAGAAGGGGAATCTTCAAAGTGCGGAGATTGCCAAAGTGA
TTAGAAAGGTTTTTGGTGGACGAAAGTGGAGAAAAATGTGAGGAGGAAGGCTAAGGAA
TTGAGTGAACAAATGGAATGAGAGGTGACGAGGAGGAGATTGATAACCTTGTAGA
GGAGCTGCTGCAACTCTGCCGAAAGAACAACGGGGGCTGTGATTACAAGGATGACGA
CGATAAGTAA**ATATGAAGATGAAGATGAAATATTTGGTGTGTCAAATAAAAAGTTGTGTGCT**
TAAGTTGTGTTTTTTCTTGGCTTGTGTGTTATGAATTTGTGGCTTTTTCTAATATTAAATGAATG
TAACATCTCATTATAATGAATAAACAAATGTTTCTATAATCCATTGTGAATGTTTGTGGATCTCT
TCTCCAGCATATACTACTGTATGTGCTATGGTATGGACTATGGAATATGATTAAAGATAAG

GTCGACGAGTCAGTAATAAACGGCGTCAAAGTGCGTTGCAGCCGGCACACACGAGTCGTGT
TTATCAACTCAAAGCACAAATACTTTTCCTCAACCTAAAAATAAGGCAATTAGCCAAAAA
CAACTTTGCGTGTAACAACGCTCAATACACGTGTCATTTATTATTAGCTATTGCTTCACC
GCCTTAGCTTTCTCGTGACCTAGTCGTCTCTTTCTTCTTCTTCTTCTATAAAAACAATA
CCCAAAGAGCTCTTCTTCTCACAATTCAGATTTCATTTCTCAAAATCTTAAAAACTTTCT
CTCAATTCTCTTACCGTGATCAAGGTAAATTTCTGTGTTCCCTTATTCTCTCAAAATCTTCG
ATTTGTGTTTTGTTTCGATCCCAATTCGTATATGTTCTTTGGTTTAGATTCTGTAAATCTTAG
ATCGAAGACGATTTTCTGGGTTTGATCGTTAGATATCATCTTAATTCGATTAGGGTTTCA
TAGATATCATCCGATTTGTTCAAATAATTTGAGTTTGTGCGAATAATTACTCTTCGATTTGT
GATTTCTATCTAGATCTGGTGTTAGTTTCTAGTTTGTGCGATCGAATTGTCGATTAATCTG
AGTTTTTCTGATTAACAGATGGATGCTTTCTCTTCCTCTTTCTCTCTCGACTATCCCAA
ATCCCAAACCTTCCGACTTCTCTAACCAACCCTTCATCACCAAATATTCCTTTACTACACA
TTTCATCCGCTAGAATTGAAGATAAGCCACAACCAACAATTATTACTAGCACCCCCATC
ACTGCCACAACGCCACCATCATCTAATTCTAATCCCCTTAAGAAGCAGAACCTACCGAA
ATATACACTTAAAAGATCCAATCCCAAGCCGCAATCATCATCACACCACCACCATCAC
TGCAGCTGCAAGAACCTTCAAGAAAGCCAGTAGAACCATCATTGCTCACCAAATCTT
CAATGCACTTGATAACTTCATAGACACATGCATTGATCCTCCAACCTCACCTTCTGTTGA
CCCCAAATACGTCTGGCAAACAACCTTTGTACCAGTTGATGAACTCCCCCGACCGAA
TGTGAGGTGGTGGAAGGCTTTCTTCCACCGTGCCTTGACGGTGCCTACTTCCGCAATG
GCCCAAACCCTCAGTTCGTCACTCACGGGCCCTACCACTTTTTCGATGGCGATGGAATG
ATACATTCCATCAAATCTCCGAAGGAAAAGCCACCTTCTGCAGCCGGTACGTGAAAA
CATAAAGTACATGACGGAACGGGATACGGGCTCTCCCGTGTTCCCTAAAGTATTCTCT
GGTTTCAATGGCCTAACAGCCTCCGCAGCCAGGGTTGCCCTAGCTGCTGCCAGAATGT
TAGCTGGCCAATTTAACCAGCCAATGGTTTGGCCTGGCAAACACTAGCTTGGCTCTA
GTTGGGGGGGAACTATTTCGCTTTAGGTGAACCAGACCTTCCATATGCTGTAAAATTAAC
ACAACTGGAGATTTAGTAACCGTAGGCCGTATGACTTTGATGGTAAGCTTATTGCGA
GCATGACTGCACATCCCAAATCGACCCTGAGACAGGTGAAGCTTTCGCATTACAGATA
CGCGCTAATGCCTCCATTTCTAACCTTCTTTAGAATCAACCCTGATGGTACAAAAGAAC
CAGACGTCCCGATTTTCTCCATGACACGGCCTTCATTTGTTTCATGATTTGGCCATCACA
AGAAAATACGCTATTTTCTGACATACAAATAGGGATCAACCCTCTTCAGATGATCGC
CGGCGGACCACTAGTGGGTTTCAGATCCCCGAAAAGTGCCACGACTCGGGGTTATTCCC
AGGTATGCTGCGGATGAAGCTGAGATGAAGTGTTTGTGTTCCGGGACTTAATCTTAT
GCATGCTGTCAATGCTTGGGATGAAGATGATGGTGATGGTATTGTTGTTGTAGGAAC
ATAGTTTGGCCATTGAGCATGTTCTTTCGGCGGATGGATTTGGTTCATGCATCCATGGTGA
AAATCAAGATTGATCTCAAGACTGGGATGGTTTGGAAACATCCAATTTCTACGAGGAG
TATGGAATTTGGAGTCATTAATCCGGCTTATGTTGGCAAGAAGAACAAGTATGTATATGC
AGCAGTTGTAGACCCTTTGCCAAAGGTGTCAGGAGTGGTGAAGCTAGACTTATCAGCT
TCGGCAGGTGAGCACCGCGACTGCGTCATTGCTAGCCAGCTGTATGAACCGGAGTGCT
TTGGAAGCGAACCATTTTTCGTGGCCAAAGACCCTAACAAACCCAAATGCAGATGAAGA
TGATGGTTATGTGGTGTCCTATGTCCATGATGAAAACACAGGAGAGTCAAGATTCTTAG
TGATGGACGCGAAGTCACCGAGTCTTGAAATTGTGGCTGCTGTAAAATTGCCTCAAAG
GGTTCCTTCTGGCTTTCATGGCCTCTTTGTGAGAGAAAGTGAGCTCAATAAATTGTACC
CATACGATGTTCCAGATTACGCTGGTTCAAGTGCCACTAATTTTTCTCTCTTGAAACAG
GCCGGTGACGTTGAAGAGAACCCAGGTCCAATGAATTCCAGCAAAGTTCATGTTTTAG

TTCTGCCATATCCTGCCCAGGGCCACATAAACCCCTTGCTCCAGTTCTGTAAACGTTTG
 GTGGCAAAAGGTGTCAAAACCACGTTTGTCAACACTGTCTTCGTCTCCAAGTCCATTC
 AACCAGATCCCAGCTGCTCAATTAACCTTTGAGACCATTTCAGATGGCCATGACGAAGG
 TGGCTATGCAGCAGCTGAAAGCCCAGAGGCCTATCTTGTAGAACTGGCCACCTTCGGC
 TCAAAGACGTTGGCTGATCTTATCAGGAAGCTTGAAGACAAGGGCGAACCAGTCCAG
 GCCGTGATTTATGATTCATTCCTGACATGGGCTCTTGATGTAGCTAAGCAATTTGGACTT
 GTTACAGCTGCTTTTTTCTACTCAGACTTGTGCTGTAAACAGCGTCTACTACCATGTTTAT
 CATGGACTTCTACCGCTCCCTCTGTCAGACTCACCTCTTTCACTTCCTGGATTGCCGCT
 GCTCCAACCCAAGGAACTCCGTCCTTTGTTTACCTTCCCGATTTCGTATCCTGCTTTTC
 GTCTCATGCTCGTTAATCAGTTCTCCAACGTTGATCAAGCAGATTGGGTCATCCTCAAC
 ATTTTCCACAAATTGGAGGAAGATGTGCTGAATTGGATAGCGAACTTTGGCGAGTGA
 TTACAGTCGGTCCAACGGTTCATCCATGTACTTGGATAAACGCCTTGAGGACGATAAA
 TGTTACGGAATAAATCTCTTCAAGCCAGACTCCTGCCTGTGCATGAACTGGCTAAATAA
 TCAACCAAAGGACTCGGTAGTTTATGTATCATTCGGTAGCTGGACAGAAATTGACGTAA
 AACAAGTGGAGGAAATAGCTTCAGCTTTGAAGGAAAGCGGATTCAAATTCTTGTGGGT
 CGTGAGATCATTTGAGAAGGAAAAGCTTCCAAGCAACTTCGCTGAGGAGACGTCGGA
 GAAAGGCTTGGTGGTGACATGGAGTCCACAGCTTGAGGTATTGGCACATGAATCCGTC
 GCTTGTTCGTCACTCATTGCGGGTTCAATTCAGTGCTGGAGGCGTTGAGTTTAGGGGT
 GCCGGTGGTGGCCGCGCCGCAGTGGACGGACCAGCCACCAACGCTAAGTTTCTGGA
 GGATATTTGGGGAGTTGGTGTGAAGGCTGGTGC GGATGAGAAAGGAATTGTGACGAG
 GGAAACGCTAGTTTCTGCATAAGGGAGATTATGGAGGGAGAAAAAGGAAAGCAGAT
 TAAAGAGAATGCCATTAAATGGAAGACTTTAGCCAAAGAAGCAGTTGATGAGGGCGG
 AAGTTCTGATAAGAACATTGACGAATTTGTTGCAGAATTGGCTGGCCAGAAGGCCAG
 AACCATCATCACCATCACCATTAGACTAGTTCGAGCTTTCGTTTCGTATCATCGGTTTCGACAA
 CGTTCGTCAAGTTCAATGCATCAGTTTCATTGCGCACACACCAGAATCCTACTGAGTTTGAGTAT
 TATGGCATTGGGAAAACGTGTTTTCTGTACCATTGTGTGCTTGTAAATTTACTGTGTTTTTATTC
 GGTTTTCGCTATCGAACTGTGAAATGGAAATGGATGGAGAAGAGTTAATGAATGATATGGTCCT
 TTTGTTCACTCTCAAATTAATATTATTGTTTTCTCTTATTGTTGTGTGTTGAATTTGAAATTATA
 AGAGATATGCAAACATTTTGTGTTGAGTAAAAATGTGTCAAATCGTGGCCTCTAATGACCGAAGT
 TAATATGAGGAGTAAAACACTGTAGTTGTACCATTATGCTTATTCCTAGGCAACAAATATATTT
 TCAGACCTAGAAAAGCTGCAAATGTACTGAATACAAGTATGTCCTCTTGTGTTTAGACATTTAT
 GAACTTTCCTTTATGTAATTTCCAGAATCCTTGTGAGATTCTAATCATTGCTTTATAATTATAGTT
 ATACTCATGGATTGTAGTTGAGTATGAAAATATTTTTAATGCATTTATGACTTGCCAATTGATT
 GACAACATGCATCAA

Figure S2. The nucleotide sequence of insertion fragment in the multigene vector AU-CU. The crocins biosynthesis genes *GjALDH2C3* (black), *GjUGT94E13* (blue), *GjCCD4a* (green), and *GjUGT74F8* (orange) were shown in this graph. P2A peptide sequences were highlighted in red. *GjALDH2C3*-2A-*GjUGT9F13* was controlled by CaMV 35S promoter (highlighted in yellow), and the terminator was heat-shock protein 18.2 terminator (Thsp) (highlighted in light blue). *GjCCD4a*-2A-*GjUGT74F8* was controlled by the AtUBQ10 promoter (highlighted in light green) and terminator from the pea *rbcs-E9* gene (Te9) (highlighted in gray).

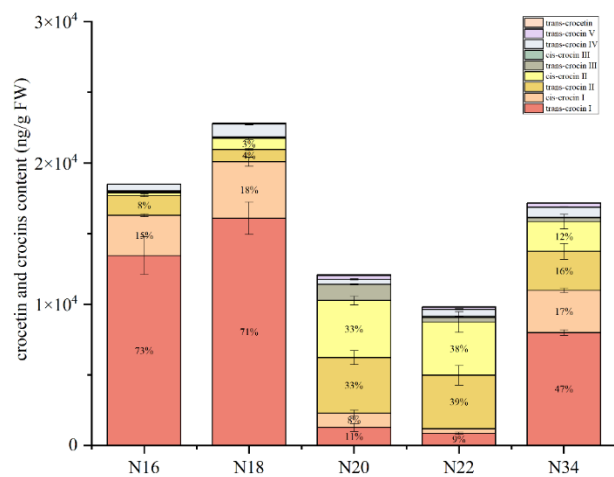


Figure S3. Crocetin and crocins production in transgenic tobacco of T1 generation lines. The data are presented as the mean values \pm SDs, $n = 3$ biologically independent samples.

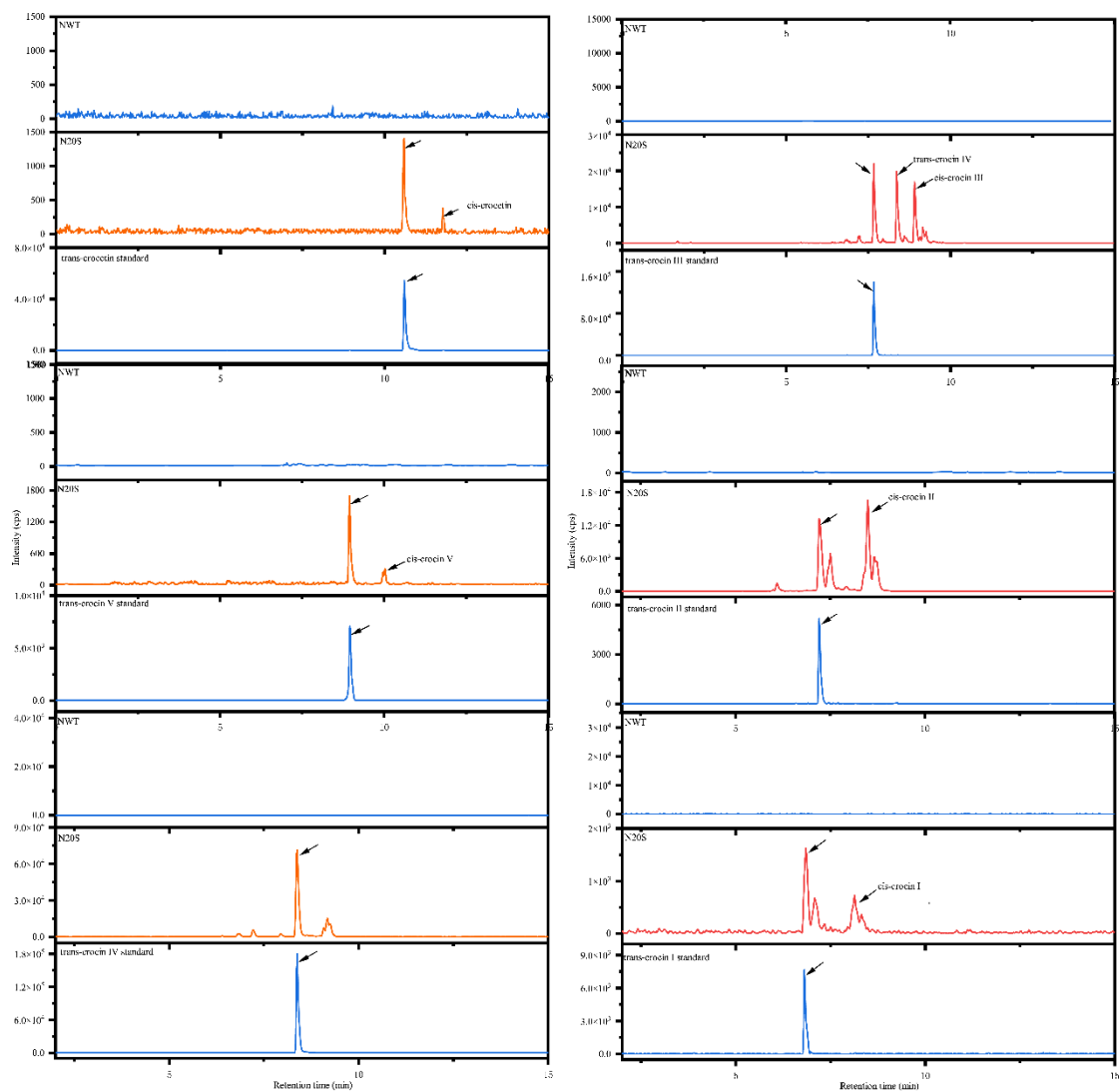


Figure S4. Identification of crocins in the stems of the transgenic tobacco N20 line.

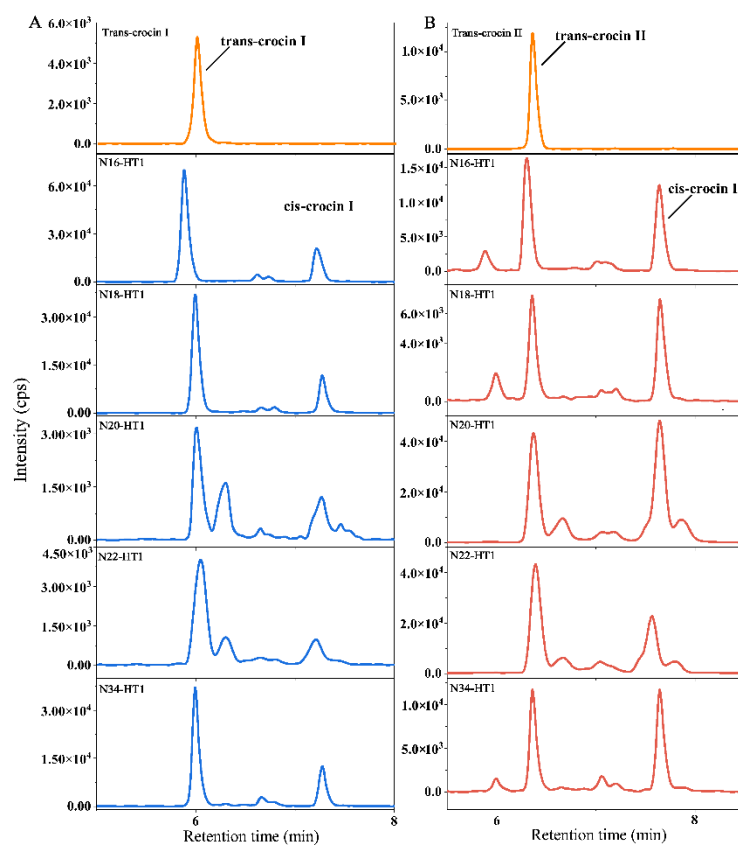


Figure S5. Analysis of crocin I and crocin II in the transgenic tobacco line N16-HT1, N18-HT1, N20-HT1, N22-HT1 and N34-HT1. (A) Ion pair 999.3/675.0 extraction chromatogram for analysis of crocin I in transgenic tobacco line. (B) Ion pair 837.3/675.0 extraction chromatogram for analysis of crocin II in transgenic tobacco line.

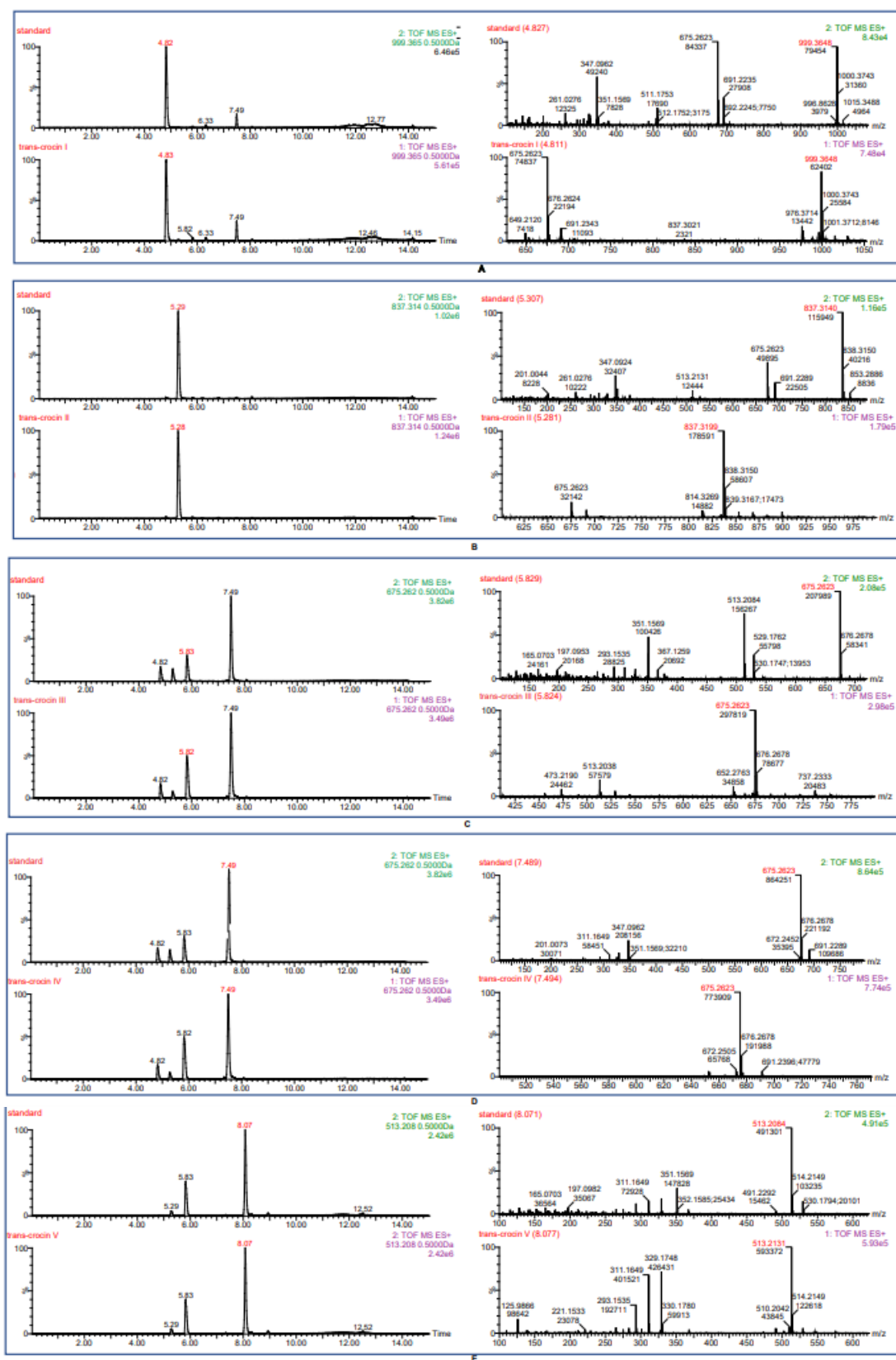


Figure S6. The UPLC-ESI-QTOF-MS/MS analysis of crocins for the standards. (A-E) The ion chromatogram, the MS and MS/MS spectra of crocin V-I.

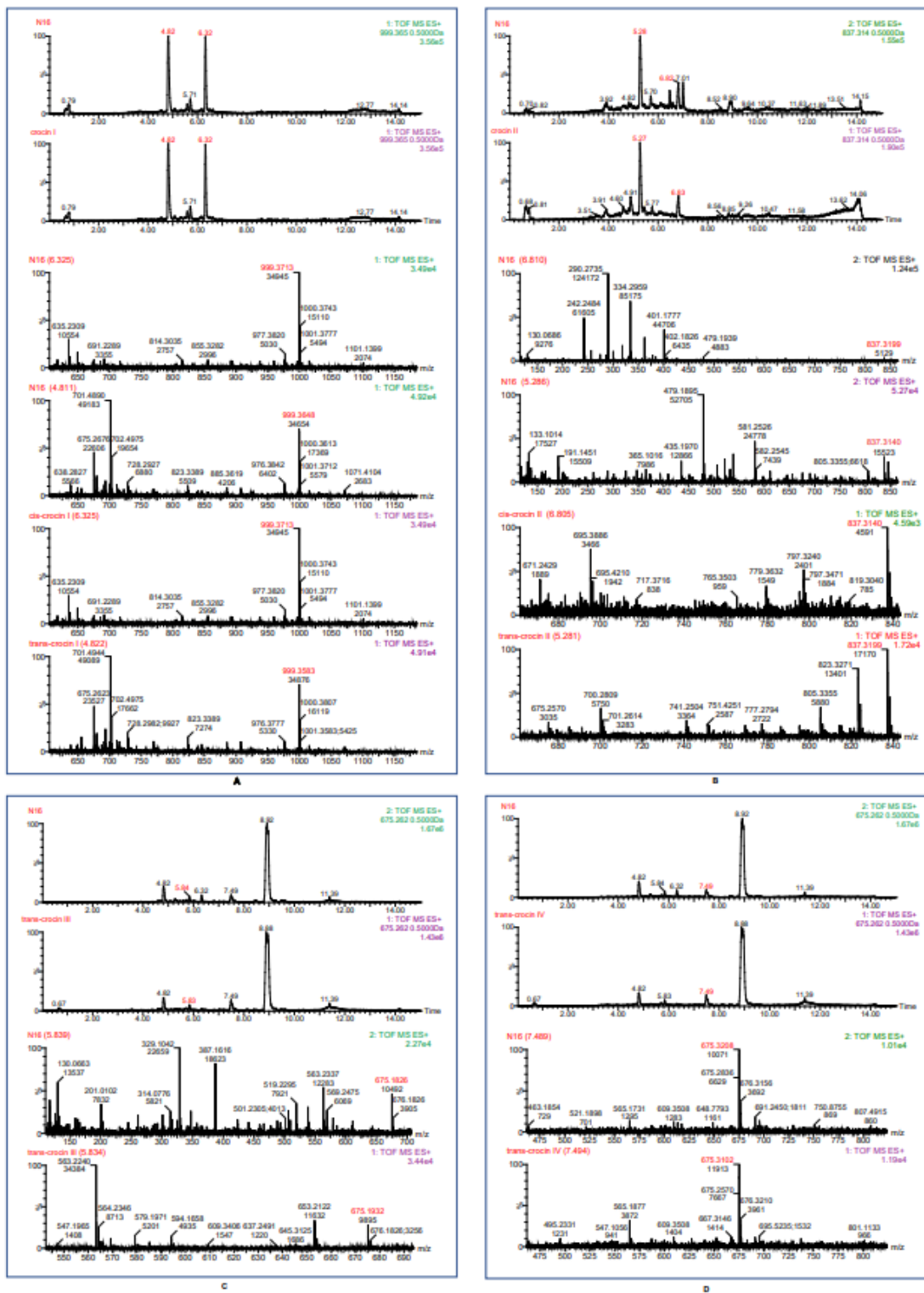
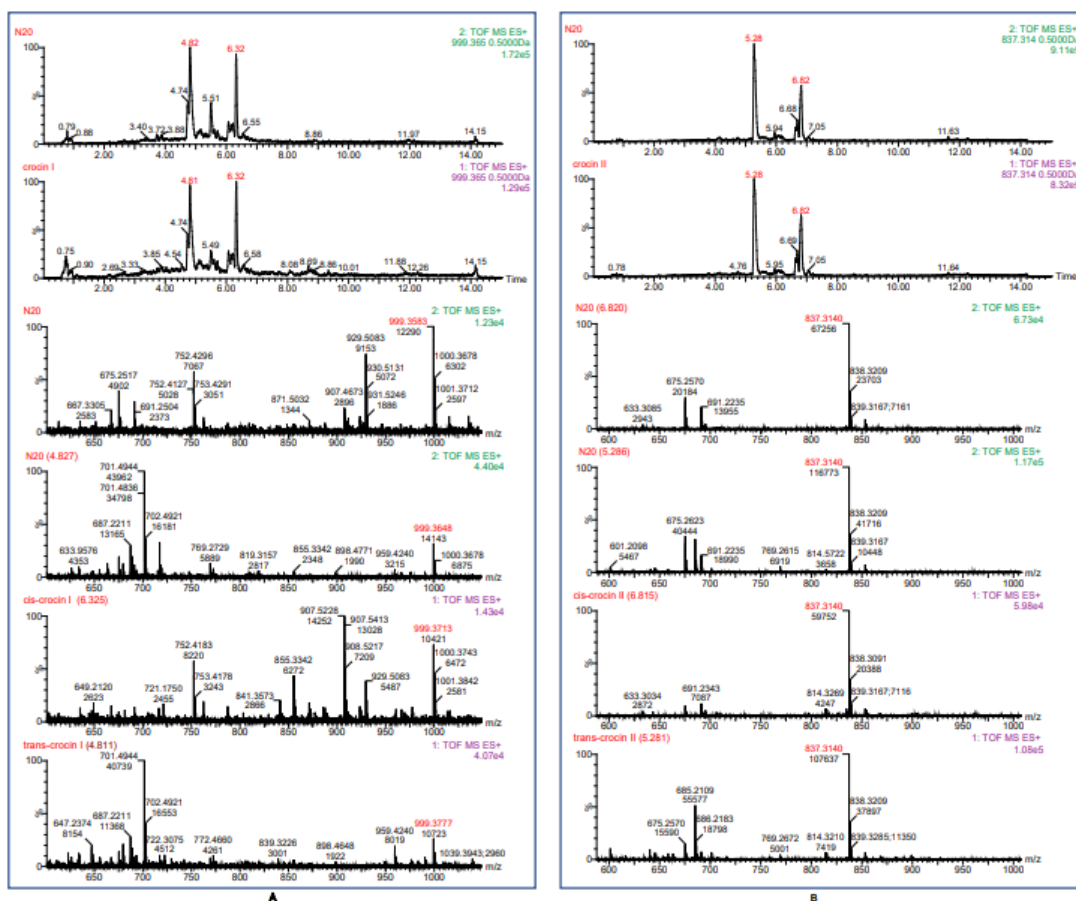


Figure S7. The UPLC-ESI-QTOF-MS/MS analysis of crocins for the N16 line. (A-D) The ion chromatogram, the MS and MS/MS spectra of crocin IV, crocin III, crocin II, and crocin I.

Figure S8. The UPLC-ESI-QTOF-MS/MS analysis of crocins for the N18 line. (A-D) The ion chromatogram, the MS and MS/MS spectra of crocin IV, crocin III, crocin II, and crocin I.



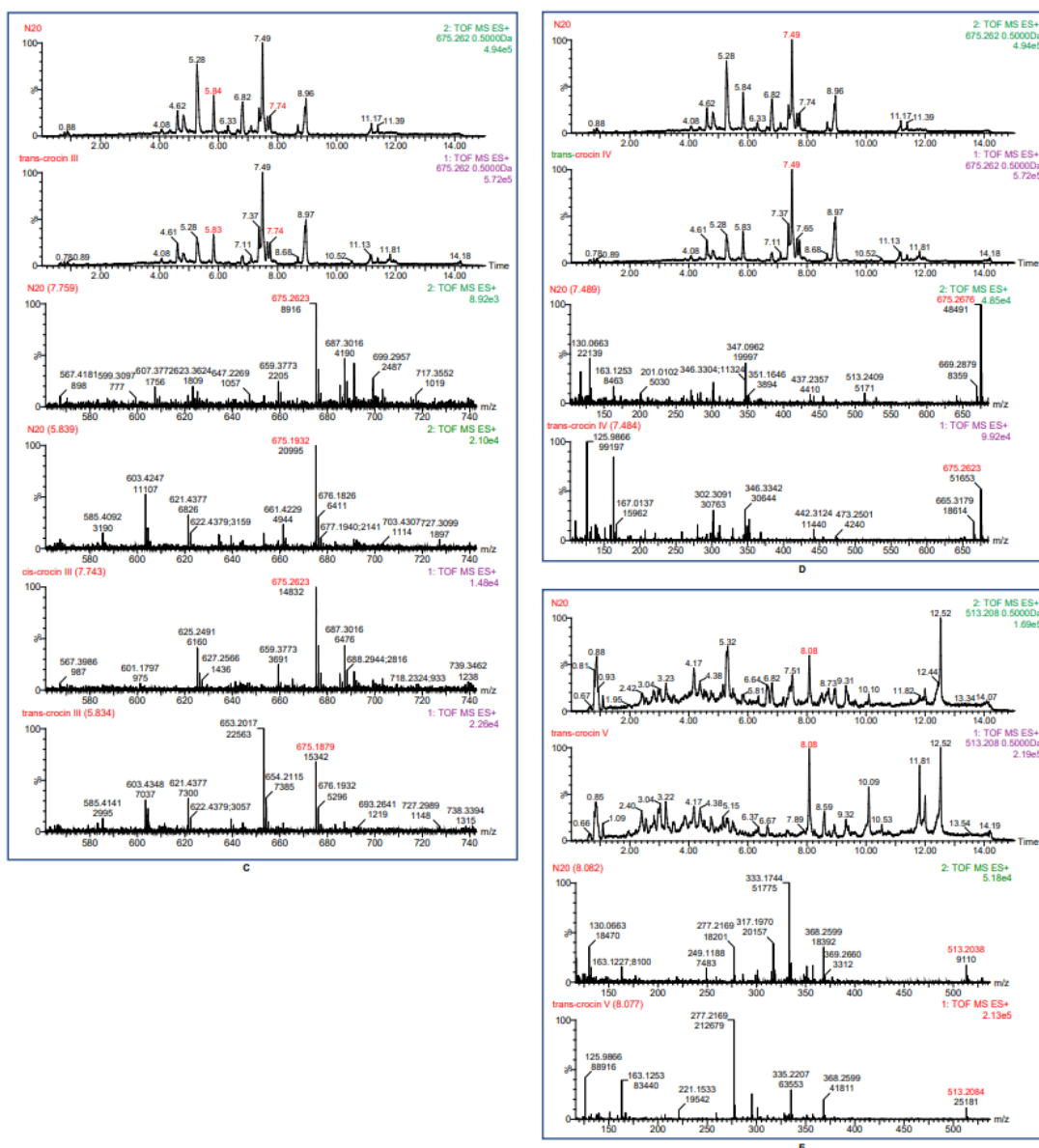


Figure S9. The UPLC-ESI-QTOF-MS/MS analysis of crocins for the N20 line. (A-E) The ion chromatogram, the MS and MS/MS spectra of crocin V, crocin IV, crocin III, crocin II, and crocin I.

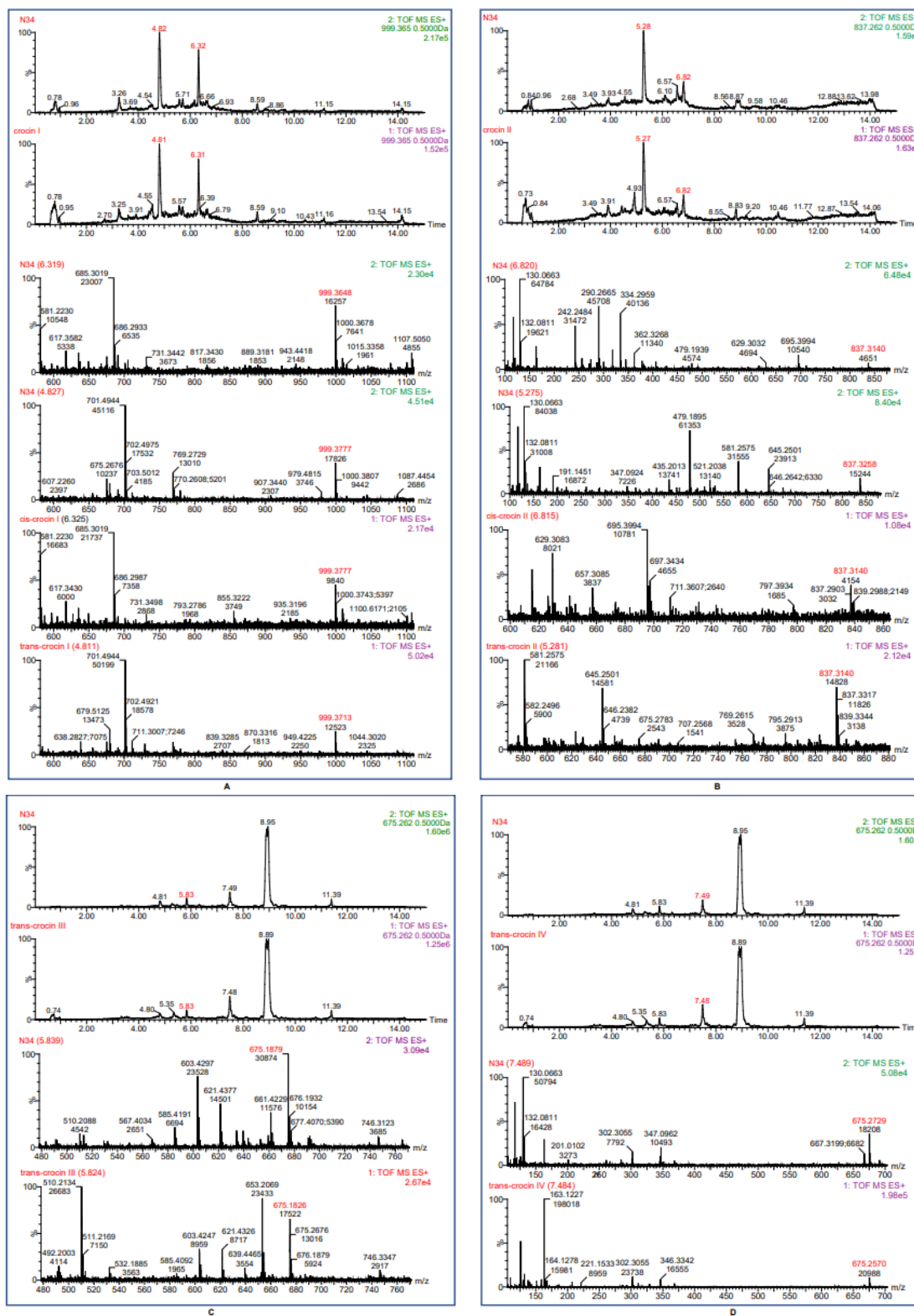


Figure S11. The UPLC-ESI-QTOF-MS/MS analysis of crocins for the N34 line. (A-D) The ion chromatogram, the MS and MS/MS spectra of crocin IV, crocin III, crocin II, and crocin I.

Primers	Sequences (5' to 3')
<i>GjCCD4a-F</i>	ATGACTGATTTATTGTGTTTCCTATACG
<i>GjCCD4a-R</i>	TTACAATTTATTGAGCTCACTTTCTCTG
<i>GjALDH2C3-F</i>	ATGGCCGTCCAAAGCAACG
<i>GjALDH2C3-R</i>	TTACAGCCAAGGTGAATTATAAATGG
<i>GjUGT74F8-F</i>	ATGAATTCCAGCAAAGTTCATGTT
<i>GjUGT74F8-R</i>	CTAGTTCTGGGCCTTCTGGCC
<i>GjUGT94E13-F</i>	ATGAAGGTACTAATGTTGCCATGG
<i>GjUGT94E13-R</i>	TTAACAGCCCCCGTTGTTCTT

Table S1. Primers for PCR detection

Primers	Sequences (5' to 3')
<i>GjCCD4a-qF</i>	CAAGATTGATCTCAAGACTGG
<i>GjCCD4a-qR</i>	GTCTACAAGTCTGCATATAC
<i>GjALDH2C3-qF</i>	GGATTCTCATGTTAAGGTTCC
<i>GjALDH2C3-qR</i>	GGATCTATTGTCTCAAACGTC
<i>GjUGT94E13-qF</i>	CAATAGAACCAGTTGAGCTAC
<i>GjUGT94E13-qR</i>	GATATTAGAAAAGTCGGGAGC
<i>GjUGT74F8-qF</i>	CAAAGGACTCGGTAGTTTATG
<i>GjUGT74F8-qR</i>	CTTCTCAAATGATCTCACGAC
<i>NbActin-qF</i>	TATTCCTAGTATTGTTGGC
<i>NbActin-qR</i>	CTGGGGTATTAAAAGTCTCA

Table S2. Primers for qRT-PC